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

August 2, 2020

Abstract

This paper argues, in the context of targeted advertising, that receivers' rational inattention or ability to independently acquire information have a non-trivial impact on the sender's optimal disclosure strategy. In our model, a monopolist has an opportunity to launch an advertising campaign and chooses a targeting strategy – which consumers to send its advertisement to. The consumers are uncertain about and heterogeneous in their valuations for the product, and are rationally inattentive in that they must incur a cost if they want to learn their true valuations. 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 but may ignore all consumers in between.

JEL-codes: D83, L15, M37.

Keywords: advertising, targeting, rational inattention, costly disclosure.

*We are grateful to Mogens Fosgerau, Alessandro Ispano, Andrey Minaev, and Adrien Vigier, as well as seminar participants at the University of Copenhagen and OLIGO 2020 virtual conference participants 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 economic literature has largely focused on the implications of the more extensive information collection 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 its 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]

The folk wisdom implies that firms benefit most from advertising to consumers who are avoiding the product by a thin margin. If we adopt adopt the informational approach favored in Economics – namely, that consumers are uncertain about their valuation for the product, and advertising makes them more optimistic regarding this valuation – then the firm should target its advertising towards consumers, whose beliefs are just a bit too pessimistic to justify a purchase. However, this reasoning ignores the fact that consumers have the opportunity to acquire information after being subjected to advertising. Indeed, it seems implausible, especially with larger purchases such as automobiles or major appliances, that a consumer would change her mind based purely on advertising. Instead, an ad would lead the consumer to investigate the product more or less closely, and affect her purchasing decision indirectly through this channel. We explore the trade-offs of targeted advertising in such a setting, where consumers strategically acquire information about the product. We show that this negates the folk wisdom outlined above and leads to quite interesting conclusions and novel optimal advertising strategies. Worth noting that we focus on persuasive advertising – that which means to improve consumers' willingness to pay for a product – rather than informative advertising meant to make consumers aware of the product.

In particular, we look at a theoretical model, where a population of rationally inattentive consumers is faced with an option to buy a product which yields an uncertain payoff. The consumers differ in their initial perceptions 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

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

select freely which consumers will receive the ad – i.e., target the ad towards the most susceptible consumers.

We discover that the optimal advertising strategy deviates quite significantly from the intuitive strategy of targeting consumers who are ex ante "almost indifferent" between buying the product and not. 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: an optimistic firm will foster search, while a pessimistic one would rather discourage it. Curiously, if the firm is uncertain enough about the marketability of its product, then it will advertise to both groups simultaneously, while ignoring consumers with average beliefs – who are ex ante "closest to indifference" between buying the product and not. This is because such consumers choose to investigate the product regardless of advertising – meaning that advertising does not have much of an effect on their behavior. 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 and what kind of information a consumer acquires).

1.1 Illustrative example

The main driving force behind our result can be illustrated using the following highly stylized example. Suppose a firm sells an item of quality $s \in \{H, L\}$ at some fixed price normalized to one. The firm values the item at zero. The consumer values a high-quality item at $v = 1 - R$ net of the price, and a low-quality item at $v = -R$. Both the firm and the consumer ex ante believe the product is of high quality with probability $p^0 = 1/3$. The firm receives a signal $y \in \{h, l\}$ about s with precision $\rho = 0.8$ (so with 80% probability the signal is correct). The firm can reveal (advertise) this signal to the consumer at some cost $c < 1/3$. After hearing this ad – if any – the consumer can choose to investigate the quality (learn it perfectly) at cost $\lambda = 0.1$. Assume for this example that the firm's signal is high, $y = h$, so observing y increases a player's belief in $s = H$ to $\alpha(p^0) = 2/3$. Assume also that if the consumer receives no ad, her belief remains at p^0.

The consumer effectively has three options: pass, buy, or investigate (and then buy if and only if $s = H$). Her expected payoffs from these options are given by:

$$\mathbb{E}U = \begin{cases} 0, & \text{if pass;} \\ p - R = p(1-R) + (1-p)(-R), & \text{if buy;} \\ p(1-R) - \lambda = p(1-R) + (1-p)0 - \lambda, & \text{if investigate.} \end{cases}$$

The consumer then passes if $R \geq \bar{R}$, investigates if $R \in (\underline{R}, \bar{R})$, and buys if $R \leq \underline{R}$, where $(\underline{R}, \bar{R}) = (0.15, 0.7)$ if the consumer does not receive the ad, and $(\underline{R}, \bar{R}) = (0.3, 0.85)$ if she does.

The firm's expected profit, depending on the consumer's decision, is given by: $\mathbb{E}\pi = 0$ if the consumer passes, $\mathbb{E}\pi = 1$ if she buys, and $\mathbb{E}\pi = \alpha(p) = 2/3$ if she investigates. Advertising to a consumer with $R \in [0.15, 0.3]$ pushes her from investigating the product to buying it immediately, increasing the firm's profit by $1/3$. Advertising to a consumer with $R \in [0.7, 0.85]$ nudges her to at least investigate the product instead of ignoring it altogether, thereby increasing the firm's profit by $2/3$. Advertising to a consumer with any other R (including the ex ante indifferent consumer with $R = 1/3$) would not affect her behavior and is thus pointless.

Of note is the fact that the firm in this example profits more from advertising to pessimistic (high-R) consumers. This is driven by the firm's optimistic belief $p_f = 2/3$ in own quality (conditional on $y = h$). Optimism leads firm to expect that any investigating consumer is likely to learn that the product quality is high. Therefore, incentivizing the otherwise unwilling consumers to investigate is valuable for the firm, while nudging the investigating consumers to buy unconditionally is less so. Conversely, if prior p^0 and signal strength ρ were low, targeting optimistic (low-R) consumers would be better for the firm.

One may wonder whether the intuition above holds under more general information acquisition processes, under which the consumer is not as restricted in her information acquisition strategy. The remainder of the paper explores a more general model, which provides insight into this and other issues.

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 properties. Section 5 concludes. Most proofs are retained in the main text, since we believe they are concise and clear enough to convey valuable intuition. However, some of the longer proofs and supplementary lemmas are relegated to the Appendix.

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.

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. Relatedly, Jerath and Ren [2019] explore the firm's incentives to limit the amount of information consumers can potentially acquire about the product. 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. Our model is different from all of these in that our sender (the firm) is constrained in which messages it can send – i.e., we explore a model of disclosure, rather than Bayesian Persuasion.

Gossner, Steiner, and Stewart [2018] demonstrate in a setting with strategic information acquisition by the consumer that attracting consumer's attention to own product – in the sense of forcing the consumer to investigate it and not other products, – is always beneficial for the firm. We demonstrate how such "attention attraction" manifests in a purely informational setting, where the consumer's attention is rationally allocated and responds to external stimuli, but cannot be exogenously directed towards a product.

There also exists an extremely rich literature on Bayesian Persuasion and information design without the information acquisition layer (rational inattention), particularly in the context of bilateral trade. For some examples see Bergemann, Brooks, and Morris [2013], Roesler and Szentes [2017], Kolotilin, Mylovanov, Zapechelnyuk, and Li [2017], and Condorelli and Szentes [2020]. We differ from this literature in allowing both the consumer to acquire information and the firm to supply it at the same time, as compared to the information structure being exogenous or chosen exclusively by one side of the market.

3 The Model

The market consists of a single firm and a continuum of consumers $\mathcal{I}$ with unit demand each. The firm offers for sale a product of unknown quality $s \in S \equiv \{H, L\}$ at some exogenously fixed price. All consumers and the firm share a common prior which assigns probability $p^0 \in (0,1)$ to the quality being high. The consumers' net product valuations (i.e., value minus price) are given by $v_L = 0$ for a low-quality product and $v_H = 1$ for a high-quality product.[2] The consumers vary in their reservation utilities in case of not buying the product, given by $R_i \in (0,1)$. Equivalently, $-R_i$ can be interpreted as the level of consumer's ex ante interest in the product. Each consumer's R_i is known to the firm (due to the use of tracking technologies) and is hereinafter referred to as the type of consumer i. The marginal costs of production (and the firm's value for the product) are normalized to zero.

