Targeted advertising and consumer information*

Job Market Paper.

Sébastien Broos†

November 2016

Abstract

Using ever-increasing amounts of data, firms are able to link the valuations of consumers with the information they possess about a product. We analyse the impact of that new ability on the advertising strategy of a monopolist. If there is a positive correlation between consumers' valuations and their information then, in contrast to the literature, better targeting often reduces prices. A lower price does not necessarily lead to a higher consumer surplus: some high-valuation/high-information consumers may stop purchasing because they stop receiving ads. Because of the interplay between the targeting and the pricing strategies, consumer surplus and welfare may be increasing in the advertising cost. Finally, we highlight that the link between valuation and information poses new problems for the estimation of returns to advertising.

Keywords: targeted advertising, information, e-commerce.

JEL classification: D42, D80, L12, M37

1 Introduction

If a major role of advertising is to convey information about products to consumers, then an obvious question comes to mind: does advertising work equally well on all consumers and, in particular, does it work on consumers who are already well informed about a product? The unsurprising answer is no, it does not work on informed consumers (Ackerberg, 2001; Blake et al., 2015), but the surprising fact is that the economics literature has not yet considered the implications of this finding. Instead, it has mostly argued that if firms have

*I would like to thank participants from the 2nd IO in the Digital Economy workshop in Louvain-la-Neuve, the Augustin Cournot Doctoral Days in Strasbourg, the Annual Conference of the Leibniz Science-Campus 2016 MACCI, EARIE Lisbon and seminar participants at Télécom ParisTech and Tilburg University. I also thank Axel Gautier, Julien Jacqmin, Anna D'Annunzio, Jorge Marcos Ramos, Jonas Teusch, Florian Schuett and Bertrand Koebel for their helpful comments. This research was funded through the ARC grant for Concerted Research Actions, financed by the French-speaking Community of Belgium.

†University of Liege, HEC Liege, Liege Competition and Innovation Institute, sbroos@ulg.ac.be

the ability to target their ads, they should aim for the high-valuation consumers so that price can be raised.

It is now possible for firms to incorporate the heterogeneity of consumer information in their advertising strategies. Indeed, the amount of data advertisers possess about consumers is massive. Facebook processes hundreds of millions of photos, likes and messages each day[1] and offers 98 targeting options to advertisers such as age, income, location, credit rating, political orientation, etc.[2] Bluekai, a data broker, has data about more than one billion consumers with, on average, fifty attributes per individual.[3] More practically, firms such as eBay are able to proxy consumer information through the recency and frequency of purchases (Blake et al., 2015). This wealth of data implies a wealth of targeting possibilities. We contend that in addition to *valuation*, firms' advertising strategies should also incorporate *information*.

Consider the following example. We observe two consumers, Jane and Joe. Jane is a fan of best-selling author Nassim Taleb[4] and regularly reads blogs and forums about Taleb's work. Joe is a casual reader who does not have strong preferences but sometimes reads essays. Taleb is on the brink of publishing a new book. His publisher, using the service of an advertiser or of a data broker, learns that Jane is a high-valuation/high-information consumer while Joe is a middle-valuation/low-information consumer. How should the publisher target its ads and price the product? The pricing trade-off is classical: setting a high price but sell only to Jane or setting a low price but sell to Jane and Joe. The targeting trade-off is more interesting. On the one hand, sending an ad to Jane seems to be a waste of money because she is likely informed about the release anyway. On the other hand, she has a high-valuation and would be willing to pay a high price. Hence, losing her would be costly. This trade-off is, to the best of our knowledge, new and has important ramifications.

In the main text, as in the example, we assume that there is a positive relationship between the information that consumers possess and their valuation for the good, at least for some consumers. We examine how a monopoly firm which knows this relationship should target its ads and the impact of that type of targeting on price, consumer surplus and welfare.

To that end, we compare three strategies, which are actually the only solutions to the profit-maximization problem we will examine. We call targeting "valuation targeting" if a

[1] Share Lab (2016), "Facebook algorithmic factory(1)", available at `https://labs.rs/en/facebook-algorithmic-factory-immaterial-labour-and-data-harvesting/`.

[2] The Washington Post (2016), "98 personal data points that Facebook uses to target ads to you", available at `https://www.washingtonpost.com/news/the-intersect/wp/2016/08/19/98-personal-data-points-that-facebook-uses-to-target-ads-to-you/`.

[3] The Economist (2014), "Getting to know you", available et `http://www.economist.com/news/special-report/21615871-everything-people-do-online-avidly-followed-advertisers-and-third-party`.

[4] Taleb is the author of "The Black Swan".

firm decides that consumer information is useless, i.e. if it targets high-valuation consumers only. This is a proxy for the way targeting is typically modelled in the literature. On the other hand, we call the strategy that takes into account consumers' information "information targeting", e.g. Taleb's publisher does not send an ad to Jane because she will buy his new book anyway. The major advantage of information targeting is that it allows the firm to send ads to consumers who have a high valuation *and* a low information, thereby avoiding a waste of ads without losing too many consumers. Because of the positive relationship between valuation and information, consumers who stop receiving ads under information targeting are those with a high valuation. Finally, the firm can abstain from advertising completely, for instance in a market with perfect information.

So long as (i) advertising is not too costly nor too cheap and (ii) there is a sufficient discrepancy between the levels of information of different consumers, then information targeting is the equilibrium strategy of the firm. In that case, some high-valuation consumers do not receive ads.

