Sparking curiosity or tipping the scales? Targeted advertising to rationally inattentive consumers*

Andrei Matveenko[1] and Egor Starkov †[1]

[1]University of Copenhagen, Department of Economics

PRELIMINARY AND INCOMPLETE, PLEASE DO NOT DISTRIBUTE
January 31, 2020

Abstract

A monopolist has an opportunity to launch an advertising campaign and chooses a targeting strategy, i.e. to which consumers to send its advertisement. The consumers differ in their initial opinions about the product and are rationally inattentive in that they must incur a cost if they want to learn their value for the product. We discover that the firm generally prefers to target consumers who are either indifferent between ignoring and investigating the product, or between investigating and buying it unconditionally. If the firm is uncertain about the consumer appeal of its product, it targets these two disjoint groups of consumers simultaneously.

JEL-codes: D83, L15, M37.

Keywords: advertising, targeting, rational inattention.

*We are grateful to Mogens Fosgerau and Adrien Vigier, as well as seminar participants at the University of Copenhagen for helpful comments and valuable feedback. This project has received funding from the European Research Council under the European Union's Horizon 2020 research and innovation programme (grant agreement No 740369).

†The authors' contributions are equal and the order is alphabetical.

E-mail addresses: `anma@econ.ku.dk` (A. Matveenko), `egor.starkov@econ.ku.dk` (E. Starkov).

1 Introduction

The recent decade has witnessed a flood of website analytics systems and smartphone apps that collect detailed information on users and enable firms to track consumers' tastes and actions with ever increasing precision. The welfare implications of such information collection are still not perfectly clear. The economic literature has largely focused on the implications of larger scope for price discrimination allowed for by the ever-growing amounts of information available to the firms. However, much less attention has been devoted to what is arguably the more widespread use of personal information by firms, namely personalized advertising.

Knowledge of the consumers' tastes allows the firm to target the advertising towards audiences that are most easily manipulated into purchasing the product, thus generating higher return per dollar spent on advertising. The value added by targeting can be quite substantial, as illustrated (albeit in a political, rather than economic context) by the story of Cambridge Analytica, a now-infamous company that has arguably played an important role in the outcomes of the 2016 US Presidential Election and the UK Brexit referendum by influencing voters via highly-targeted political advertising on Facebook.[1]

We build a theoretical model designed to explore the trade-offs of targeted advertising. In this model a population of rationally inattentive consumers is faced with an option to buy a product, but the payoff which this product brings is unknown to the consumer. The consumers differ in their prior belief regarding the product's payoff, and they can acquire costly information about it. The firm has an option of sending an ad to a subset of consumers, improving their belief about the product, and the firm can select freely which consumers will receive the ad.

We discover that the optimal advertising strategy is not straightforward. In particular, if the advertisement is not too powerful, while the cost of advertising is significant, then it is optimal for the firm to target consumers in two disjoint groups. The first group is the pessimistic consumers who ignore the product without investigating it – sending an ad to these consumers will render them interested enough to acquire more information about the product, which converts into sales. The second group is the optimistic consumers who are close to buying the product but acquire a little bit of information "just to be sure" – advertising to these consumers will "tip the scales" and convince them to buy the product without investigation. The relative focus on the two groups depends on the firm's belief about its own product. Curiously, if the firm is uncertain enough then it will advertise to both groups simultaneously, while ignoring consumers with average beliefs, who choose to investigate the product regardless of advertising. It is more valuable for the firm to affect the consumers' information acquisition decisions at the extensive margin (whether a consumer acquires any information or not) than at the intensive margin (how much

[1] See, for instance, `https://www.theguardian.com/uk-news/2018/mar/23/leaked-cambridge-analyticas-blueprint-for-trump-victory`

information a consumer acquires).

The remainder of this paper is organized as follows. Section 2 contains a review of the relevant literature. Section 3 describes the model, and Section 4 derives the optimal advertising strategy and explores its comparative statics. Section 5 concludes.

2 Literature Review

Goldfarb and Tucker [2019] present an excellent survey of works on digital economics, see chapter 6 for survey of literature dealing with the implications of improved consumer tracking technologies.

One of the seminal papers on targeting is that by Iyer, Soberman, and Villas-Boas [2005]. They model advertising as generating awareness of the product, and obtain that in the presence of competition, firms prefer to target consumers with a strong preference for their product, rather than those close to indifference. In the end, two firms in the market target their advertising to non-overlapping populations and do not fight for the "median consumer". We show that with informative advertising and a subsequent information acquisition stage by the consumers, even a single monopolistic firm would cover similar groups (fans and/or haters but not the undecided ones), although for very different reasons.

Subsequent work on targeting includes papers by Athey and Gans [2010], Bergemann and Bonatti [2011], Farahat and Bailey [2012], Chen and Stallaert [2014], and Anderson, Baik, and Larson [2019]. These papers explore (theoretically and in the field) the effects of the technology that enables targeting and, equivalently, of the legislation which limits it, such as GDPR in Europe. They find multiple surprising results pertaining to the price and amounts of advertising, as well as its social value. However, none of these papers account for the possibility that consumers acquire information independently, which, as we show, can drive predictions to quite a significant extent.

Literature on strategic information acquisition in various settings is vast, going back at least to the sequential sampling model of Wald [1945]. As of lately, it has also been closely intertwined with literature on rational inattention, pioneered by Sims (see Sims [2010] and Mackowiak, Matejka, and Wiederholt [2018] for recent surveys). Hébert and Woodford [2017] and Morris and Strack [2017] show that under certain conditions, the Wald problem can be represented as a static problem with information cost given by mutual information, as in our model. Matějka and McKay [2015] show that rationally inattentive consumers exhibit logit behavior from the observer's point of view. We employ some of the analytical tools developed in that paper.