The firm has an opportunity to advertise its product. In particular, at the beginning of the game it receives a private signal $y \in Y \equiv \{h, l\}$ with precision $\rho \in (1/2, 1)$, meaning $\mathbb{P}(h|H) = \mathbb{P}(l|L) = \rho$. The firm can verifiably disclose this signal in an advertisement, and it chooses which consumers $\mathcal{T}(y) \subseteq \mathcal{I}$ will receive the ad. We see y as a piece of hard but inconclusive evidence that the firm can disclose to the selected consumers. This may include awards, rankings standings, critics' reviews, results of external quality evaluations etc. Hereinafter we will refer to $\mathcal{T} : Y \to 2^{\mathcal{I}}$ as the firm's (pure) advertising strategy. The cost of such advertising campaign depends on its size and is given by $c \cdot |\mathcal{T}(y)|$ for some per-consumer cost $c > 0$. The consumers do not explicitly observe the firm's choice of $\mathcal{T}(y)$. More generally, we can allow for mixed strategies $\tau : Y \to \Delta(2^{\mathcal{I}})$, which lead the firm to randomize between different sets of consumers given y. We will use $\tau(\mathcal{T}|y)$ to denote the probability with which consumer set $\mathcal{T}$ is targeted given y, and $\tau(y)$ to denote the whole probability measure.

[2] Our results will hold if the consumers' tastes are idiosyncratic, as long as the firm's information (described further) has some predictive power regarding the consumers' valuations.

Each consumer i chooses whether to purchase the product or not. Prior to making the decision (but after receiving the ad, if any), the consumer has an opportunity to acquire a noisy signal $x_i \in \mathbb{R}$ about her valuation for the product. The consumer can choose any distribution of quality-contingent signals $\mathcal{G}_i(x_i|s) : S \to \Delta(\mathbb{R})$. However, generating a signal is costly, as described further. Upon observing x_i, each consumer updates her belief using Bayes' rule. Given this updated belief, the consumer decides whether to purchase the product so as to maximize her expected payoff.

While we assume that consumers use Bayes' rule to update their belief upon receiving any ad or private signal, we allow for one possible deviation from Bayesianism when no ad is received. In particular, we analyze a special case of the model in which consumers are *cursed* in the sense of Eyster and Rabin [2005]. In our setting, being cursed means that a consumer does not update her belief if she receives no ad.[3] This version provides a valuable insight into the model with fully *Bayesian* consumers, who update their beliefs upon both receiving an ad and the absence thereof.

The overall timing of the model is as follows:

1. state s is drawn by the nature, not observed by anyone;
2. the firm observes y and chooses the advertising strategy $\mathcal{T}(y)$;
3. every targeted consumer $i \in \mathcal{T}$ receives the ad and every $i \in \mathcal{I}$ updates her belief from the prior p^0 to the interim belief p_i^1;
4. every consumer $i \in \mathcal{I}$ selects an information acquisition strategy $\mathcal{G}_i(x_i|s)$;
5. every consumer $i \in \mathcal{I}$ observes signal x_i and updates her belief to posterior p_i^2;
6. every consumer $i \in \mathcal{I}$ decides whether to purchase the product given p_i^2;
7. payoffs are realized.

The consumers' purchasing decisions at stage 6 are mechanical: consumer i buys the product if and only if the expected utility from doing so is positive:

$$p_i^2 \cdot 1 + (1 - p_i^2) \cdot 0 \geq R_i \Leftrightarrow p_i^2 \geq R_i$$

(we break ties in firm's favor for concreteness). We will henceforth take this purchasing strategy as given and focus on other strategic layers of the game.

It will prove convenient to use the beginning of stage 4 as point of reference. Let $\mathcal{D}(p_i^1, R_i, s)$ denote the probability with which consumer i purchases the item (conditional on her optimal information acquisition strategy) when the realized quality is s. Let $\mathcal{D}(p_i^1, R_i)$ denote the respective unconditional probability from the consumer's perspective

[3] 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 [2018] in the lab. Markets with cursed consumers have been explored by Matysková and Šípek [2017] and Ispano and Schwardmann [2018].

as of stage 4:[4]

$$\mathcal{D}(p_i^1, R_i) \equiv \sum_{s \in S} p_i^1(s) \mathcal{D}(p_i^1, R_i, s).$$

Let $\mathcal{D}_f(p_i^1, R_i)$ represent the probability that consumer i buys the product *as perceived by the firm*, conditional on its respective signal y and the implemented advertising strategy. Given these probabilities, we can define the players' payoffs. In particular, consumer i's expected utility as of stage 4 is given by

$$\sum_{s \in S} p_i^1(s) \left[\mathcal{D}(p_i^1, R_i, s) v_s + (1 - \mathcal{D}(p_i^1, R_i, s)) R_i \right] - \lambda \kappa(\mathcal{G}_i; p_i^1), \tag{1}$$

where $\lambda \kappa(\mathcal{G}_i; p_i^1)$ is the cost of generating the signal structure $\mathcal{G}_i$, proportional to the expected reduction in entropy between the consumer's interim and posterior beliefs:[5,6]

$$\kappa(\mathcal{G}_i; p_i^1) \equiv -\sum_{s} p_i^1(s) \log p_i^1(s) + \sum_{s} \sum_{x_i} \mathcal{G}_i(x_i) p_i^2(s|x_i) \log p_i^2(s|x_i),$$

with $\lambda \in \mathbb{R}_{++}$ being the information cost factor. In the above, $\mathcal{G}_i(x_i) \equiv \sum_s p_i^1(s) \mathcal{G}_i(x_i|s)$ is the unconditional probability of observing signal x_i.

The firm's expected profit conditional on y and strategy $\mathcal{T}(y)$ is given by

$$\int_{i \in \mathcal{I}} \mathcal{D}_f(p_i^1, R_i) di - c \cdot |\mathcal{T}(y)|. \tag{2}$$

This expression estimates the expected profit as of stage 4 conditional on the consumers' interim beliefs $\left\{p_i^1\right\}_{i \in \mathcal{I}}$. However, in equilibrium, the firm knows that all consumers' prior beliefs are p^0 and knows exactly how these beliefs will react to ads or lack thereof, hence (2) is also a valid representation of the firm's expected profit as of stage 2.

We will be looking for a Perfect Bayesian Equilibrium of the game which consists of the firm's mixed advertising strategy $\tau(y)$ and the consumers' information acquisition strategy $\mathcal{G}(x_i|s)$ such that:

1. information acquisition strategy $\mathcal{G}(x_i|s)$ maximizes consumer i's expected payoff (1);
2. mixed advertising strategy $\tau(y)$ only assigns positive weight to $\mathcal{T}$ which maximize

[4] From this point onwards we incorporate part of equilibrium reasoning into our definitions. In particular, we assume that consumers' behavior from stage 4 onwards only depends on her reservation utility R_i and the induced interim belief p_i^1, but not on her identity i (this incorporates the purchasing strategy we fixed for stage 6, but also requires that all consumers use the same belief updating rules on and off the equilibrium path). This assumption ensures that R_i and p_i^1 serve as sufficient characteristics of consumer i as of stage 4 and we do not need to keep careful track of consumers' labels when defining purchasing probabilities and payoffs.

[5] It can be shown that the optimal $\mathcal{G}_i$ has finite support in our problem (see Matějka and McKay [2015]), hence we assume this from the start.

[6] For a detailed treatment of entropy and information theory see Cover and Thomas [2012].

the firm's expected profit (2) given signal $y \in \{h, l\}$;[7]

3. firm's belief is updated using Bayes' rule.

Recall that we have fixed the consumers' purchasing behavior at stage 6 as "i buys iff $p_i^2 \geq R_i$", hence we do not include this in the equilibrium definition. Furthermore, we do not include the consumers' belief updating rules in the equilibrium definition, since, as mentioned previously, we explore different versions of the model in which consumers follow different updating rules. Finally, whenever the equilibrium advertising strategy $\tau(y)$ is degenerate, i.e., assigns probability one to some pure strategy $\mathcal{T}(y)$ for all y, we will use $\mathcal{T}(y)$ to refer to the equilibrium strategy.

4 Analysis

4.1 Consumers' Problem

We solve the model via backwards induction. It was already mentioned in the setup that in stage 6, consumer i buys the product if and only if $p_i^2 \geq R_i$. Moving up in the timing, we now need to solve the consumers' information acquisition problem in stage 4. This problem is, however, analogous to one of the problems explored by Matějka and McKay [2015] (see Problem 1 in Section F1 of their Online Appendix). We thus use their results to characterize the solution to the information acquisition problem in our model. This solution is described below; an interested reader is welcome to refer to their paper for more details.