Compared to valuation targeting, some high-valuation consumers stop purchasing the good because they have not received ads and thus, are not informed any more. This reduces demand and *always* causes the monopolist to lower its price. This is one of our main results: better targeting, in the sense of being able to link information and valuation, may lead to a lower price. We use the conditional because compared to no advertising, the information targeting price may be lower or higher.

Interestingly, a lower price does not necessarily imply a higher consumer surplus. This is in line with the literature on targeted advertising but for opposite reasons. Typically, in a monopoly setting, consumer surplus will decrease because the price increases with targeting but, on the other hand, better targeting leads to more transactions with high-valuation consumers (who, by definition, obtain the most surplus from transaction). In our model, information targeting may reduce the number of transactions and, importantly, the consumers who stop buying because they are not targeted any more are exactly those who would have obtained the most surplus from the purchase. This may lead to a decrease in consumer surplus. The impact of information targeting on welfare is ambiguous for the same reasons.

That the monopolist chooses not to target some high-valuation consumers leads to a non-monotone relationship between the advertising cost and the price. This result also holds for consumer surplus and welfare: a higher advertising cost may, through the interplay of the pricing and the advertising strategies, lead to a higher consumer surplus and welfare.

Our theoretical model thus emphasizes the important role of information on market outcomes. This also matters for empirical research and is in line with Ackerberg (2001) and Blake et al. (2015) who show that returns to advertising can not be estimated properly without taking into account consumers' information. We highlight that the link between valuation and information poses additional problems for empirical research. In particular, the use of experiments to compute returns to advertising should take into account that, to be able to compare two markets, consumers should not only have the same valuations but also the same link between valuation and information. Otherwise, returns will be biased in unpredictable directions.

The next Section discusses the related literature. Section 3 sets up the basic model and the major hypotheses. Sections 4, 5 and 6 examine, respectively, the conditions under which information targeting is optimal, its impact on prices, consumer surplus and welfare, and the first-best. Section 7 covers possible issues empirical researchers may face if there is a link between valuation and information. In Section 8, we discuss our main assumptions and analyse some extensions. Section 9 concludes.

2 Related literature

First and foremost, this paper is part of the literature on targeted advertising (Athey and Gans, 2010; Bergemann and Bonatti, 2011; Brahim et al., 2011; Esteban et al., 2001; Esteves and Resende, 2016) and in particular that which is concerned with exogeneously informed consumers (Meurer and Stahl, 1994; Xu et al., 2012).[5] Typically, this literature finds that in the absence of competitive constraints, targeting is a way for firms to reach consumers with a high valuation and therefore to increase their market power. As a result, a monopolist that can target its ads will increase its price (Esteban et al., 2001; Hernandez-Garcia, 1997). Our main contribution to this literature is to add (i) heterogeneity to consumer information and (ii) the possibility for advertisers to include this heterogeneous information in their advertising strategy. These additions may lead to vastly different results.

Our paper is also linked to Johnson and Myatt (2006) and, more globally, to the literature on information provision (Bar-Isaac et al., 2010; Lewis and Sappington, 1994; Ottaviani and Prat, 2001; Saak, 2008). More precisely, Johnson and Myatt (2006) argue that advertising can contain two types of information: hype or real. Information is *hype* if a consumer learns the product's existence, price, availability and any objective quality but *not* his own valuation, i.e. from this information, all consumers infer the same common valuation. If information is *real*, a consumer learns his subjective preference for the product: his own valuation. The provision of hype information leads to an outward shift of the demand curve while the provision of real information leads to a rotation of the demand curve: some consumers are willing to pay more because their valuation is higher than expected, but others learn that the product is not a good fit for them. A major difference between the

[5] This is also considered in an extension in Iyer et al. (2005). See Bagwell (2007) and Renault (2016) for literature reviews on advertising with mentions/sections on targeted advertising.

literatures on targeted advertising and on information provision is that the former assumes that without advertising, consumers do not buy and thus that information is hype (there is a demand shift), while the latter assumes that consumers have a common ex-ante valuation and, after learning their true valuation, some may stop buying (there is a demand rotation). Information in that case is thus real. Both assumptions can be accommodated in our model and do not change the results. The details are discussed in Section 8.3.

Finally, this paper is related to the recent empirical literature on the measurement of advertising effectiveness. This literature is discussed more in detail in Section 7.

3 The baseline model

3.1 Consumers

There is a unit mass of consumers, each of whom has use for maximum one unit of a product. Consumers have valuations v distributed on $[0, 1]$ according to log-concave density function $f(v)$ and the corresponding log-concave cumulative distribution function $F(v)$.[6] They do not incur any nuisance cost from receiving ads.

Consumers are imperfectly informed about the good: they may not know that the good exists. There are two possible sources of information. First, there is an "information function" $g(v)$ which links the valuations of consumers and the probability they have of being informed about the good and its price without receiving ads, i.e. a consumer with valuation v has a probability $g(v)$ of being informed about the good and to learn the price and his (true) valuation. We assume that $g(v)$ is invertible and log-concave. Second, information can be transmitted through advertising: a consumer who receives an ad is perfectly informed.

There are two important questions that must be considered regarding consumer information. First, how uninformed are uninformed consumers? In the main case, they have no information about the product: they do not know its existence nor its characteristics. Therefore, uninformed consumers do not buy. Formally, this is equivalent to assuming that they have a valuation of 0. As discussed in Section 2, this is a common assumption in most (targeted) advertising models. In Section 8.3, uninformed consumers have a common ex-ante valuation $\tilde{v} \in (0, 1]$. It does not change the nature of our conclusions although it may affect under which circumstances they hold.