Closest to ours is the paper by Matysková [2018], who characterizes an optimal Bayesian Persuasion mechanism subject to the constraint that the receiver may engage in costly information acquisition after hearing the sender's message. Our model is different in that our sender (the firm) is constrained in which messages it can send, and doing so is costly.

Relatedly, Jerath and Ren [2019] explore the firm's incentives to limit the amount of information consumers can potentially acquire about the product. Gossner, Steiner, and Stewart [2018] demonstrate in a rational inattention setting that attracting consumer's attention to own product is always beneficial for the firm. Bloedel and Segal [2018] study a general Bayesian Persuasion problem with a rationally inattentive receiver who must exert a cost to understand the sender's message. Jain and Whitmeyer [2019] look at a similar model with competing senders; they find that competition encourages disclosure.

There also exists an extremely rich literature on Bayesian Persuasion and information design without the information acquisition layer, particularly in application to bilateral trade. These papers analyze and compare various information structures within the context of a buyer-seller interaction. For some examples see Bergemann, Brooks, and Morris [2013], Roesler and Szentes [2017], Kolotilin, Mylovanov, Zapechelnyuk, and Li [2017], and Condorelli and Szentes [2020]. A strand of this literature has focused, in particular, on the possible effects of the improvements in tracking technology that allow the firm to better price discriminate among consumers with different tastes. Some examples include Bergemann, Bonatti, and Smolin [2018], Preuss [2018], Carroni, Ferrari, and Righi [2019], and Ichihashi [2019].

3 The Model

3.1 The Setup

The market consists of a single firm and a continuum of consumers $\mathcal{I}$ with unit demand. The firm offers for sale a single product of unknown quality $s \in S = \{H, L\}$. All consumers' valuation of a high-quality product is $v = 1$, while a low-quality product is valued at $v = 0$. Consumer $i \in \mathcal{I}$ initially believes that the quality is high with probability $g_i \in (0, 1)$.[2] The firm believes the quality of its product is high ($s = H$) with probability $g_{f,0}$. The price of the product is R.

The firm can advertise its product. In particular, it has a private signal $y_i \in \{h, l\}$ distributed conditional on quality according to $\mathbb{P}(h|H) = \mathbb{P}(l|L) = \rho > 1/2$. The firm can verifiably disclose this signal in an ad, and it chooses a set $\mathcal{T} \subseteq \mathcal{I}$ of consumers to whom to send the ad; the cost of such advertising campaign depends on its size and is given by $c \cdot |\mathcal{T}|$ for some per-person cost $c > 0$.

Each consumer chooses whether to purchase the item at a given price or not. Prior to making the decision (but after receiving the ad), the consumer has an opportunity to receive a signal $x_i \in \mathbb{R}$, which is informative about the product quality. The consumer can also choose the distribution of signals, $f(x_i, s) \in \Delta(\mathbb{R} \times S)$, where $\Delta(\mathbb{R} \times S)$ is the set of all probability distributions on $\mathbb{R} \times S$. Upon receiving a signal, each consumer updates her

[2] We also use alternative notation for beliefs, $g_i(s)$, understanding it as $g_i(H) = g_i$ and $g_i(L) = 1 - g_i$.

belief using Bayes rule. However, generating a signal is costly, and we assume the cost to be equal to $\lambda\kappa(f)$, where

$$\kappa(f) \equiv -\sum_s \mathbb{P}(s)\log(\mathbb{P}(s)) + \sum_s \sum_x \mathbb{P}(s|x)\log(\mathbb{P}(s|x))$$

is the expected reduction in entropy between the consumer's prior and posterior beliefs, and λ is a cost factor.[3] Given the updated belief, the consumer decides whether to purchase the product at price R so as to maximize her expected payoff.

The overall timing of the model is as follows:

1. the firm observes y and chooses $\mathcal{T}$;
2. the targeted consumers receive ads and update their prior beliefs g_i;
3. the consumers select their information acquisition strategy $f(x_i, s)$;
4. the consumers observe their signals x_i and update their posterior beliefs $\tilde{g}_i$;
5. the consumers decide whether to purchase the product given $\tilde{g}_i$;
6. payoffs are realized.

Note that the consumers' purchasing decision is mechanical: consumer i buys the product if and only if $\tilde{g}_i \geq R$ (break ties in firm's favor for simplicity). Therefore, we will henceforth take this purchasing strategy as given and focus on other strategic layers of the game.

Let $\mathcal{P}(g_i)(s)$ denote the probability with which consumer i purchases the item (conditional on her information acquisition strategy and the firm's advertising strategy) when the true quality is s. Let $\mathcal{P}(g_i)$ denote the respective unconditional probability from the consumer's perspective:

$$\mathcal{P}(g_i) = \sum_{s \in S} g_i(s) \cdot \mathcal{P}(g_i)(s),$$

and $\mathcal{P}^f(g_i)$ represent the probability that consumer i buys the product as perceived by the firm. Given these probabilities, we can define the players' payoffs. In particular, consumer i's expected utility from information acquisition strategy is given by

$$g_i(1-R) \cdot \mathcal{P}(g_i)(H) + (1-g_i)(-R) \cdot \mathcal{P}(g_i)(L) - \lambda\kappa(f), \tag{1}$$

while the firm's expected profit is given by

$$R \cdot \int_{i \in \mathcal{I}} \mathcal{P}^f(g_i) di - c \cdot |\mathcal{T}|. \tag{2}$$

We will be looking for a Perfect Bayesian Equilibrium of the game which consists of the firm's advertising strategy $\mathcal{T}$ and the consumers' information acquisition strategy f such that:

[3] We assume (w.l.o.g. given the remainder of the model) that f has finite support. For a detailed treatment of entropy, see, for example, Cover and Thomas [2012].