The first step to solving the consumer's problem is realizing that given some interim belief, every signal x_i that the consumer can choose to acquire must induce a different action. If two signals x_i induce the same action, the consumer can save on information costs by pooling the two signals into one. This insight was explored in detail by Caplin and Dean [2013]. In our setting, this means that the consumer's optimal signal structure $\mathcal{G}_i$ will assign positive probability to at most two distinct signals – a recommendation to buy and a recommendation to pass.

One of the findings of Caplin and Dean [2013] is that the consumer's posterior p_i^2 after a given recommendation generated by the optimal signal structure $\mathcal{G}_i$ is independent of her interim belief p_i^1 (as long as she decides to acquire any information at all given p_i^1). In particular, if we define

$$\bar{p}_i \equiv \frac{e^{\frac{R_i}{\lambda}} - 1}{e^{\frac{1}{\lambda}} - 1} \qquad \text{and} \qquad \hat{p}_i \equiv \frac{1 - e^{-\frac{R_i}{\lambda}}}{1 - e^{-\frac{1}{\lambda}}} = e^{\frac{1-R_i}{\lambda}} \bar{p}_i, \tag{3}$$

then consumer i buys the product without investigation whenever $p_i^1 \geq \hat{p}_i$, and passes

[7] Hereinafter we assume that if the firm is indifferent between advertising and not to consumer i, it does *not* advertise. Our results are not dependent on this restriction, up to indifference-specific formulations.

on the product without investigation whenever $p_i^1 \leq \bar{p}_i$. If, however, $p_i^1 \in (\bar{p}_i, \hat{p}_i)$, then consumer i investigates the product in such a way as to generate a binary recommendation, with a recommendation to buy bringing her belief up to $p_i^2 = \hat{p}_i$, and a recommendation to pass pushing her belief down to $p_i^2 = \bar{p}_i$.

Note that cutoffs $\bar{p}_i$ and $\hat{p}_i$ vary across consumers, since they depend on reservation utility R_i. This also means that we can invert the result: given some fixed belief p_i, there exist cutoffs $\hat{R}_i$ and $\bar{R}_i$ such that consumer i with interim belief p_i investigates the product if $R_i \in (\hat{R}_i, \bar{R}_i)$, passes if $R_i \geq \bar{R}_i$, and buys if $R_i \leq \hat{R}_i$. These cutoffs can be obtained by inverting the expressions in (3).

The most relevant for our purposes are probabilities $\mathcal{D}(p_i, R_i)$ generated by consumer i's optimal information acquisition strategy. These represent the probability with which consumer i expects to eventually buy the product, as evaluated at stage 4. These probabilities are given by

$$\mathcal{D}(p_i^1, R_i) = \begin{cases} 0, & \text{if } \hat{\mathcal{D}}(p_i^1, R_i) \leq 0; \\ \hat{\mathcal{D}}(p_i^1, R_i), & \text{if } 0 < \hat{\mathcal{D}}(p_i^1, R_i) < 1; \\ 1, & \text{if } 1 \leq \hat{\mathcal{D}}(p_i^1, R_i); \end{cases} \tag{4}$$

$$\text{where } \hat{\mathcal{D}}(p_i^1, R_i) = \frac{p_i^1 \left(e^{\frac{1}{\lambda}} - 1\right) - \left(e^{\frac{R_i}{\lambda}} - 1\right)}{\left(e^{\frac{1}{\lambda}} - e^{\frac{R_i}{\lambda}}\right)\left(1 - e^{-\frac{R_i}{\lambda}}\right)}. \tag{5}$$

Now we can use Theorem 1 in Matějka and McKay [2015] to calculate the respective choice probabilities conditional on the true product quality s:

$$\mathcal{D}(p_i^1, R_i, H) = \frac{\mathcal{D}(p_i^1, R_i) e^{\frac{1-R_i}{\lambda}}}{\mathcal{D}(p_i^1, R_i) e^{\frac{1-R_i}{\lambda}} + (1 - \mathcal{D}(p_i^1, R_i))}, \tag{6}$$

$$\mathcal{D}(p_i^1, R_i, L) = \frac{\mathcal{D}(p_i^1, R_i) e^{-\frac{R_i}{\lambda}}}{\mathcal{D}(p_i^1, R_i) e^{-\frac{R_i}{\lambda}} + (1 - \mathcal{D}(p_i^1, R_i))}. \tag{7}$$

4.2 Firm's Problem: Preliminaries

From this point onwards we explore the firm's decision in stage 2 of the game – namely, its choice of advertising strategy $\mathcal{T}$.

Given the consumers' information acquisition strategy, the firm can calculate its expected sales. In particular, the firm can compute $\mathcal{D}_f(p_i^1, R_i)$, the expected probability that consumer i will buy the product conditional on her interim belief p_i^1. This probability is given by

$$\mathcal{D}_f(p_i^1, R_i) = p_f \mathcal{D}(p_i^1, R_i, H) + (1 - p_f)\mathcal{D}(p_i^1, R_i, L) \tag{8}$$

where p_f is the firm's belief which incorporates both the prior belief p^0 and the good private signal, i.e., $p_f = \frac{p^0\rho}{p^0\rho+(1-p^0)(1-\rho)}$. The question now is how the consumer's interim belief p_i^1 responds to the firm's advertisement or a lack of one.

The firm can only advertise using hard information 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 signal h is

$$p_i^1(h) = \alpha(p^0) \equiv \frac{p^0\rho}{p^0\rho + (1-p^0)(1-\rho)} > p^0, \tag{9}$$

which coincides with p_f, while upon hearing an ad with signal l, the belief is

$$p_i^1(l) = \beta(p^0) \equiv \frac{p^0(1-\rho)}{p^0(1-\rho) + (1-p^0)\rho} < p^0. \tag{10}$$

Since advertising with signal l is costly and depresses the consumer's belief, it is never optimal for the firm to send it.[8] This is true as long as consumers' beliefs react at least somewhat rationally to the lack of advertising, in the sense of beliefs not dropping further than they would have after observing a low signal.

Lemma 1. *In any equilibrium, if in the absence of an ad consumer i's interim belief is $p_i^1 \geq \beta(p^0)$, then $i \notin \mathcal{T}(l)$ for all $\mathcal{T}(l)$ in the support of the equilibrium strategy $\tau(l)$.*

Proof. Sale probability $\mathcal{D}_f(p_i^1, R_i)$ is weakly increasing in p_i^1, which follows immediately from (4)-(8). Hence, sending an ad with $y = l$ to such consumer i would yield less expected sales than not advertising, while costing $c > 0$ to the firm. Therefore, not advertising to i is optimal. □

The condition in the lemma holds for all consumers in all cases that we consider. Hence from this point onwards we focus on the problem of a firm with a high private signal $y = h$, implicitly assuming that the firm with signal l never advertises ($\mathcal{T}(l) = \emptyset$).

4.3 Firm's Problem with Cursed Consumers

We begin by solving a version of the problem with *cursed* consumers who react to advertisements, but not to the lack thereof. In other words, if consumer i is cursed and she receives an ad with signal h or l then her belief changes to $\alpha(p^0)$ or $\beta(p^0)$ respectively, but if she hears nothing then her belief remains at p^0, regardless of the firm's equilibrium strategy (even though a Bayesian consumer would infer from the firm's strategy that silence is suggestive of bad news). Consumers' cursedness is common knowledge. The next subsection demonstrates the connection of this problem to the one with fully Bayesian consumers.

[8] Although disclosing adverse information may, more generally, be optimal in richer settings; see Smirnov and Starkov [2020] for one example and a survey of related papers.