Second, what is the structure of information and in particular, what is the link between the information function and the valuations of consumers? We assume that the valuation of a consumer and its probability to be informed are uncorrelated or positively correlated,

[6] Many standard distributions such as the uniform or the normal distribution respect log-concavity. For more on log-concavity, see Bagnoli and Bergstrom (2005).

i.e. $g'(v) \geq 0$. What this means exactly is the topic of Section 4. We discuss the case of a negative correlation in Section 8.2.

3.2 The firm

A monopolist sells one good. It is produced at a constant marginal cost which is normalized, without loss of generality, to 0. It knows the information function and the distribution of the valuation but not necessarily the valuation of each consumer. It is still able to target ads very precisely, for instance because ads are sent through an intermediary which knows the valuation of each consumer, but cannot reveal it for privacy reasons. The firm is therefore unable to price discriminate and sets a uniform price p. Allowing the firm to have more precise information about consumers' valuations and to price discriminate does not change the nature of our conclusions and is discussed in details in Section 8.

Advertising costs a per consumer[7] and is perfectly informative: after receiving an ad, consumers know their valuation v with certainty. Advertising is also perfectly precise: an ad intended for consumer v will reach consumer v. The monopolist has three advertising strategies, which are subsumed in the following profit-maximization problem:

$$\max_{p,\bar{v}} \left[(p-a) \int_p^{\bar{v}} f(v)dv + p \int_{\bar{v}}^1 f(v)g(v)dv \right]. \tag{1}$$

There are two strategic variables: the price p and the upper targeting bound $\bar{v}$. The price is a double instrument. It is what buyers pay but it is also the lower targeting bound: consumers with $v \in [p, \bar{v}]$ receive ads. Given p, the firm never sends ads to consumers with $v < p$ because they never buy. Because $g'(v) \geq 0$, it may be that high-valuation consumers have a high probability to be informed. Hence, the firm may decide that it is not necessary to send ads to these consumers. Formally, it can set an upper targeting bound $\bar{v}$ such that consumers with $v \in [\bar{v}, 1]$ do *not* receive ads.

Note that setting a lower bound $\underline{v}$ such that consumers with $v \in [p,\underline{v}]$ do not receive ads is never an optimal strategy. Indeed, the information function is non-decreasing in v. Therefore, if it is optimal not to target consumers in $[p,\underline{v}]$, it cannot be optimal to target consumers with $v > \underline{v}$: they are (weakly) more likely to be informed than those with $v \in [p,\underline{v}]$. This point is proven formally in Appendix A.

Denote by p^* and $\bar{v}^*$ the price and upper targeting bound which solve the profit-maximization problem. There are three possibilities regarding $\bar{v}^*$: it is either interior, a lower corner solution ($\bar{v}^* = p^*$) or an upper corner solution ($\bar{v}^* = 1$). These solutions represent the three different advertising strategies available to the monopolist.

[7] The linearity of the cost function is similar to that in Iyer et al. (2005). The more usual convexity assumption implies decreasing returns to scale to advertising. There is no reason for this here since (i) consumers cannott be reached twice inadvertently and (ii) ads do not miss their targets.

First, the monopolist may indeed rely on the fact that high-valuation consumers are highly informed and not send ads to some of them. In that case, $p^* < \bar{v}^* < 1$: consumers with $v \in [p^*, \bar{v}^*)$ receive ads but those with $v \in [\bar{v}^*, 1]$ do not. While the former buy with certainty, the latter do not because some of them will not be informed. The trade-off for the firm is thus the following: saving the ad cost but losing some consumers. We call this strategy information targeting.

Second, the monopolist may choose to send ads to all consumers with $v > p^*$: $\bar{v}^* = 1$. This is costly but ensures that all potential buyers are informed. Information is thus ignored, in the sense that whether a particular consumer receives an ad does not depend on the information function. This is "valuation targeting".

Finally, the firm may choose to abstain from advertising ("no advertising"): $\bar{v}^* = p^*$. Because of the information function, some consumers are still informed and profit is positive.

Valuation targeting is of particular interest because it is a proxy for the way the economics literature has generally understood targeted advertising. Coming back to Equation 1, if targeting is on valuation only, the second term collapses and only the first remains. If there is no advertising, it is the first term which disappears and the second that remains. This implies that the information function only plays a role in the case of information targeting or if there is no advertising.

4 When is information targeting optimal?

The first question we want to tackle is that of the conditions under which information targeting is optimal, i.e. the conditions under which a monopolist maximizes its profit by choosing (i) $\bar{v}^* < 1$ and (ii) $p^* < \bar{v}^*$. In other words, when is $\bar{v}^*$ an interior solution? If the solution is indeed interior, the optimal price and upper targeting bound satisfy the following first-order conditions:[8]

$$\frac{\partial \Pi}{\partial p} = \int_{\bar{v}^*}^{1} f(v)g(v)dv + \int_{p^*}^{\bar{v}^*} f(v)dv - (p^* - a)f(p^*) = 0, \tag{2}$$

$$\frac{\partial \Pi}{\partial \bar{v}} = (p - a)f(\bar{v}^*) - pf(\bar{v}^*)g(\bar{v}^*) = 0 \Longleftrightarrow \bar{v}^* = g^{-1}\left(1 - \frac{a}{p^*}\right). \tag{3}$$

[8]For details on second-order conditions, see Appendix C. They may require an additional assumption on the information function and the distribution of valuation in certain cases.