1. information acquisition strategy $f(x_i, s)$ maximizes consumer i's expected payoff (1);
2. advertising strategy $\mathcal{T}(y)$ maximizes the firm's expected profit (2) given signal y;
3. beliefs are updated using Bayes' rule whenever possible.

4 Analysis

4.1 Consumers' Problem

According to Corollary 1 from Matějka and McKay [2015], the optimal choice behavior of the consumer in such setting can be found as a solution to a simpler maximization problem that is stated in terms of state-contingent choice probabilities alone. In particular, instead of choosing an information acquisition strategy f, consumer i with prior g_i (possibly influenced by the firm) can maximize over the state-contingent choice probabilities $\mathcal{P}(g_i)(s)$ that this strategy induces. That is, the consumer's problem is:

$$\max_{\{\mathcal{P}(g_i)(s)|s\in S\}} \{g_i(1-R)\cdot\mathcal{P}(g_i)(H) + (1-g_i)(-R)\cdot\mathcal{P}(g_i)(L) - \lambda\kappa(g_i)\} \tag{3}$$

$$\text{subject to: } \mathcal{P}(g_i)(s) \in [0,1] \qquad \forall s \in S, \tag{4}$$

where $\kappa(g_i)$ denotes the expected reduction in entropy between the prior and the posterior beliefs:

$$\begin{aligned}\kappa(g_i) =& \underbrace{-\left[\mathcal{P}(g_i)\log\mathcal{P}(g_i) + (1-\mathcal{P}(g_i))\log\left(1-\mathcal{P}(g_i)\right)\right]}_{\text{prior uncertainty}} - \\ &-\sum_{s\in S}\left(\underbrace{-\left[\mathcal{P}(g_i)(s)\log\mathcal{P}(g_i)(s) + (1-\mathcal{P}(g_i)(s))\log\left(1-\mathcal{P}(g_i)(s)\right)\right]}_{\text{posterior uncertainty in state } s} g_i(s)\right).\end{aligned} \tag{5}$$

Corollary 2 in Matějka and McKay [2015] provides the normalization condition which we can use to find $\mathcal{P}(g_i)$:

$$g_i\frac{e^{\frac{1-R}{\lambda}}}{\mathcal{P}(g_i)e^{\frac{1-R}{\lambda}} + (1-\mathcal{P}(g_i))} + (1-g_i)\frac{e^{-\frac{R}{\lambda}}}{\mathcal{P}(g_i)e^{-\frac{R}{\lambda}} + (1-\mathcal{P}(g_i))} = 1.$$

Solving for $\mathcal{P}(g_i)$ yields the interior solution for $\mathcal{P}(g_i)$:

$$\mathcal{P}(g_i) = \frac{g_i\left(e^{\frac{1}{\lambda}}-1\right) - \left(e^{\frac{R}{\lambda}}-1\right)}{\left(e^{\frac{1}{\lambda}}-e^{\frac{R}{\lambda}}\right)\left(1-e^{-\frac{R}{\lambda}}\right)} \cap [0,1].$$

4.2 Firm's Problem

The firm can only use hard information in its advertising in our model, meaning that given a private signal y, the firm can disclose y to consumers in an ad or stay silent, but cannot modify it or send any other messages. Therefore, consumer i's interim belief upon receiving an ad with a good signal is

$$\alpha(g_i) \equiv \frac{g_i \rho}{g_i \rho + (1 - g_i)(1 - \rho)}, \tag{6}$$

while upon hearing an ad with a bad signal, the belief is

$$\beta(g_i) \equiv \frac{g_i(1 - \rho)}{g_i(1 - \rho) + (1 - g_i)\rho}. \tag{7}$$

Since advertising with a bad signal is costly and depresses the consumer's belief, it is never optimal for the firm to do so. Hence we assume that the firm with a bad signal never advertises ($\mathcal{T}(l) = \varnothing$), and from this point onwards we only consider the problem of a firm with a good private signal. One implication of this is that in equilibrium, if consumer i expects to be targeted by a firm with a good signal but does not receive any ad, he infers that the firm's private signal is bad, and his belief drops to $\beta(g_i)$.

We begin by solving a version of the problem with *cursed* consumers which react to news, but not to the lack thereof. In other words, if consumer i receives the ad then his belief increases to $\alpha(g_i)$, but if consumer i hears nothing then his belief remains at g_i.[4] Note that the benefit from advertising is smaller when consumers are cursed, since they do not penalize silence as fully sophisticated consumers do. Therefore, if advertising to a given consumer i is optimal when i is cursed, advertising is also optimal when i is sophisticated.