The condition in Lemma 1 holds trivially for all $i \in \mathcal{I}$ since $p^0 > \beta(p^0)$, thus $\mathcal{T}(l) = \emptyset$ and we focus on the high signal $y = h$. The probability that (cursed) consumer i purchases the product is given by $\mathcal{D}_f(p_i^1, R_i) = \mathcal{D}_f(\alpha(p^0), R_i)$ after ad h and $\mathcal{D}_f(p_i^1, R_i) = \mathcal{D}_f(p^0, R_i)$ in the absence of any ad. The firm's expected profit from targeting set $\mathcal{T}(h)$ of consumers with ad h is given by (2), which can then be rewritten as

$$\left[\int_{i\in\mathcal{T}(h)} \mathcal{D}_f(\alpha(p^0), R_i)di + \int_{i\in\mathcal{I}\setminus\mathcal{T}(h)} \mathcal{D}_f(p^0, R_i)di\right] - c\cdot|\mathcal{T}(h)|.$$

Therefore, if we let

$$\mathcal{A}(R_i) \equiv \mathcal{D}_f(\alpha(p^0), R_i) - \mathcal{D}_f(p^0, R_i)$$

denote the ad effect depending on the consumer's initial prior, then the firm will find it optimal to target consumer $i \in \mathcal{I}$ if and only if $\mathcal{A}(R_i) > c$. Notably, the resulting strategy $\mathcal{T}_c(h) \equiv \{i|\mathcal{A}(R_i) > c\}$ is uniquely optimal given consumers' beliefs, hence the firm will never want to mix. Furthermore, beliefs of cursed consumers are independent of the firm's strategy (conditional on a given observation), hence the equilibrium is also unique. Note that uniqueness is meant up to indifferences in both cases – i.e., there may exist other equilibria, in which the firm or some consumer choose a different action when indifferent between some actions available to them (advertise or not to consumer i; buy the product or not).

The linearity of ad campaign cost in its size is the source of separability obtained above: that the decision to advertise to consumer i only depends on R_i, but not on the whole distribution of R in the population. This also implies that instead of choosing which consumers $\mathcal{T}(y) \subseteq \mathcal{I}$ to target with signal y, the firm can equivalently choose which consumer *types* $\mathcal{R}(y) \equiv \{R_i|i \in \mathcal{T}(y)\}$ to target given its private signal y. Hereinafter, we refer to $\mathcal{R}$ as the reserve-representation of a given pure targeting strategy $\mathcal{T}$.

This subsection is summarized by (and proves) the following Proposition.

Proposition 1. *If consumers are cursed, the firm's optimal advertising strategy in the unique equilibrium is pure and given by $\mathcal{R}(l) = \emptyset$ and $\mathcal{R}(h) = \{R_i|\mathcal{A}(R_i) > c\}$.*

In what follows, we will use $\mathcal{T}_c$ to refer to the firm's advertising strategy in the equilibrium with cursed consumers, and $\mathcal{R}_c$ for its reserve-representation.

The following subsection demonstrates that the same advertising strategy is optimal when consumers are Bayesian. Section 4.5 then demonstrates that this optimal strategy has the bipartite structure announced in the Introduction.

4.4 Firm's Problem with Bayesian Consumers

We now expand the analysis above to allow for fully Bayesian consumers, as originally intended. Such consumer i would still update her belief to $\alpha(p^0)$ or $\beta(p^0)$ after an ad with signal h or l respectively. However, in case she does not receive an ad, she would make inference from that fact based on the firm's equilibrium advertising strategy, potentially implying $p_i^1 \neq p^0$. Notably, $p_i^1 \in [\beta(p^0), \alpha(p^0)]$ in the absence of an ad for all i, so Lemma 1 applies, and in all equilibria with Bayesian consumers $\mathcal{T}(l) = \emptyset$ is uniquely optimal after $y = l$.

We now show that the optimal advertising strategy in this case is the same as with cursed consumers (in one equilibrium). The intuition behind this lies in the exact way cursed and Bayesian consumers differ in our problem. In particular, if consumer i does not expect to receive an ad, then it does not matter for the firm whether she is cursed or Bayesian, since in the absence of an ad her belief remains at $p_i^1 = p^0$ in either case. Therefore, if advertising to i is not optimal if i is cursed, it is also not optimal when she is Bayesian, since both the cost and the benefit of an ad are the same. Conversely, if consumer i does expect to receive an ad and it was optimal to advertise to her when she was cursed, then it is surely optimal to advertise if she is Bayesian. This is because an ad improves her belief to $\alpha(p^0)$ in either case, but her reaction to the absence of an ad is weakly worse when she is Bayesian ($p_i^1 \leq p^0$).

Therefore, when a firm is developing its advertising strategy, it is a safe bet for it to assume that all consumers are cursed and unresponsive to the lack of advertising – regardless of whether this is actually true, or consumers are sophisticated but correctly anticipate the firm's strategy. This idea is formalized in our first theorem below.

Theorem 1. *There exists an equilibrium of the game with Bayesian consumers, in which* $\tau(\mathcal{T}_c(y)|y) = 1$ *for* $y \in \{h, l\}$.

Proof. According to Proposition 1, $\mathcal{T}_c$ is such that $\mathcal{R}_c(l) = \emptyset$ and $\mathcal{R}_c(h) = \{R_i | \mathcal{A}(R_i) > c\}$. Optimality of $\mathcal{R}_c(l) = \emptyset$ is given by Lemma 1.

Now consider $y = h$. If consumer i is Bayesian and $R_i \in \mathcal{R}_c(h)$, then the actual ad effect in terms of firm-expected increase in sale probability is

$$\hat{\mathcal{A}}(R_i) \equiv \mathcal{D}_f(\alpha(p^0), R_i) - \mathcal{D}_f(\beta(p^0), R_i), \tag{11}$$

which is weakly greater than $\mathcal{A}(R_i)$ for all R_i, since $\mathcal{D}_f(\beta(p^0), R_i) \leq \mathcal{D}_f(p^0, R_i)$ (see Lemma 2 in the Appendix). Hence if advertising to consumer i was optimal when she was cursed, it is still optimal to advertise if she is Bayesian. Conversely, if $R_i \notin \mathcal{R}_c(h)$ then the ad effect is still given by $\mathcal{A}(R_i) \leq c$ (consumer does not expect to receive an ad, hence does not update her belief after not receiving one), – so advertising to this consumer is not optimal. $\square$

Note that this result is specific to costly disclosure settings, but does not in any way rely on the consumers' information acquisition layer of the problem. I.e., it applies equally well if the cost of information is $\lambda = +\infty$, but loses its bite (while remaining formally true) when the cost of advertising is $c = 0$, since in the latter case the firm has no reason to ever withhold a positive signal from any consumer.

Hereinafter we refer to the equilibrium in Theorem 1 as the *cursed equilibrium* of the game with Bayesian consumers. The following subsection illustrates and explores this equilibrium in greater detail. Section 4.6 then studies other equilibria of the game.

4.5 Cursed Equilibrium Characterization

In this section we analyze the cursed equilibrium of the game with Bayesian consumers (on-path equivalent to the unique equilibrium with cursed consumers) and state our main result. However, before we can do that, additional notation and definitions are in order. To begin with, let us introduce two threshold consumer types:

$$\bar{R} \equiv \min\{R \mid \mathcal{D}_f(p^0, R) = 0\},$$
$$\underline{R} \equiv \max\{R \mid \mathcal{D}_f(\alpha(p^0), R) = 1\}.$$

Here $\bar{R}$ denotes the type of consumer, who, when her interim belief equals p^0, is indifferent between passing on the product and investigating it (here $\bar{R}$ is analogous to $\bar{R}_i$ in Section 4.1). On the other hand, $\underline{R}$ denotes the type of consumer, who, after receiving a good ad, is indifferent between buying the product immediately and investigating it. Both thresholds are well defined since $\mathcal{D}_f$ is weakly decreasing and continuous in R_i (see Lemma 2 in the Appendix), and $\mathcal{D}_f(p, 1) = 0$, $\mathcal{D}_f(p, 0) = 1$ for any $p \in (0, 1)$. We will consider two cases depending on the relation between these two thresholds.

Definition. *Advertisement ρ is* strong *if there exists $R \in [0, 1]$ such that $\mathcal{D}_f(p^0, R) = 0$ and $\mathcal{D}_f(\alpha(p^0), R) = 1$, and* weak *otherwise.*

Equivalently, an ad is strong if $\underline{R} \geq \bar{R}$ and weak otherwise. In words, an ad is strong if there exists some type $R \in [\bar{R}, \underline{R}]$ – which is such that a (cursed) consumer of this type is definitely passing on the product without an ad, and would definitely buy it after an ad. Another way to frame this division is in terms of a bound on ρ: an ad is strong if and only if $\rho \geq \bar{\rho}$ for some cutoff value $\bar{\rho} > 1/2$, which depends on model parameters.[9]

We are now ready to state the main result, claiming that the optimal advertising strategy in the cursed equilibrium has a bipartite structure, targeting groups of consumers centered at $\underline{R}$ and $\bar{R}$.