To satisfy requirement (i) assuming that $p < \bar{v}$, we need that:[9]

$$\left.\frac{\partial \Pi}{\partial \bar{v}}\right|_{\bar{v}=1} = (p^* - a)f(1) - p^* f(1)g(1) < 0 \Longleftrightarrow 1 - \frac{a}{p^*} < g(1). \tag{4}$$

In words, consumers with the extreme valuation of 1 should have a sufficiently high probability to be informed without advertising, such that sending them ads would be a waste of resources. More generally, this can be extended to $\bar{v}$ smaller than any threshold valuation. This also rules out the case of $a = 0$. Obviously, if it is costless, there is no reason to abstain from advertising.

Requirement (ii) has two different implications. First,we have:

$$\left.\frac{\partial \Pi}{\partial p}\right|_{p=\bar{v}^*} = \int_{\bar{v}^*}^{1} f(v)g(v)dv - (\bar{v}^* - a)f(\bar{v}^*) < 0 \Longleftrightarrow a < \bar{v}^* - \frac{\int_{\bar{v}^*}^{1} f(v)g(v)dv}{f(\bar{v}^*)} \equiv a^*. \tag{5}$$

Because p^* is only implicitly defined, we can not go much further in interpreting this threshold.

Second, consumers should not be "too informed". To take an extreme case, if all consumers are perfectly informed, there is no reason to advertise. Formally:

$$\left.\frac{\partial \Pi}{\partial \bar{v}}\right|_{\bar{v}=p^*} = (p^* - a)f(p^*) - p^* f(p^*)g(p^*) > 0 \Longleftrightarrow g(p^*) < 1 - \frac{a}{p^*}. \tag{6}$$

If we combine this condition with that of requirement (i) (Equation 4), we obtain:

$$g(p^*) < 1 - \frac{a}{p^*} < g(1). \tag{7}$$

A necessary (but not sufficient) condition for this to hold is that consumers are heterogeneous with regards to the information they possess. Therefore, information functions such as $g(v) = k$ with $k \in [0, 1]$ do not lead to information targeting.[10] Intuitively, in that case, it would not make sense to have a targeting strategy based on information because all consumers have the same information. To use an analogy, there would be no reason to set different prices for different consumers if they all have the same valuation.

This condition can also be interpreted in terms of the advertising cost, it should neither

[9] Note that the case where $f(1) = 0$ is uninteresting. Indeed, in that case, because there are no consumers with $v = 1$, the valuation targeting scenario implies that the targeted consumers are those with $v \in [p, \tilde{v}]$ where $\tilde{v}$ is the highest v with strictly positive density. Therefore, the corner solution is not defined as $\bar{v} = 1$ but as $\bar{v} = \tilde{v}$. This does not change any of the conclusions.

[10] Moreover, technically, a function of this form is not invertible and thus violates our invertibility assumption.

be too high nor too low:

$$p^*(1 - g(1)) < a < p^*(1 - g(p^*)). \tag{8}$$

If the advertising cost is too low, it is better to target on valuation only. If it is too high, it is better not to advertise. The following Proposition summarizes these conditions.

Proposition 1. *If:*

1. $g(p^*) < 1 - \frac{a}{p^*} < g(1)$.
2. $0 < a < a^*$.

Then, information targeting is optimal ($p^* < \bar{v}^* < 1$). Otherwise, $\bar{v}^*$ is a corner solution and information targeting is not optimal.

There are many sets of sufficient conditions that satisfy this Proposition. Let us take two examples to clarify it.

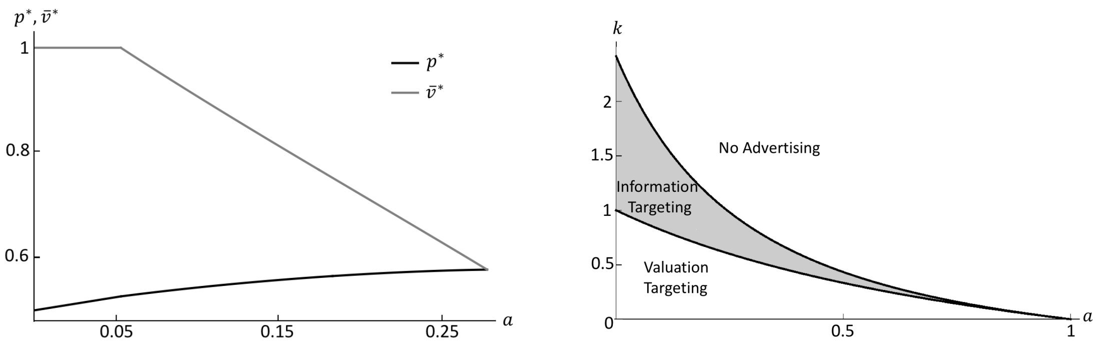
Example 1. Suppose that $v \sim U[0,1]$ and that $g(v) = 0.9v$. Then, it can be shown that if $a \in [0, 0.0526)$ valuation targeting is optimal, if $a \in [0.0526, 0.2773)$ information targeting is optimal and finally, if $a \geq 0.2773$, no advertising occurs. This case is illustrated on Figure 1a.

Example 2. Suppose that $v \sim U[0,1]$ and that $g(v) = kv$ with $0 \leq k \leq +\infty$. The interesting feature of this example is that a higher k implies that more consumers are informed. Therefore, we can establish, albeit in a particular case, how the level of information influences the possibility to observe information targeting. From the first condition of Proposition 1,[11] we must have $kp^* < 1 - a/p^* < k$, which can be rewritten as $p^*(1-k) < a < p(1 - kp^*)$. This is depicted on Figure 1b. Information targeting is most likely when both the advertising cost and consumer information are not too high nor too low.