Given the consumer's information acquisition strategy, the firm can compute sales probabilities. In particular, the firm can compute $\mathcal{P}^f(g_i)$, the expected probability that consumer i will buy the product without (before) the ad, and $\mathcal{P}^f(\alpha(g_i))$, the corresponding probability after the consumer is exposed to the ad. The firm's expected profit from targeting set $\mathcal{T}$ of consumers is given by (2), which can then be rewritten as

$$R \cdot \left[\int_{i \in \mathcal{T}} \mathcal{P}^f(\alpha(g_i)) di + \int_{i \in \mathcal{I} \setminus \mathcal{T}} \mathcal{P}^f(g_i) di \right] - c \cdot |\mathcal{T}|,$$

[4]Notion of cursedness for general games was proposed by Eyster and Rabin [2005]. Empirical evidence of cursedness of consumers in various settings has been obtained by Li and Hitt [2008] and Brown, Camerer, and Lovallo [2012] in the field, and Jin, Luca, and Martin [2018] and Deversi, Ispano, and Schwardmann [2019] in the lab. Markets with cursed consumers have been explored by Matysková and Šípek [2017] and Ispano and Schwardmann [2018].

where the purchasing probabilities are given by

$$\mathcal{P}^f(g_i) = g_f \frac{\mathcal{P}(g_i)e^{\frac{1-R}{\lambda}}}{\mathcal{P}(g_i)e^{\frac{1-R}{\lambda}} + (1-\mathcal{P}(g_i))} + (1-g_f)\frac{\mathcal{P}(g_i)e^{-\frac{R}{\lambda}}}{\mathcal{P}(g_i)e^{-\frac{R}{\lambda}} + (1-\mathcal{P}(g_i))}.$$

Note that the firm's belief g_f here incorporates both the prior belief $g_{f,0}$ and the good private signal.

Therefore, the firm will find it optimal to target consumer $i \in \mathcal{I}$ if and only if $\mathcal{A}(g_i) \geq c$, where

$$\mathcal{A}(g_i) \equiv \mathcal{P}^f(\alpha(g_i)) - \mathcal{P}^f(g_i)$$

denotes the ad effect depending on the consumer's original prior. Therefore, due to linearity of costs, instead of choosing which consumers $\mathcal{T} \subseteq \mathcal{I}$ to target, the firm effectively chooses which consumers' priors $\mathcal{T}_f \equiv \{g_i | i \in \mathcal{T}\}$ to target. According to the above, the firm's optimal advertising strategy conditional on good private signal is given by $\mathcal{T}_f(y) = \{g | \mathcal{A}(g) > c\}$ when consumers are cursed (assuming the firm does not advertise when indifferent).

Going back to sophisticated consumers, we claim that $\mathcal{T}_f(y)$ as defined above constitutes an equilibrium. Indeed, for any $g \in \mathcal{T}_f(y)$ the actual ad effect is

$$\hat{\mathcal{A}}(g) \equiv \mathcal{P}^f(\alpha(g)) - \mathcal{P}^f(\beta(g)) > \mathcal{A}(g),$$

hence advertising to this consumer is still optimal. Conversely, for any $g \notin \mathcal{T}_f(y)$ the ad effect is still given by $\mathcal{A}(g) \leq c$ (consumer does not expect to receive an ad, hence does not update his belief after not receiving one), so advertising to this consumer is not optimal. Hereinafter we refer to this as the *cursed equilibrium.* The following sections illustrate and explore this equilibrium in greater detail; section **??** looks at other equilibria of the game and justifies the selectino of the cursed equilibrium as the one with the least amount of advertising.

4.3 Illustrative Example

We begin by looking at a particular example and state the general result after that.

Figure 1 depicts the expected purchasing probabilities $\mathcal{P}^f(\alpha(g_i))$ and $\mathcal{P}^f(g_i)$. The effect of the ads on purchasing probabilities of cursed consumers (i.e., the difference between the two probabilities) as a function of a consumer's original prior g_i is depicted in Figure 2. Notably, it has two peaks, so if advertising cost c is moderately high then it is optimal for the firm to target two disjoint groups of consumers.

The right peak in Figure 2 is at the consumer who is indifferent between investigating the product and buying it immediately. Consumers with slightly worse priors conduct their investigations into the product (however uninformative) for sure, so the firm faces a risk

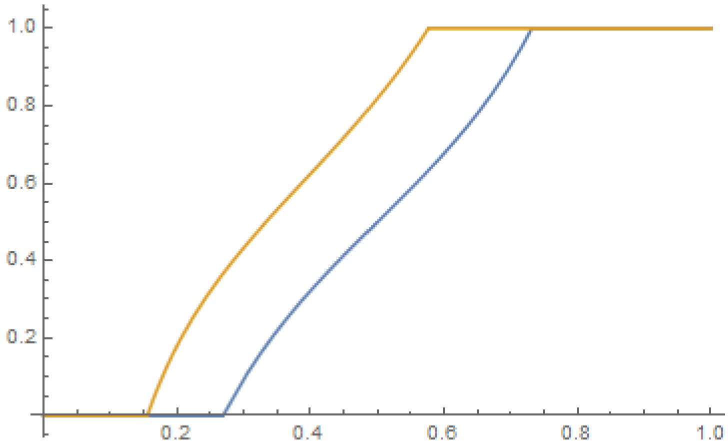


Figure 1: Purchasing probabilities before and after ad, as expected by the firm: $\mathcal{P}^f(g_i)$ (blue) and $\mathcal{P}^f(\alpha(g_i))$ (yellow). $\lambda = 1$, $g_f = \frac{1}{2}$, $q = 2$, $R = 0.5$.