Theorem 2. *If the cursed equilibrium prescribes a non-zero level of advertising ($\mathcal{T}_c \neq \emptyset$), then the optimal targeting strategy satisfies the following:*

1. $\mathcal{R}_c(l) = \emptyset$;

[9] Equivalence of this formulation is easy to see given that $\bar{R}$ does not depend on ρ, while $\underline{R}$ is increasing in ρ, and for $\rho = 1/2$ we have $\bar{R} > \underline{R}$.

2. *if the ad is weak:* $\mathcal{R}_c(h) = (\underline{R}_d, \underline{R}_u) \cup (\bar{R}_d, \bar{R}_u)$*, with* $\underline{R}_d \leq \underline{R} \leq \underline{R}_u, \bar{R}_d \leq \bar{R} \leq \bar{R}_u$*, and with one of the two intervals possibly empty;*
3. *if the ad is strong:* $\mathcal{R}_c(h) = (\bar{R}_d, \underline{R}_u)$*, with* $\bar{R}_d \leq \bar{R} \leq \underline{R} \leq \underline{R}_u$*.*

Proof. See Appendix. □

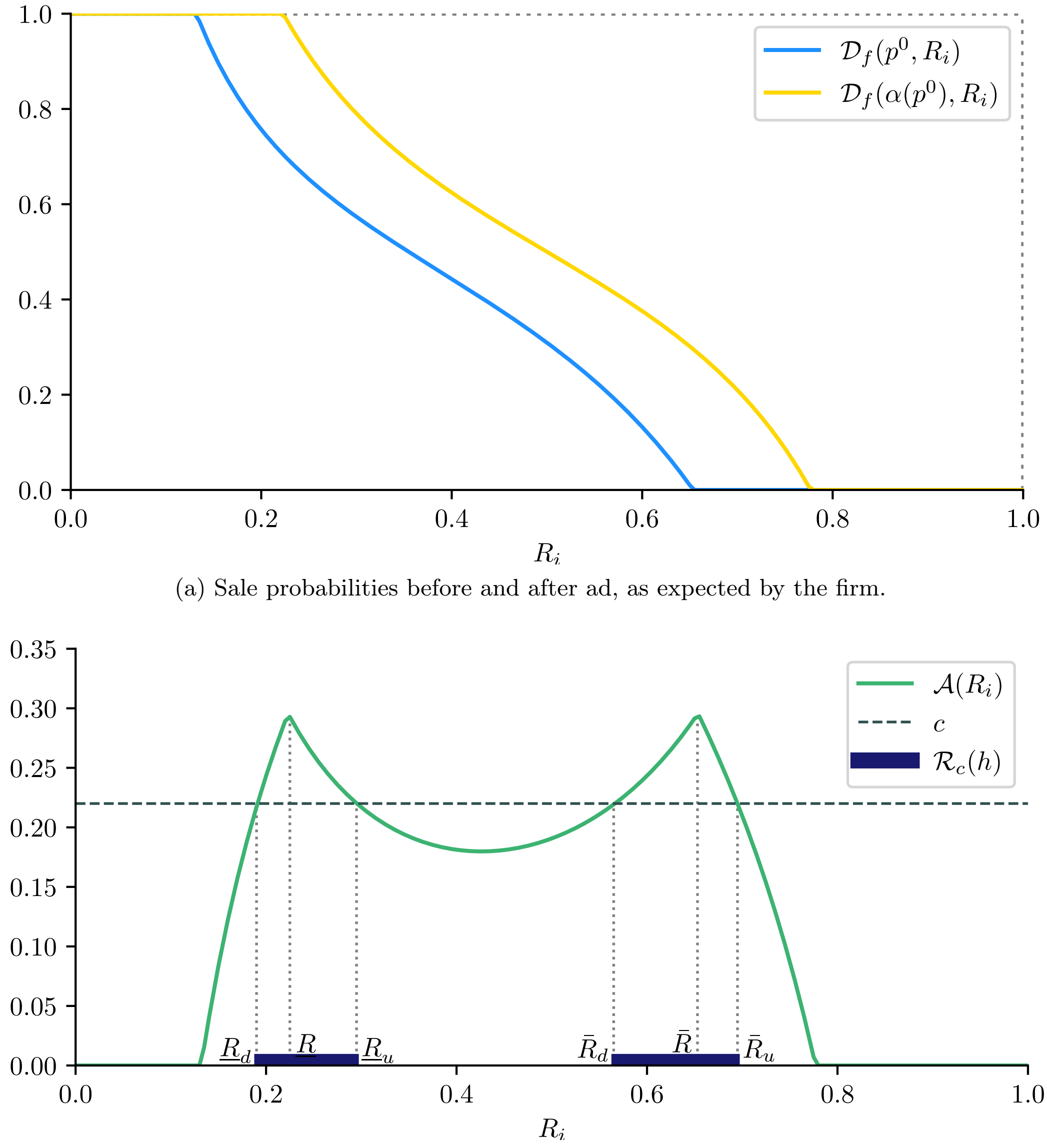
Let us focus on the case of a weak ad. Theorem 2 says that in the cursed equilibrium the firm targets some group of consumers that are close to being indifferent between buying the product and investigating it ($R_i \approx \underline{R}$), and/or some other group close to indifference between investigating and not buying ($R_i \approx \bar{R}$). If the ad is weak but the advertising cost is low then the two groups may merge into one – then $\mathcal{R}_c = (\underline{R}_d, \bar{R}_u)$, and the values $\underline{R}_u = \bar{R}_d$ cannot be pinned down uniquely.

The result may be easier to grasp by looking at Figure 1, which plots the cursed equilibrium for $\lambda = 0.35$, $p^0 = 1/3$, $\rho = 2/3$, and $c = 0.22$. Panel (a) depicts the expected sale probabilities with and without an ad, $\mathcal{D}_f(\alpha(p_i^0), R_i)$ and $\mathcal{D}_f(p_i^0, R_i)$ respectively, as functions of consumers' reservation utility R_i. Panel (b) shows the effect of ads on expected sales as a function of R_i, as well as the equilibrium advertising strategy.[10] As one can see from panel (b), the ad effect is double-peaked in R_i, which leads to the optimal advertising strategy being bipartite, depending on c. Theorem 2 claims that this twin-peak structure is holds generally.

The left peak in Figure 1b is at $R_i = \underline{R}$, i.e., the consumer who is indifferent – after receiving the ad – between investigating the product and buying it immediately. By advertising to this consumer and those with slightly smaller R_i, the firm eliminates any risk that these consumers investigate the product on their own and arrive to the (possibly incorrect) conclusion that the product is bad. The possible sale is converted into a sure sale – or an almost sure sale in case of consumers with R_i slightly above $\underline{R}$. On the other hand, the right peak is at $R_i = \bar{R}$ – the consumer who is indifferent (given no ad) 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 her 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 2 demonstrates that this is partly due to the firm being not sure of the quality of its own product (in Figure 1 we have $p_f = \frac{1}{2}$). In particular, if the firm's belief (equal to the common prior adjusted for good news $y = h$) dictates that product is likely good or likely bad, then its targeting strategy $\mathcal{R}_c$ focuses more on one of the two groups. The firm with a good product prefers to target consumers who would otherwise either not even look at the product (high R_i), or not put much effort

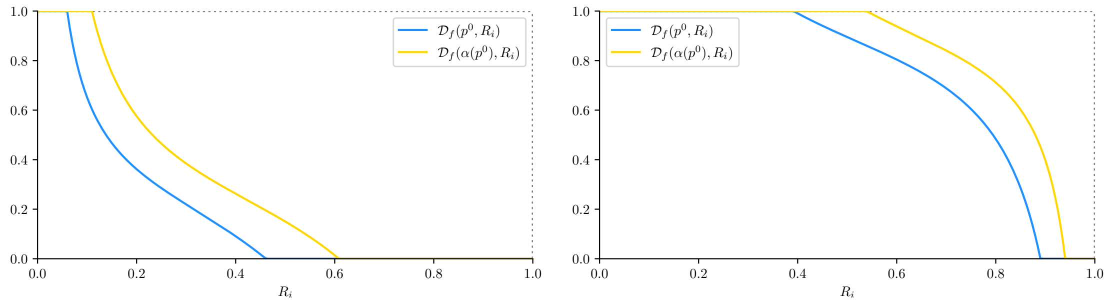
[10] Recall that $\mathcal{A}(R_i)$ measures the ad effect for cursed consumers and those Bayesian consumers who do not expect an ad. The actual ad effect for expecting Bayesian consumers is larger, see Sections 4.4 and 4.6.