As shown by the examples, Proposition 1 applies to numerous situations. What is required is that advertising is not too cheap nor too costly and that some consumers are more informed than others if they have a higher valuation for the good.

Of course, the impact of consumer information is not limited to this set-up. For instance, if $g(v) = k$ with $k \in [0,1]$, the conditions of Proposition 1 are not respected and information targeting does not arise. Yet, information affects profit and targeting strategies but in a simpler way: the outside option of not advertising becomes more profitable. We now turn to the effects of information targeting on prices and welfare.

[11] The second condition is weaker in this case and can thus be ignored.



(a) Example 1: Price and targeting bound if $v \sim U[0,1]$ and $g(v) = 0.9v$.

(b) Example 2: Optimal strategies if $v \sim U[0,1]$ and $g(v) = kv$.

Figure 1: Examples 1 and 2.

5 The impact of information targeting on prices, consumer surplus and welfare

Suppose that the conditions of Proposition 1 are fulfilled. To be able to compare prices, we "force" the corner solutions as if information targeting was not available, for instance for technological reasons. For clarity, we use the following notation:

Notation. *We denote the optimal price by*

- p_i *for "information targeting" if* $\bar{v} = \bar{v}^*$.
- p_v *for "valuation targeting" given that* $\bar{v} = 1$.
- p_n *for "no advertising" given that* $\bar{v} = p$.

Therefore, p_n and p_v should be understood as the optimal prices given that the upper targeting bound is, respectively, forced to be equal to the price and to 1. The first-order conditions with regard to price are the following:[12]

$$\frac{\partial \Pi}{\partial p} = \int_{\bar{v}^*}^{1} f(v)g(v)dv + \int_{p_i}^{\bar{v}} f(v)dv - (p_i - a)f(p_i) = 0, \tag{9}$$

$$\left.\frac{\partial \Pi}{\partial p}\right|_{\bar{v}=1} = \int_{p_v}^{1} f(v)dv - (p_v - a)f(p_v) = 0, \tag{10}$$

$$\left.\frac{\partial \Pi}{\partial p}\right|_{\bar{v}=p} = \int_{p_n}^{1} f(v)g(v)dv - p_n f(p_n)g(p_n) = 0. \tag{11}$$

[12]In the case of valuation targeting and no advertising, the second-order conditions with regard to prices can safely be ignored because (i) the product of two log-concave functions is a log-concave function and (ii) log-concavity implies an increasing hazard rate. Hence, the second-order conditions are satisfied, see Tirole (1988). For the second-order conditions related to the information advertising case, see Appendix C.

First, information targeting always lowers the price compared to valuation targeting: $p_i < p_v$. In other words, when the monopolist targets in function of both valuation *and* information, it sets a lower price than when it only takes valuation into account. The reason is simply that $\bar{v}^* < 1$. To see why, suppose that the two prices are equal and that they are lowered by a small amount. We have the two traditional effects of a change in price in a monopoly setting: (i) the firm loses the price differential on consumers to whom it sells (the "margin effect") but (ii) it gains some additional consumers (the "demand effect"). That demand effect is the same whether the firm uses information or valuation targeting because the involved consumers receive ads and are thus perfectly informed. However, the margin effect is different. Indeed, the lost profit is computed on consumers who actually buy. For consumers between p and $\bar{v}^*$, it does not make a difference: they all buy regardless of the targeting strategy. But for those between, $\bar{v}^*$ and 1, the effect is smaller under information targeting: all consumers buy under valuation targeting but only those who are informed make a purchase under information targeting. Formally, this can be seen by rewriting the first-order condition in the valuation targeting case as follows:

$$\left.\frac{\partial \Pi}{\partial p}\right|_{\bar{v}=1} = \int_{\bar{v}^*}^{1} f(v)dv + \int_{p_v}^{\bar{v}^*} f(v)dv - (p_v - a)f(p_v) = 0. \tag{12}$$

This conclusion contrasts strongly with the literature on targeted advertising. In general, targeting is seen as a device that enables firms to reach consumers with a high valuation so that prices can be raised. Here, better targeting lowers the price because the firm is willing to trade a reduction in demand against a reduction in cost.

In contrast, there is no *a priori* ranking of p_n and p_i. To see this, compare again the first-order conditions. Clearly, the first term in Equation 11 is smaller than the sum of the first two terms in Equation 9: the margin effect under no advertising is weaker because more consumers do not buy due to a lack of information. However, the second term in equation 11 is also smaller than the last term in Equation 9.[13] On the one hand, demand diminishes less under no advertising because consumers do not receive ads and hence, only those who are informed through the information function are lost (those who are uninformed do not buy anyway). On the other hand, a lost consumer induces a bigger loss under no advertising than under information targeting because the advertising cost is not incurred if consumers do not receive ads.

This is a somewhat strange result because p_n is constant in a, p_i is not, and we know that for some high a (at worst, $a = 1$), no advertising is always the optimal strategy and thus $p_i = p_n$. Hence, we would expect a clear-cut ranking. We do not obtain it because p_i is not necessarily monotonically increasing in a. The intuition (see Appendix B for details)

[13] We have that $pf(p)g(p) < (p-a)f(p) \iff g(p) < 1 - \frac{a}{p} \iff p < \bar{v}^*$.

is that a change in a has a double effect on p_i:

$$\frac{\mathrm{d}p_i}{\mathrm{d}a} = \frac{\partial p_i}{\partial a} + \frac{\partial p_i}{\partial \bar{v}^*}\frac{\partial \bar{v}^*}{\partial a}. \tag{13}$$

The direct effect leads to an increase in price to compensate the increase in cost. But there is also an indirect effect. A hike in a leads to a lower $\bar{v}^*$, which in turn pressures p_i downwards because of the lower margin effect (the same as when we compared p_i and p_v).