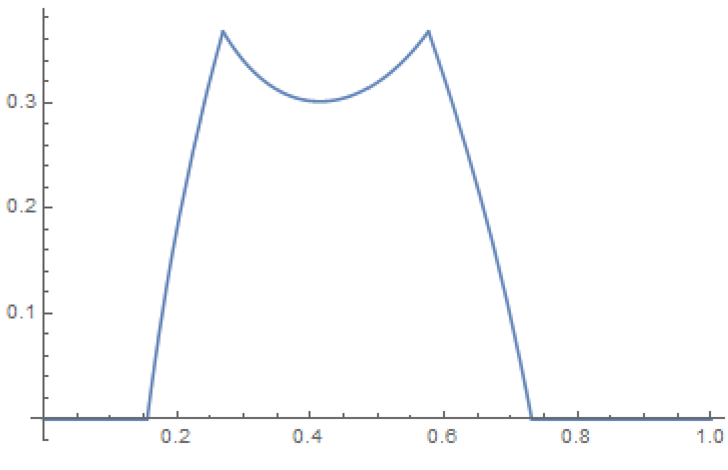
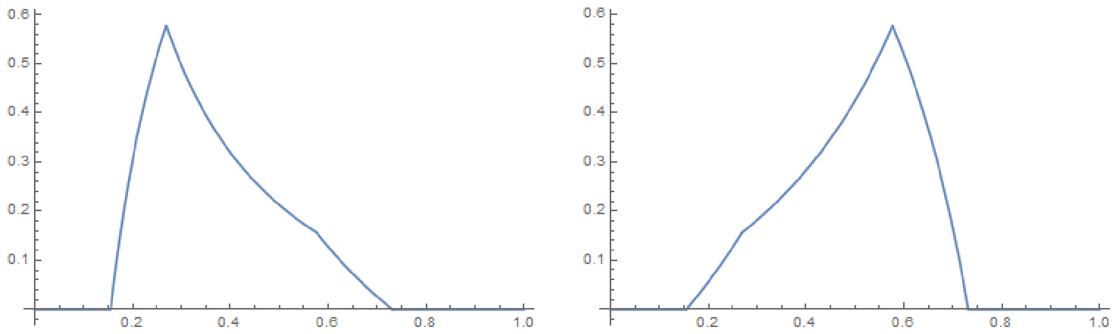


Figure 2: Ad effect $\mathcal{A}(g_i)$. $\lambda = 1$, $g_f = \frac{1}{2}$, $\rho = 2/3$, $R = 0.5$.



(a) $g_f = 1$

(b) $g_f = 0$

Figure 3: Ad effect $\mathcal{A}(g_i)$. $\lambda = 1$, $\rho = 2/3$, $R = 0.5$.

of these consumers arriving to the conclusion that the product is bad and not making a purchase. The firm thus benefits greatly from advertising to these consumers, since this eliminates the investigation risk and nudges these consumers to buy the product with no further deliberations.

The left peak is at the consumer who is indifferent between investigating the product and walking by. By advertising to this consumer, the firm does not automatically generate a sale, but rather sparks the consumer's interest and leads him to investigate the product rather than simply walk past it.

It is quite surprising that the firm prefers to primarily target two disjoint groups of consumers, while consumers who investigate the product regardless of seeing the ad are less profitable to advertise to. Figure 3 demonstrates that this is due to the firm being not sure of the quality of its own product ($g_f = \frac{1}{2}$ in Figure 2).

In particular, if the firm is sure that its product is surely good ($g_f = 1$) or surely bad ($g_f = 0$) then its targeting strategy is monotone in c. The firm with a good product prefers to target consumers who would otherwise either not even look at the product, or not put much effort into investigating its quality. Such a firm realizes that any consumer that investigates its product carefully enough will realize that it is worth buying. Hence the firm tries to induce the consumers to investigate its product carefully.

A firm with a bad product has the opposite incentives: it would like to discourage the information acquisition as much as possible, since it knows that by looking closely at its product the consumers will mostly get discouraged from buying it. Therefore, it primarily targets the consumers for which the ad can tip the scales in favor of buying the product without further investigation.

A firm that is uncertain about the quality of its own product thus prefers to hedge and diversify its targeting across the two options above. Curious, however, is exactly the fact that such a firm targets both very optimistic and very pessimistic consumers, while neglecting consumers with average priors, who are investigating the product both with and without the ad. A possible intuition behind this result is that the firm knows that its product is either good, or bad, with positive probabilities, hence going for one of the extremes is optimal – while targeting consumers with moderate opinions would be optimal if the firm knew it has an average-quality product for sure.

Finally, even for a firm which is not confident in its own product, the double-peaked strategy only arises if the amount ρ of information contained in its ad is low enough relative to the information acquisition/contemplation cost factor λ. In particular, if the ad is very persuasive (ρ very high) then there exists a group of consumers whom the firm can convert – raise their purchasing probability $\mathcal{P}(g_i)$ from zero to one, i.e., take consumers who did not even contemplate about buying the product and convince them to buy it without second thought. Advertising to this group would be strictly optimal for the firm regardless of its opinion about its product. In other words, the two consumer groups described above merge into one group in this scenario. The higher is the contemplation cost λ, the lower is the informativeness cutoff for ρ, above which this happens.

4.4 Formal Result

Introduce two threshold types:

$$\bar{g} \equiv \min\{g_i \mid \mathcal{P}^f(\alpha(g_i)) = 1\}$$
$$\underline{g} \equiv \max\{g_i \mid \mathcal{P}^f(g_i) = 0\}$$

(both are well defined since $\mathcal{P}^f$ is continuous and crosses both zero and one for $R < 1$).

We will consider two cases depending on the relation between these two thresholds.