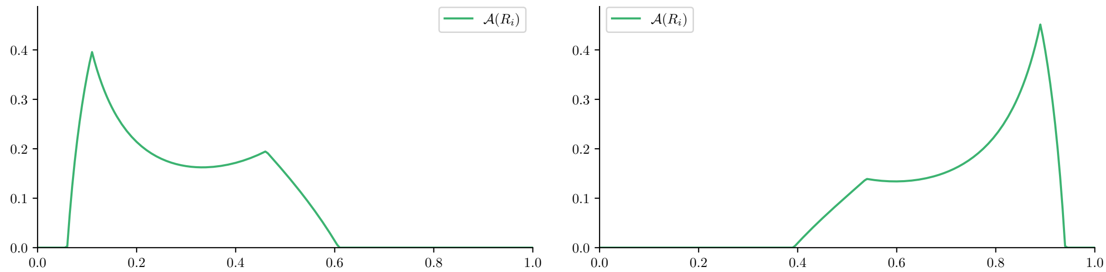
(a) Sale probabilities before and after ad, as expected by the firm.

(b) Ad effect $\mathcal{A}(p_i^0)$ and the optimal advertising strategy $\mathcal{R}(h)$.

Figure 1: The cursed equilibrium.



(a) Sale probabilities, $p^0 = 1/6$ (b) Sale probabilities, $p^0 = 5/7$



(c) Ad effect, $p^0 = 1/6$ (d) Ad effect, $p^0 = 5/7$

Figure 2: Effects of prior p^0.

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 (low R_i). 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.

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 – i.e., if the ad is weak, which is the case we explored so far. In contrast, if the ad is strong (ρ is high) then there exists a group of consumers whom the firm can fully convert – increase the sale probability 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. This group would trivially be the most profitable for the firm to advertise to, as compared to consumers who investigate the product – and thus buy it with probability strictly between zero and one, – in one of the two cases (with or without the ad). In other words, the two consumer groups described above merge into one group in this scenario.

4.6 Other Equilibria with Bayesian Consumers

The cursed equilibrium is not the unique equilibrium in our model when consumers are Bayesian. The multiplicity stems from the fact that equilibria are, to some extent, self-reinforcing: benefit from advertising to some consumer i is larger when she expects to receive an ad (from a firm with signal h) than when she does not. This is because in the former case sending the ad increases her belief from $\beta(p)$ to $\alpha(p)$ – since by not receiving the ad she infers that $y = l$, – while in the latter case the increase is from $p > \beta(p)$ to $\alpha(p)$.

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 any equilibrium with Bayesian consumers, for any $y \in \{h, l\}$: any $\mathcal{T}$ in the support of the firm's equilibrium strategy $\tau(y)$ is such that $\mathcal{T}_c(y) \subseteq \mathcal{T}$.*

Proof. By Lemma 1 and the consumers' belief updating rule, $\tau(\emptyset|l) = 1$, hence for the remainder of the proof consider $y = h$. Consider an arbitrary equilibrium of the game. Benefit from advertising to Bayesian consumer i (in terms of purchasing probability) is between $\hat{\mathcal{A}}(R_i)$ and $\mathcal{A}(R_i)$, since her belief p_i^1 after no ad is $p_i^1 \in [\beta(p^0), p^0]$. The consumer is targeted if and only if the ad benefit is strictly larger than c. It is true that $\hat{\mathcal{A}}(R_i) \geq \mathcal{A}(R_i)$ (see proof of Theorem 1), and in the cursed equilibrium consumer with prior p is targeted if and only if $\mathcal{A}(p) > c$, hence if the consumer is targeted in the cursed equilibrium, she is also targeted in the equilibrium under consideration. □

The proposition above validates the cursed equilibrium as the one least tainted by advertising. This is on top of its appeal as being the unique equilibrium – up to the firm's indifference – when the consumers are cursed, which is an assumption that is empirically appealing by itself in this setting (see footnote 3 for references to works providing empirical and experimental evidence supporting receivers' cursedness in disclosure games).

However, if one insists on consumers being fully Bayesian, a reasonable question to ask would be: "what other equilibria are there, except for the cursed equilibrium?" This is a question that we can answer in this setting, and in a way that seconds our main result. We begin by characterizing another extreme point of the equilibrium set by saying that the equilibrium with the most advertising (exists and) has the exact same structure as the cursed equilibrium.

Proposition 3. *If consumers are Bayesian, there exists an equilibrium with the firm's advertising strategy $\bar{\mathcal{T}}$ such that for any $y \in \{h, l\}$ and any $\mathcal{T}$ in the support of the firm's strategy $\tau(y)$ played in any other equilibrium, $\mathcal{T} \subseteq \bar{\mathcal{T}}(y)$.*

Furthermore, strategy $\bar{\mathcal{T}}$ is such that its reserve-representation $\bar{\mathcal{R}}(y) = \{R_i | i \in \bar{\mathcal{T}}(y)\}$ satisfies properties 1-3 in Theorem 2.

Proof. See Appendix. □

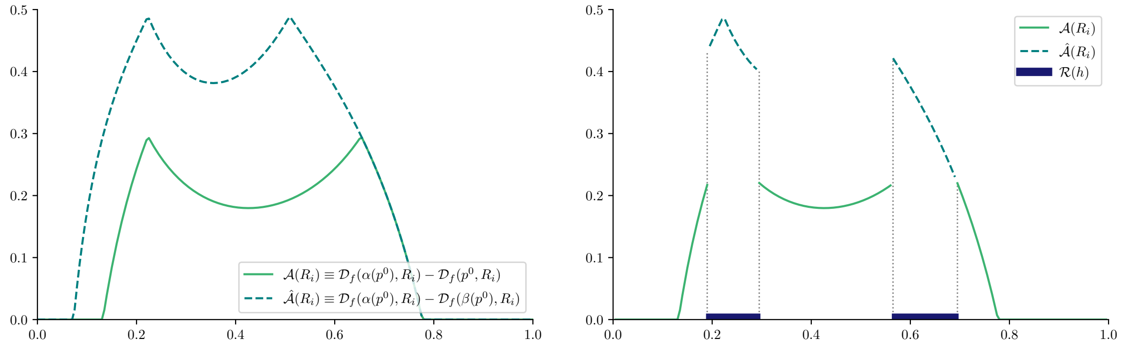


Figure 3: Ad effects $\mathcal{A}(R_i)$ and $\hat{\mathcal{A}}(R_i)$ for cursed and unexpecting Bayesian vs expecting Bayesian consumers, respectively (left panel); the actual ad effect in the cursed equilibrium (right panel).

Both Propositions 2 and 3 can be illustrated using Figure 3. The left panel plots the two values of ad effect, depending on the type of consumer i: $\mathcal{A}(R_i)$ applies when i is either cursed, or is Bayesian but does not expect to receive any ad; $\hat{\mathcal{A}}(R_i)$ as defined in (11) is the ad effect for consumer i who expects to receive an ad in equilibrium. In other words, in equilibrium with Bayesian consumers, the ad effect is given by $\hat{\mathcal{A}}(R_i)$ for all $i \in \mathcal{T}(h)$ and by $\mathcal{A}(R_i)$ for all $i \in \mathcal{I} \backslash \mathcal{T}(h)$. So targeting strategy $\mathcal{T}(h)$ is an equilibrium with Bayesian consumers if and only if $\hat{\mathcal{A}}(R_i) \geq c$ for all $i \in \mathcal{T}(h)$ and $\mathcal{A}(R_i) < c$ for all $i \notin \mathcal{T}(h)$.