In some cases, the sum of these two effects is positive for all a, for instance if $v \sim U[0,1]$ and $g(v) = v$, but in others it can be negative, for example if $v \sim$ Kumaraswamy$(2,5)$ and $g(v) = v$.[14]

Because, as explained earlier, p_n is constant in a and $p_n = p_i$ for sufficiently high a, a necessary condition to obtain $p_n < p_i$ is that p_i should *not* be monotonically increasing in a for all a.

We summarize these results in the following Proposition.

Proposition 2. *Information targeting does not necessarily increase prices:*

1. $p_i < p_v$.

2. *If* p_i *is strictly monotonically increasing in* a*,* $p_i < p_n$.

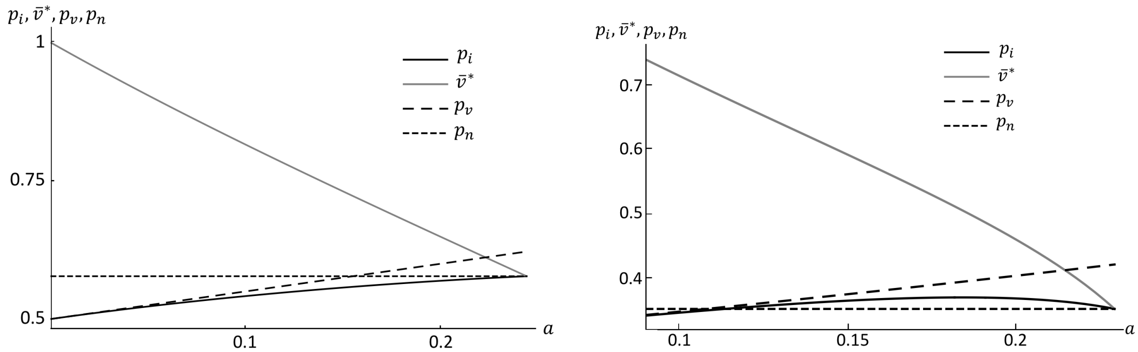
3. *If* p_i *is not strictly monotonically increasing in* a*, it may be that* $p_i \geq p_n$.

Proof. See Appendix B. ■

The impact of information targeting on price therefore depends highly on the counterfactual situation. Moreover, this Proposition contrasts strongly with the literature on targeted advertising. In general, targeting is seen as a device that enables firms, especially in a monopolistic context (Esteban et al., 2001; Hernandez-Garcia, 1997), to reach consumers with a high valuation so that prices can be raised. The reason for that difference is the correlation between valuation and information. The Proposition is illustrated in Figure 2.

The effect of information targeting on consumer surplus also depends on the counter-

[14] The use of this relatively obscure distribution is due to the simplicity of its closed-form expression. It is closely related to the Beta distribution, which is more commonly used in economics. Given enough computing power, there is no limitation on using other, more common, distributions. For more details on the Kumaraswamy distribution, see Appendix E or, for a more complete treatment, Jones (2009).



(a) $v \sim U[0,1]$ and $g(v) = v$.

(b) $v \sim$ Kumaraswamy$(2,5)$ and $g(v) = v$.

Figure 2: Prices and targeting bound with different valuation distributions.

Notes: so that Figure 2b is as clear as possible, we only show the range where $a \in [0.09, 0.2288]$. If $a > 0.2288$, information targeting is not optimal any more.

factual. Formally, consumer surplus is:

$$CS_i = \int_{p_i}^{\bar{v}^*} f(v)(v - p_i)dv + \int_{\bar{v}^*}^{1} f(v)g(v)(v - p_i)dv, \tag{14}$$

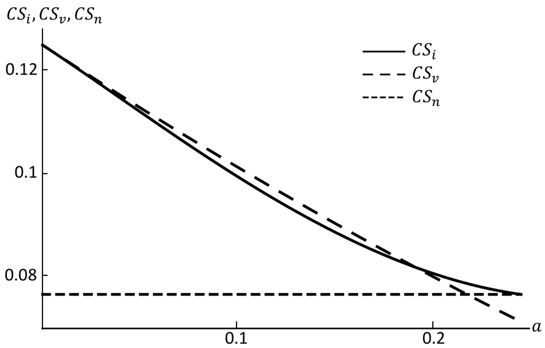
$$CS_v = \int_{p_v}^{1} f(v)(v - p_v)dv = \int_{p_v}^{\bar{v}^*} f(v)(v - p_v)dv + \int_{\bar{v}^*}^{1} f(v)(v - p_v)dv, \tag{15}$$