Definition. *An advertisement q is:*

- weak *if $\underline{g} < \bar{g}$;*
- strong *if $\underline{g} \geq \bar{g}$.*

To clarify, an ad is strong if there is some prior belief $g_i \in [\bar{g}, \underline{g}]$ with which a consumer was definitely not buying the product before an ad and would definitely buy it after an ad. If the set of such priors is empty, we say that the ad is weak (this is the case in Figure 1). This separation can equivalently be phrased in terms of a bound on q: an ad is strong if and only if $\rho \geq \bar{\rho}$ for some cutoff value $\bar{\rho} > 1$. This is easy to see given that $\underline{g}$ does not depend on ρ, while $\bar{g}$ is decreasing in ρ, and for $\rho = 1/2$ we have $\bar{g} > \underline{g}$.

We formulate our result in two stages. First, the following lemma gives a good understanding of the environment.

Lemma 1. *Function $\mathcal{A}(g_i)$ is continuous, weakly increasing for $g_i \leq \min\left\{\underline{g}, \bar{g}\right\}$, weakly decreasing for $g_i \geq \max\left\{\underline{g}, \bar{g}\right\}$. It is also constant for $g_i \in [\bar{g}, \underline{g}]$ if the ad is strong, and strictly convex for $g_i \in [\underline{g}, \bar{g}]$ if the ad is weak.*

Proof. Since $\mathcal{P}^f(g)$ is a strictly increasing function truncated to $[0,1]$ and $\alpha(g_i) > g_i$, there exist $\bar{\bar{g}}, \underline{\underline{g}} \in [0,1]$ such that:

$$\mathcal{A}(g_i) = \begin{cases} 0 & \text{for } g_i \leq \underline{\underline{g}}, \\ \mathcal{P}^f(\alpha(g_i)) & \text{for } g_i \in [\underline{\underline{g}}, \min\left\{\underline{g}, \bar{g}\right\}], \\ \mathcal{P}^f(\alpha(g_i)) - \mathcal{P}^f(g_i) & \text{for } g_i \in [\min\left\{\underline{g}, \bar{g}\right\}, \max\left\{\underline{g}, \bar{g}\right\}], \\ 1 - \mathcal{P}^f(g_i) & \text{for } g_i \in [\max\left\{\underline{g}, \bar{g}\right\}, \bar{\bar{g}}], \\ 0 & \text{for } g_i \geq \bar{\bar{g}}. \end{cases}$$

The continuity and monotonicities follow immediately from this representation. If the ad is strong then $\mathcal{A}(g_i) = 1$ for $g_i \in (\bar{g}, \underline{g})$. If the ad is weak then the second derivative of $\mathcal{A}(g_i)$ for $g_i \in (\min\left\{\underline{g}, \bar{g}\right\}, \max\left\{\underline{g}, \bar{g}\right\})$ is given by

$$\begin{aligned} &\frac{d^2}{(dg_i)^2}\left[\mathcal{P}^f(\alpha(g_i)) - \mathcal{P}^f(g_i)\right] = \\ &= \frac{d^2}{d\alpha(g_i)^2}\mathcal{P}^f(\alpha(g_i)) \cdot \left(\frac{d}{dg_i}\alpha(g_i)\right)^2 + \frac{d}{d\alpha(g_i)}\mathcal{P}^f(\alpha(g_i)) \cdot \frac{d^2}{(dg_i)^2}\alpha(g_i) - \frac{d^2}{(dg_i)^2}\mathcal{P}^f(g_i) \\ &= \frac{2\left(\frac{\rho}{1-\rho} - 1\right)\left[g_f(1-g_i)^3\left(1 - e^{-\frac{R}{\lambda}}\right)^2 + \frac{\rho}{1-\rho}(1-g_f)(g_i)^3\left(e^{\frac{1-R}{\lambda}} - 1\right)^2\right]}{\left(e^{\frac{1}{\lambda}} - 1\right)\left(e^{\frac{1-R}{\lambda}} - 1\right)\left(1 - e^{-\frac{R}{\lambda}}\right)\frac{\rho}{1-\rho}\left[(1-g_i)g_i\right]^3} \end{aligned}$$

which is strictly positive since $\frac{\rho}{1-\rho} > 1$. □

This lemma implies the main result.

Theorem. *In a cursed equilibrium, if some advertising is optimal (i.e., $\mathcal{T} \neq \varnothing$) then the firm's optimal targeting strategy is given by:*

1. *$\mathcal{T}_f = (\underline{g}_l, \underline{g}_r) \cup (\bar{g}_l, \bar{g}_r)$ if the ad is weak, with $\underline{g}_l \leq \underline{g} \leq \underline{g}_r, \bar{g}_l \leq \bar{g} \leq \bar{g}_r$, and with one of the two intervals possibly empty;*
2. *$\mathcal{T}_f = (\bar{g}_l, \underline{g}_r)$ if the ad is strong, with $\bar{g}_l \leq \bar{g} \leq \underline{g} \leq \underline{g}_r$.*

Proof. As argued previously, the firm targets consumer $i \in \mathcal{I}$ if and only if $\mathcal{A}(g_i) > c$. From the assumption that $\mathcal{T} \neq \varnothing$ it follows that $\max_{g \in [0,1]} \mathcal{A}(g) > c$. From the continuity of $\mathcal{A}(g)$, its upper contour set $\{g | \mathcal{A}(g) > c\}$ is open in $[0,1]$ for any c. If the ad is strong then by Lemma 1 the maximum of $\mathcal{A}(g)$ is attained by all $g \in \left[\bar{g}, \underline{g}\right]$. By the monotonicity of $\mathcal{A}(g)$ for $g \leq \bar{g}$ and for $g \geq \underline{g}$ described in Lemma 1, part 2 follows.