The equilibrium multiplicity is produced by the fact that $\hat{\mathcal{A}}(R) > \mathcal{A}(R)$ for all R, which can lead to quite arbitrary targeting sets being self-sustaining in equilibrium. However, the point of Proposition 3 and Figure 3 is that $\hat{\mathcal{A}}(R)$ satisfies all the same properties that $\mathcal{A}(R)$ does, and yields (in the eqiulibrium with the largest $\mathcal{T}(h)$) qualitatively the same targeting strategy as in the cursed equilibrium.

In the end, while we cannot guarantee that the optimal advertising strategy in any equilibrium has the nice bipartite structure as outlined in Theorem 2, the optimal target groups must belong to one of the two broad pools as defined there (one group close to indifference between acquiring a signal and buying unconditionally and another group close to indifference between acquiring a signal and ignoring the product). The only caveat is that the benefit of advertising to some of these consumers is only outweighed by the implicit threat of these consumers growing skeptical of the product quality in case they do not see it advertised. The existence of such a threat in the context of advertising is debatable, but may be more justified in other settings.

5 Discussion

In this paper, we explore the optimal ad targeting strategy of a monopolistic firm which is facing a population of rationally inattentive consumers with heterogeneous outside options. We show that this strategy is generally bimodal, focusing on 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 relative focus on these two groups depends on the firm's expectation regarding its own product.

In our analysis we make a connection to behavioral literature studying "cursed" consumers, who deviate from being Bayesian by making no inferences from observing a lack of signal. While an empirically relevant phenomenon, this "cursedness" contradicts the paradigm of a fully rational, fully Bayesian economic agent used in Economic literature. We show in this paper that in costly disclosure games, this distinction may sometimes be ignored as not affecting the equilibrium strategies.

While we see this paper as primarily normative, it may be interesting to see whether similar strategies are employed in the real world – or, if not, whether using our strategy would improve upon the strategies currently used in real-life scenarios.

The model abstacts from some relevant aspects of the setting, such as flexible pricing (including price discrimination) and competitive effects arising in the presence of multiple sellers. Exploring these extensions in greater detail would be important for understanding the scope of applicability of our results. We expect that accounting for competition will not have any significant effect on our results: it does not matter much whether the firm tries to attract consumers in a game of tug-of-war versus their outside options or versus a competing product. Pricing, on the other hand, has the scope to alter the nature of the problem quite significantly, both via price signaling of the firm's private information as a competing communication channel and personalized pricing as an alternative way to influence consumers' information acquisition decisions.

References

S. P. Anderson, A. Baik, and N. Larson. Price Discrimination in the Information Age: Prices, Poaching, and Privacy with Personalized Targeted Discounts. SSRN Scholarly Paper ID 3428313, Social Science Research Network, Rochester, NY, June 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.

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.

A. Caplin and M. Dean. Behavioral Implications of Rational Inattention with Shannon Entropy. Technical Report w19318, National Bureau of Economic Research, Cambridge, MA, August 2013.

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*, 128(2):681–709, February 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: A model and an Experimental Investigation of Vague Disclosure. SSRN Scholarly Paper ID 3275418, Social Science Research Network, Rochester, NY, 2018.

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

A. Farahat and M. C. Bailey. How effective is targeted advertising? In *Proceedings of the 21st international conference on World Wide Web*, WWW '12, pages 111–120, Lyon, France, April 2012. Association for Computing Machinery. ISBN 978-1-4503-1229-5. doi: 10.1145/2187836.2187852.

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

O. Gossner, J. Steiner, and C. Stewart. Attention Please! SSRN Scholarly Paper ID 3300084, Social Science Research Network, Rochester, NY, November 2018.

B. Hébert and M. Woodford. Rational Inattention and Sequential Information Sampling. Working Paper 23787, National Bureau of Economic Research, September 2017. Series: Working Paper Series.

A. Ispano and P. Schwardmann. Competition Over Cursed Consumers. SSRN Scholarly Paper ID 3208067, Social Science Research Network, Rochester, NY, May 2018.

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

K. Jerath and Q. Ren. Consumer Attention Allocation and Firm Information Design. working paper, 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, National Bureau of Economic Research, 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. SSRN Scholarly Paper ID 3161174, Social Science Research Network, Rochester, NY, April 2018.

L. Matysková and J. Šípek. Manipulation of Cursed Beliefs in Online Reviews. SSRN Scholarly Paper ID 2960745, Social Science Research Network, Rochester, NY, April 2017.

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

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, January 2010. doi: 10.1016/B978-0-444-53238-1.00004-1.

A. Smirnov and E. Starkov. Bad news turned good: reversal under censorship. *American Economic Journal: Microeconomics*, forthcoming, 2020. doi: 10.1257/mic.20190379.

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

Appendix

Lemma 2. *Function $\mathcal{D}_f(p,R)$ is weakly decreasing in R.*

Proof. It is immediate from (4)-(8) that $\mathcal{D}_f(p,R)$ is continuous in R and differentiable in R almost everywhere, with the exception of edge cases in (4). If R is such that $\mathcal{D}(p,R) \in \{0,1\}$ then $\mathcal{D}_f(p,R) = \mathcal{D}(p,R)$, hence monotonicity holds. Otherwise $\mathcal{D}(p,R) \in (0,1)$ – then take the partial derivative of $\mathcal{D}_f(p,R)$:

$$\frac{\partial \mathcal{D}_f(p,R)}{\partial R} = \left(\frac{\partial \mathcal{D}(p,R)}{\partial R} - (1-\mathcal{D}(p,R))\frac{\mathcal{D}(p,R)}{\lambda}\right) \cdot \\ \cdot \left(\frac{p_f e^{\frac{1-R}{\lambda}}}{\left(\mathcal{D}(p,R)e^{\frac{1-R}{\lambda}} + (1-\mathcal{D}(p,R))\right)^2} + \frac{(1-p_f)e^{-\frac{R}{\lambda}}}{\left(\mathcal{D}(p,R)e^{-\frac{R}{\lambda}} + (1-\mathcal{D}(p,R))\right)^2}\right).$$

The second large bracket in the above is nonnegative, since all numerators and denominators are strictly positive ($R \in (0,1)$). In the first bracket, $(1-\mathcal{D}(p,R))\frac{\mathcal{D}(p,R)}{\lambda} > 0$, and

$$\frac{\partial \mathcal{D}(p,R)}{\partial R} = \frac{\partial \hat{\mathcal{D}}(p,R)}{\partial R} = -\frac{e^{\frac{R}{\lambda}}\left[p\left(e^{\frac{1}{\lambda}} - e^{\frac{R}{\lambda}}\right)^2 + (1-p)\left(e^{\frac{R}{\lambda}} - 1\right)^2 e^{\frac{1}{\lambda}}\right]}{\lambda\left[\left(e^{\frac{1}{\lambda}} - e^{\frac{R}{\lambda}}\right)\left(e^{\frac{R}{\lambda}} - 1\right)\right]^2} \leq 0.$$

Therefore, $\frac{\partial \mathcal{D}_f(p,R)}{\partial R} \leq 0$. □

Lemma 3. *Function $\mathcal{A}(R_i)$ is continuous, weakly increasing for $R_i \leq \min\{\underline{R}, \bar{R}\}$, weakly decreasing for $R_i \geq \max\{\underline{R}, \bar{R}\}$. It is also constant for $R_i \in [\bar{R}, \underline{R}]$ if the ad is strong, and strictly convex for $R_i \in [\underline{R}, \bar{R}]$ if the ad is weak.*

Proof. Since $\mathcal{D}_f(p,R)$ is a strictly decreasing function of R truncated to $[0,1]$, is weakly increasing in p, and $\alpha(p^0) > p^0$, there exist $\bar{\bar{R}}, \underline{\underline{R}} \in [0,1]$ such that:

$$\mathcal{A}(R_i) = \begin{cases} 0, & \text{if } R_i \geq \bar{\bar{R}}, \\ \mathcal{D}_f(\alpha(p^0), R_i), & \text{if } R_i \in [\max\{\underline{R}, \bar{R}\}, \bar{\bar{R}}], \\ \mathcal{D}_f(\alpha(p^0), R_i) - \mathcal{D}_f(p^0, R_i), & \text{if } R_i \in [\min\{\underline{R}, \bar{R}\}, \max\{\underline{R}, \bar{R}\}], \\ 1 - \mathcal{D}_f(p^0, R_i), & \text{if } R_i \in [\underline{\underline{R}}, \min\{\underline{R}, \bar{R}\}], \\ 0, & \text{if } R_i \leq \underline{\underline{R}}. \end{cases}$$

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

$$\frac{d^2\mathcal{A}(R)}{dR^2} = \frac{p_f - p^0}{\lambda^2 p^0(1-p^0)} \cdot e^{\frac{R}{\lambda}} \left[p^0 \frac{e^{\frac{R}{\lambda}} + 1}{\left(e^{\frac{R}{\lambda}} - 1\right)^3} + (1-p^0) \frac{e^{\frac{1}{\lambda}}\left(e^{\frac{1}{\lambda}} + e^{\frac{R}{\lambda}}\right)}{\left(e^{\frac{1}{\lambda}} - e^{\frac{R}{\lambda}}\right)^3} \right],$$

which is strictly positive since $p_f > p^0$. □

Proof of Theorem 2. Part 1 holds by Proposition 1 (or, equivalently, Lemma 1).