$$CS_n = \int_{p_n}^{1} f(v)g(v)(v - p_n)dv. \tag{16}$$

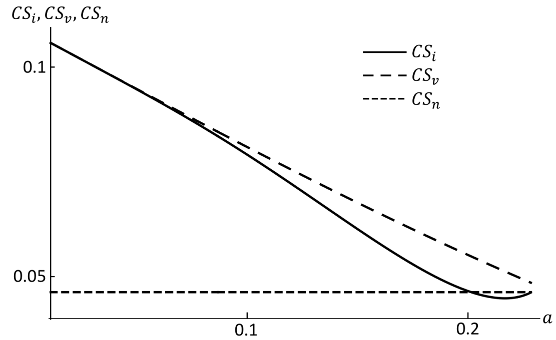
A move from valuation to information targeting impacts consumers in two ways. There is a price effect which is unambiguously positive: consumers pay a lower price. But there is also a demand effect which is ambiguous: more consumers buy because of the lower price but other consumers stop purchasing because $\bar{v}^* < 1$. The identity of these consumers is important: those who stop buying have $v \in [\bar{v}^*, 1]$ while the new consumers have $v \in [p_i, p_v]$. Clearly, the former have a much stronger (negative) effect: they are exactly those who would benefit the most from a purchase. The monopolist does not internalize this at all since it can only set a single price.[15] The global impact of information targeting on consumer surplus is therefore ambiguous. A lower price is not necessarily synonymous with a higher consumer surplus.

The ambiguity remains if we compare information targeting with no advertising but for different reasons. There is always a positive demand effect because at least a few consumers receive ads and become informed. If $p_i < p_n$, demand increases even more, all consumers pay less and consumer surplus must increase. However, if $p_n < p_i$, demand decreases and consumers pay more: the impact of information targeting is ambiguous.

[15] See Section 8.1 for the case of perfect price discrimination.



(a) $v \sim U[0,1]$ and $g(v) = v$.



(b) $v \sim$ Kumaraswamy$(2, 5)$ and $g(v) = v$.

Figure 3: Consumer surplus with different valuation distributions.

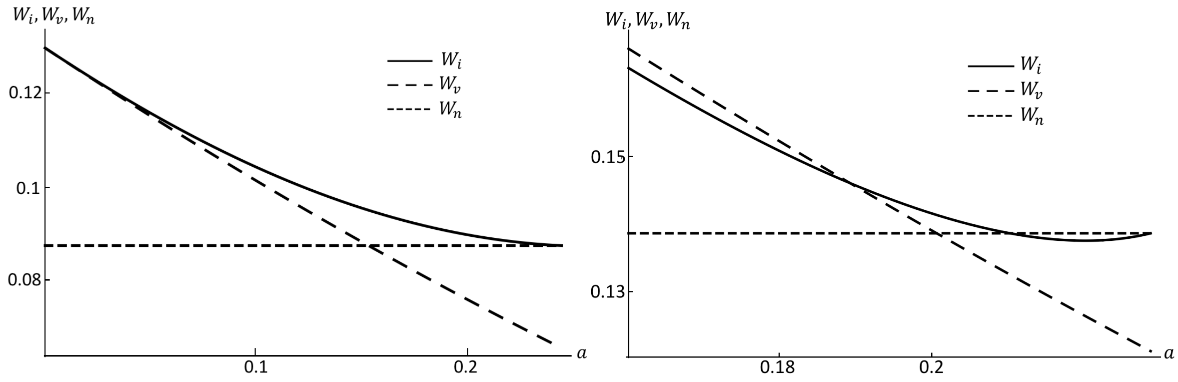
The non-monotonicity of p_i carries on to consumer surplus: if p_i decreases in a, consumer surplus may increase in the advertising cost. These effects are illustrated in Figure 3 which displays a case where consumer surplus is monotonically decreasing in a everywhere (Figure 3a) and a case where is it is monotonically increasing in some parameter range (Figure 3b).

The ambiguous effects of information targeting on consumer surplus also affect welfare: W_i, W_v, W_n, defined as the sum of profit and consumer surplus in each case. It is therefore unclear *a priori* whether the switch to information targeting increases or decreases total surplus despite a possible lower price. The only exception is if p_i is monotonically increasing for all a. Then, welfare under information targeting is higher than under no advertising because both consumer surplus and profit are higher.

The non-monotonicity result also carries on to welfare. Generally, welfare decreases in a, but it may also rise (see Figure 4) if consumer surplus increases sufficiently so that the loss in profit is more than compensated. This result is possible not only because of the non-monotonicity of p_i but also because of the availability of two strategic variables: when a increases, the monopolist's profit is reduced (weakly) less than if it could only choose the price. These results are summarized in the following Proposition.

Proposition 3. *The impact of information targeting on consumer surplus and welfare is ambiguous except if p_i is monotonically increasing for all a, then $CS_i > CS_n$ and $W_i > W_n$. A lower price guarantees neither a higher consumer surplus nor a higher welfare. Under information targeting, if p_i is not monotonically increasing for all a, consumer surplus and welfare are not necessarily monotonically decreasing in a.*

Proof. In the text and in the examples in Figures 3 and 4. ■



(a) $v \sim U[0,1]$ and $g(v) = v$.

(b) $v \sim \text{Kumaraswamy}(2,5)$ and $g(v) = v$.

Figure 4: Welfare with different valuation distributions.

Note: so that the different effects are displayed clearly, we have restricted the parameter range to $a \in [0.16, 0.2288]$ on Figure 4b

6 First-best

Finally, we would like to compare the information targeting outcome with what a social planner would choose. The social planner solves the following problem:

$$\max_{p,\bar{v}} W. \tag{17}$$

Where

$$\begin{aligned} W = CS + \Pi = (p-a)\int_p^{\bar{v}} f(v)dv + p\int_{\bar{v}}^1 f(v)g(v)dv \\ + \int_p^{\bar{v}} f(v)(v-p)dv + \int_{\bar{v}}^1 f(v)g(v)(v-p)dv. \end{aligned} \tag{18}$$

The first-order conditions, after some algebra, are:

$$\frac{\partial W}{\partial p} = (p-a)f(p) = 0, \tag{19}$$

$$\frac{\partial W}{\partial \bar{v}} = \bar{v}(1-g(\bar{v})) - a = 0. \tag{20}$$

Let us call p_w and $\bar{v}_w$ the price and targeting bound that maximize welfare. From the first-order conditions, $p_w = a$ and we get the usual result that price should be equal to marginal cost.