If the ad is weak then by Lemma 1 $\mathcal{A}(g)$ is strictly convex for all $g \in [\underline{g}, \bar{g}]$. Together with the monotonicity of $\mathcal{A}(g)$ in the remaining regions, this implies that $\arg\max_{g \in [0,1]} \mathcal{A}(g) \in \{\underline{g}, \bar{g}\}$. Consider two cases depending on whether there exists a $\tilde{g} \equiv \arg\min_{g \in (\underline{g}, \bar{g})} \mathcal{A}(g)$. If it does then consider $\mathcal{A}(g)$ separately on $[0, \tilde{g}]$ and $[\tilde{g}, 1]$. On both intervals $\mathcal{A}(g)$ is single-peaked, hence quasi-concave, meaning that its upper contour sets $\{g | \mathcal{A}(g) > c\}$ are convex within each interval and include the respective peaks $\underline{g}$ and $\bar{g}$. If $\tilde{g}$ does not exist then $\mathcal{A}(g)$ is strictly monotone on $[\underline{g}, \bar{g}]$. Then $\mathcal{A}(g)$ is single-peaked on $[0, 1]$, so again its upper contour set $\{g|\mathcal{A}(g) > c\}$ is convex and includes the global maximum.[5] This proves part 1 of the theorem. □

The theorem says that it is always optimal for the firm to target some group of consumers that are close to being indifferent between buying the product and investigating it, and/or some group close to indifference between investigating and not buying. If the ad is weak but the advertising cost is low then the two groups may merge into one – then $\mathcal{T}_f = [\underline{g}_l, \bar{g}_r]$, and the values $\underline{g}_r = \bar{g}_l$ cannot be pinned down uniquely.

4.5 Comparative Statics

Comparative statics sketch:

- If g_f increases then under weak ad left interval expands and the right interval shrinks, while under strong ad the unique interval shifts left.
- If λ decreases then under weak ad the intervals move away from each other and (shrink or expand?). Under strong ad we may switch to weak ad, but either way the interval (shrinks or expands?).
- If c decreases then intervals grow - trivial.
- If ρ increases while weak then ads target more diverse groups. If ρ increases while strong then the single interval grows – trivial.

4.5.1 Firm's Belief

We begin by investigating the effect of changes in the firm's belief, g_f. The illustrative example in Section 4.3 shows the big picture: if the ad is weak then as the firm's belief in own product g_f improves, it stops targeting optimistic consumers with beliefs close to $\bar{g}$ and begins targeting pessimistic consumers with beliefs close to $\underline{g}$. For intermediate beliefs the firm may target both groups simultaneously. In this section we look more closely at this dimension of comparative statics and explore how exactly does the firm's targeting strategy depend on g_f.

Proposition 1. *As g_f increases, the firm's optimal targeting strategy changes as follows:*

1. *if the ad is weak then $[\underline{g}_l, \underline{g}_r]$ expands and $[\bar{g}_l, \bar{g}_r]$ shrinks (from both edges);*
2. *if the ad is strong then $[\bar{g}_l, \underline{g}_r]$ decreases (both edges decrease).*

Proof. TBD □

[5] Note that the UCS may or may not include the second point $\{\underline{g}, \bar{g}\}$.

4.6 Other Equilibria

The cursed equilibrium is not the unique equilibrium in our model. The multiplicity stems from the fact that equilibria are to some extent self-reinforcing: benefit from advertising to some consumer i is larger when he expects to receive an ad (from a firm with a good signal) in equilibrium than when he does not. This is because in the former case sending the ad increases his belief from $\beta(g)$ to $\alpha(g)$, while in the latter the increase is from $g > \beta(g)$ to $\alpha(g)$.

However, this also implies that the cursed equilibrium features the smallest amount of advertising among all equilibria, as formalized by the following proposition.

Proposition 2. *In the cursed equilibrium, set $\mathcal{T}_f$ of targeted consumers is the smallest among all equilibria of the game.*

Proof. Fix any equilibrium of the game. In this equilibrium, consumer with prior g is targeted if $\hat{\mathcal{A}}(g) > c$ and only if $\hat{\mathcal{A}}(g) \geq c$. In the cursed equilibrium, consumer with prior g is targeted if and only if $\mathcal{A}(g) > c$. If $\hat{\mathcal{A}}(g) \geq \mathcal{A}(g)$, hence if the consumer is targeted in the cursed equilibrium, he is also targeted in the equilibrium under consideration. □

The proposition above provides a justification for looking at the cursed equilibrium as the one least tainted by advertising. This is on top of its appeal as being the unique equilibrium (up to indifference) when the consumers are cursed, which is an empirically appealing assumption in this setting.

5 Conclusion

We find an optimal targeting strategy of a monopolistic firm who is facing rationally inattentive consumers with heterogeneous beliefs. We show that it is built around two distinct groups of consumers: (i) the ones who are relatively optimistic about the product and are close to buying it without information acquisition and (ii) the ones who are relatively pessimistic about the product and are close to acquire some information about the product. The optimal targeting strategy also depends on the firm's knowledge about its own product.

The model can be expanded to a situation with several competing firms. The implications of the model can be tested, for instance, on data on political advertising in the United States.

References

S. P. Anderson, A. Baik, and N. Larson. Price Discrimination in the Information Age: Prices, Poaching, and Privacy with Personalized Targeted Discounts. 2019.