As argued in text, the firm targets consumer $i \in \mathcal{I}$ with $y = h$ if and only if $\mathcal{A}(R_i) > c$. From the assumption that $\mathcal{T}_c \neq \emptyset$ it follows that $\max_{R\in[0,1]} \mathcal{A}(R) > c$. From the continuity of $\mathcal{A}(R)$ together with $\mathcal{A}(0) = \mathcal{A}(1) = 0$, it follows that the upper contour set $\{R | \mathcal{A}(R) > c\}$ is open for any $c > 0$. If the ad is strong then by Lemma 3 the maximum of $\mathcal{A}(R)$ is attained by all $R \in [\bar{R}, \underline{R}]$. By the monotonicity of $\mathcal{A}(R)$ for $R \leq \bar{R}$ and for $R \geq \underline{R}$ described in Lemma 3, part 3 of the statement follows.

If the ad is weak then by Lemma 3 $\mathcal{A}(R)$ is strictly convex for all $R \in [\underline{R}, \bar{R}]$. Together with the monotonicity of $\mathcal{A}(R)$ in the remaining regions, this implies that $\arg\max_{R\in[0,1]} \mathcal{A}(R) \in \{\underline{R}, \bar{R}\}$. Consider two cases depending on whether there exists a $\tilde{R} \equiv \arg\min_{R\in(\underline{R},\bar{R})} \mathcal{A}(R)$. If it does then consider $\mathcal{A}(R)$ separately on $[0, \tilde{R}]$ and $[\tilde{R}, 1]$. On both intervals $\mathcal{A}(R)$ is single-peaked, hence quasi-concave, meaning that its upper contour sets $\{R | \mathcal{A}(R) > c\}$ are convex within each interval and include the respective peaks $\underline{R}$ and $\bar{R}$. If $\tilde{R}$ does not exist then $\mathcal{A}(R)$ is strictly monotone on $[\underline{R}, \bar{R}]$. Then $\mathcal{A}(R)$ is single-peaked on $[0, 1]$, so again its upper contour set $\{R | \mathcal{A}(R) > c\}$ is convex and includes the global maximum.[11] This proves part 2 of the theorem. □

Proof of Proposition 3. The proof is constructive. Let $\hat{\mathcal{A}}(R_i)$ be defined as in (11). Let $\bar{\mathcal{R}}(l) \equiv \emptyset$ and $\bar{\mathcal{R}}(h) \equiv \{R_i | \hat{\mathcal{A}}(R_i) > c\}$ and, correspondingly, $\bar{\mathcal{T}}(l) \equiv \emptyset$ and $\bar{\mathcal{T}}(h) \equiv \{i \in \mathcal{I} | R_i \in \bar{\mathcal{R}}(h)\}$.

Firstly, we show that there exists an equilibrium with $\bar{\mathcal{T}}$ as the firm's advertising strategy. Any consumer $i \in \bar{\mathcal{T}}(h)$ updates her belief to $p_i^1 = \alpha(p^0)$ if she receives ad h and to $p_i^1 = \beta(p^0)$ if she receives no ad or ad l – since she expects to not receive an ad if and only if $y = l$. The ad effect is given by $\hat{\mathcal{A}}(R_i) > c$, hence it is strictly optimal to advertise to i. On the other hand, any consumer $i \notin \bar{\mathcal{T}}(h)$ updates to $p_i^1 = \alpha(p^0)$ after ad h, $p_i^1 = \beta(p^0)$ after ad l (both are zero-probability events from her perspective), and $p_i^1 = p^0$ in the absence of any ad. The ad effect is then given by $\mathcal{A}(R_i) < \hat{\mathcal{A}}(R_i) \leq c$, hence advertising to i is strictly suboptimal.

Secondly, we show that in any other equilibrium, the firm's ad strategy $\mathcal{R}$ is such that $\mathcal{R} \subseteq \bar{\mathcal{R}}$ (which implies $\mathcal{T} \subseteq \bar{\mathcal{T}}$). Fix any such equilibrium and the respective firm's strategy.

[11] Note that the UCS may or may not include the second point $\{\underline{R}, \bar{R}\}$ in this case.

By Lemma 1, $\mathcal{T}(l) = \emptyset = \bar{\mathcal{T}}(l)$, hence we only need to show the claim for $y = h$. The interim belief of Bayesian consumer i who does not receive any ad must be $p_i^1 \in [\beta(p^0), p^0]$ in this equilibrium. It cannot be lower because the worst private information the firm can have is $y = l$, and observing such signal directly would lead the consumer to have belief $\beta(p^0)$, hence no firm's action can drop consumer i's belief below $\beta(p^0)$. On the other hand, p_i^1 cannot be higher than p^0 because the firm never advertises signal $y = l$ (see Lemma 1) – so p_i^1 after no ad is given by

$$p_i^1 = \frac{p^0\left(\rho(1-\sigma_i)+1-\rho\right)}{p^0\left(\rho(1-\sigma_i)+1-\rho\right)+(1-p^0)\left(\rho+(1-\rho)(1-\sigma_i)\right)} < p^0,$$

$$\text{where } \sigma_i \equiv \int_{\mathcal{T}\in 2^{\mathcal{I}}} \mathbb{I}\{i \in \mathcal{T}\} d\tau(\mathcal{T}|h) \in [0,1]$$

is the probability that signal $y = h$ is transmitted to i, and $\mathbb{I}\{\cdot\}$ is the indicator function. It follows that the ad effect A_i on i is such that $A_i \in [\mathcal{A}(R_i), \hat{\mathcal{A}}(R_i)]$. Advertising to i is profitable if and only if $A_i > c$, hence only if $\hat{\mathcal{A}}(R_i) > c$, hence only if $i \in \bar{\mathcal{T}}(h)$.

Finally, we need to show that $\bar{\mathcal{R}}$ satisfies the properties listed in the Theorem. Recall that $\bar{\mathcal{R}}(h) = \{R_i|\hat{\mathcal{A}}(R_i) > c\}$, while for the cursed equilibrium we have $\mathcal{R}_c(h) = \{R_i|\mathcal{A}(R_i) > c\}$. The two are defined in terms of $\hat{\mathcal{A}}(R_i) = \mathcal{D}_f(\alpha(p^0), R_i) - \mathcal{D}_f(\beta(p^0), R_i)$ and $\mathcal{A}(R_i) = \mathcal{D}_f(\alpha(p^0), R_i) - \mathcal{D}_f(p^0, R_i)$ respectively. Consider now a ficticious ("prime") world in which all players share a common prior $p' \equiv \beta(p^0)$ and signal informativeness is $\rho' \equiv \frac{\rho^2}{\rho^2+(1-\rho)^2}$. Then $\alpha'(p') = \alpha(p^0)$. But then for any i, $\hat{\mathcal{A}}(R_i) = \mathcal{A}'(R_i)$, where $\mathcal{A}'(R)$ is calculated same as $\mathcal{A}(R)$, except with ρ' instead of ρ. Consequently, Theorem 2, as applied to this ficticious world, implies that $\mathcal{R}'_c(h) \equiv \{R_i|\mathcal{A}'(R_i) > c\}$ satisfies all the required properties. But for a given i, we have $R_i \in \mathcal{R}'_c(h) \iff R_i \in \bar{\mathcal{R}}(h)$, meaning that $\bar{\mathcal{R}}(h)$ satisfies all required properties as well. $\square$