S. Athey and J. S. Gans. The Impact of Targeting Technology on Advertising Markets and Media Competition. *American Economic Review*, 100(2):608–613, May 2010. ISSN 0002-8282. doi: 10.1257/aer.100.2.608.

D. Bergemann and A. Bonatti. Targeting in advertising markets: implications for offline versus online media. *The RAND Journal of Economics*, 42(3):417–443, 2011. ISSN 1756-2171. doi: 10.1111/j.1756-2171.2011.00143.x.

D. Bergemann, B. Brooks, and S. Morris. The Limits of Price Discrimination. *American Economic Review*, 105(052):921–957, 2013. ISSN 0002-8282. doi: 10.1257/aer.20130848.

D. Bergemann, A. Bonatti, and A. Smolin. The Design and Price of Information. *American Economic Review*, 108(1):1–48, Jan. 2018. ISSN 0002-8282. doi: 10.1257/aer.20161079.

A. W. Bloedel and I. Segal. Persuasion with Rational Inattention. *SSRN Electronic Journal*, 2018. ISSN 1556-5068. doi: 10.2139/ssrn.3164033.

A. L. Brown, C. F. Camerer, and D. Lovallo. To Review or Not to Review? Limited Strategic Thinking at the Movie Box Office. *American Economic Journal: Microeconomics*, 4 (2):1–26, 2012. ISSN 1945-7669. doi: 10.2307/23249431.

E. Carroni, L. Ferrari, and S. Righi. The price of discovering your needs online. *Journal of Economic Behavior & Organization*, 164:317–330, Aug. 2019. ISSN 0167-2681. doi: 10.1016/j.jebo.2019.06.008.

J. Chen and J. Stallaert. An economic analysis of online advertising using behavioral targeting. *MIS Quarterly*, 38(2):429–450, 2014.

D. Condorelli and B. Szentes. Information Design in the Holdup Problem. *Journal of Political Economy*, pages 000–000, Jan. 2020. ISSN 0022-3808, 1537-534X. doi: 10.1086/704574.

T. M. Cover and J. A. Thomas. *Elements of information theory*. John Wiley & Sons, 2012.

M. Deversi, A. Ispano, and P. Schwardmann. Spin Doctors. 2019.

E. Eyster and M. Rabin. Cursed Equilibrium. *Econometrica*, 73(5):1623–1672, Sept. 2005. ISSN 0012-9682, 1468-0262. doi: 10.1111/j.1468-0262.2005.00631.x.

A. Farahat and M. Bailey. How Effective is Targeted Advertising? 2012.

A. Goldfarb and C. Tucker. Digital Economics. *Journal of Economic Literature*, 57(1): 3–43, Mar. 2019. ISSN 0022-0515. doi: 10.1257/jel.20171452.

O. Gossner, J. Steiner, and C. Stewart. Attention Please! 2018.

B. Hébert and M. Woodford. Rational inattention and sequential information sampling. Technical report, 2017.

S. Ichihashi. Online Privacy and Information Disclosure by Consumers. 2019.

A. Ispano and P. Schwardmann. Competition over Cursed Consumers. 2018.

G. Iyer, D. Soberman, and J. M. Villas-Boas. The Targeting of Advertising. *Marketing Science*, 24(3):461–476, Aug. 2005. ISSN 0732-2399, 1526-548X. doi: 10.1287/mksc.1050.0117.

V. Jain and M. Whitmeyer. Competing to Persuade a Rationally Inattentive Agent. *arXiv:1907.09255 [econ]*, July 2019. arXiv: 1907.09255.

K. Jerath and Q. Ren. Consumer Attention Allocation and Firm Information Design. 2019.

G. Z. Jin, M. Luca, and D. Martin. Is No News (Perceived as) Bad News? An Experimental Investigation of Information Disclosure. Working Paper 21099, 2018.

A. Kolotilin, T. Mylovanov, A. Zapechelnyuk, and M. Li. Persuasion of a Privately Informed Receiver. *Econometrica*, 85(6):1949–1964, 2017. ISSN 0012-9682. doi: 10.3982/ECTA13251.

X. Li and L. M. Hitt. Self-selection and information role of online product reviews. *Information Systems Research*, 19(4):456–474, 2008.

B. Mackowiak, F. Matejka, and M. Wiederholt. Rational inattention: A disciplined behavioral model. *Goethe University Frankfurt mimeo*, 2018.

F. Matějka and A. McKay. Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model. *American Economic Review*, 105(1):272–298, 2015.

L. Matysková. Bayesian Persuasion with Costly Information Acquisition. 2018.

L. Matysková and J. Šípek. Manipulation of Cursed Beliefs in Online Reviews. 2017.

S. Morris and P. Strack. The Wald problem and the equivalence of sequential sampling and static information costs. *Unpublished manuscript, June*, 2017.

M. Preuss. Search, Learning, and Tracking. 2018.

A.-K. Roesler and B. Szentes. Buyer-optimal learning and monopoly pricing. *American Economic Review*, 107(7):2072–2080, 2017.

C. A. Sims. Rational Inattention and Monetary Economics. In B. M. Friedman and M. Woodford, editors, *Handbook of Monetary Economics*, volume 3, pages 155–181. Elsevier, Jan. 2010. doi: 10.1016/B978-0-444-53238-1.00004-1.

A. Wald. Sequential tests of statistical hypotheses. *The annals of mathematical statistics*, 16(2):117–186, 1945.