Unsurprisingly, we find that so long as information targeting is optimal for the monopolist, $\bar{v}_w > \bar{v}^*$. The social planner takes into account that consumers with a high v have a high impact on consumer surplus and, accordingly, targets more of them. If $1 - a < g(1)$,[16] even a social planner does not use valuation targeting: $\bar{v}_w < 1$. It is socially inefficient to waste money on ads to consumers who are very likely to be informed even if it implies that there is a (small) probability that they do not purchase.

Proposition 4. *$p_w = a < p_i$ and, if information targeting is optimal for the firm, $\bar{v}^* < \bar{v}_w \leq 1$. A monopolist using information targeting prices too high and targets too narrowly.*

Proof. The price result is obtained directly from the first-order condition. For the second result, we have that:

$$\left.\frac{\partial W}{\partial \bar{v}}\right|_{\bar{v}=\bar{v}^*} = \bar{v}^* \left(1 - 1 + \frac{a}{p_i}\right) - a = a\left(\frac{\bar{v}^*}{p_i} - 1\right) > 0. \tag{21}$$

Given that $p_i < \bar{v}^*$, this expression is strictly positive. ■

This is reminiscent of the classical result of Shapiro (1980): a monopolist advertises too little because it can not extract the surplus of high valuation consumers who, here, are most likely informed anyway.

Interestingly, because $\bar{v}_w < 1$, a is not incurred for all consumers and the welfare-maximizer allows the monopolist to make a positive profit on consumers to which ads are not sent.

7 Returns to advertising

Our results are closely linked to the literature on the estimation of returns to advertising. In particular, Blake et al. (2015) analyze the returns to search engine ads for eBay. They show that two categories of user queries should be distinguished. "Brand" queries are queries which contain the name of a brand, eBay in their paper. "Non-brand" queries are those that do not. For instance, "shoes ebay" would be a brand query but "used shoes" would not. Through a field experiment, they estimate that the returns for brand queries are not statistically different from zero. For non-brand queries, they find the same result if they consider all users. However, returns are positive for users who have not recently purchased on eBay or those who do not buy frequently. As the recency and the frequency of purchase respectively increases and decreases, the effect of advertising becomes more important. In a nutshell, the less informed consumers are, the more effective advertising is. In a different setting, ads for new yoghurts, Ackerberg (2001) reaches the same conclusion. Both Blake et al. (2015) and Ackerberg (2001) show that the information that consumers have should

[16]This comes from $\left.\frac{\partial W}{\partial \bar{v}}\right|_{\bar{v}=1} < 0$.

be taken into account and especially, that a major distinction has to be drawn between consumers who are likely informed and those who are likely not.

Second, Lewis and Rao (2015) explain that computing returns, even with field experiments with millions of observations, is extremely difficult. Any additional issue that may affect the estimation of returns is thus important. The difficulty stems from two issues. First, individual-level sales are very volatile. Second, there is a selection bias because of the targeted nature of the advertising, i.e. ads are not targeted to consumers through luck but because these consumers are selected as those to whom sending ads is the most profitable. In other words, the characteristics of targeted consumers are different from those of consumers who do not receive ads. Therefore, computing returns by comparing, *within the same market*, consumers who have seen ads and have purchased and those who have not seen ads and have purchased is misleading. One solution to this problem is to perform an experiment with two similar markets, one where advertising arises and one where it does not. Then, the revenues and costs from the two markets should be compared.

Using our model, we would like to highlight two additional difficulties an empirical researcher could face because of the link between information and valuation. Suppose that a researcher observes a firm and tries to compute the return to the firm's ads. Let us start by computing what the true returns are. There are two markets (A and B) and consumers are identical in both markets. On market A the firm uses information targeting while on market B it does not advertise, for instance because it is technologically impossible. This would be a perfect natural experiment. The true returns are:

$$\text{True returns} = \frac{p_i \int_{\bar{v}^*}^{1} f(v)g(v)dv + p_i \int_{p_i}^{\bar{v}^*} f(v)dv - p_n \int_{p_n}^{1} f(v)g(v)dv}{a \int_{p_i}^{\bar{v}^*} f(v)dv - 0} - 1. \tag{22}$$

The first difficulty stems from a possible lack of information about the firm's optimal strategy. It may be that the researcher does not know that the firm uses information targeting and that he believes valuation targeting is used. Then, a possible mistake is to compute revenues from market A as stemming only from consumers who see ads. Indeed, under valuation targeting, all purchasers see ads. This leads to an underestimation of returns, which would be computed in the following way:[17]

$$\frac{p_i \int_{p_i}^{\bar{v}^*} f(v)dv - p_n \int_{p_n}^{1} f(v)g(v)dv}{a \int_{p_i}^{\bar{v}^*} f(v)dv - 0} - 1. \tag{23}$$

[17] At first sight, it could be argued that an experienced researcher would never make such a mistake. However, not all returns estimations are performed by experienced researchers and even the selection bias is regularly ignored in some industry publications (see Lewis and Rao (2015), Blake et al. (2015) and Abraham (2008) for details).

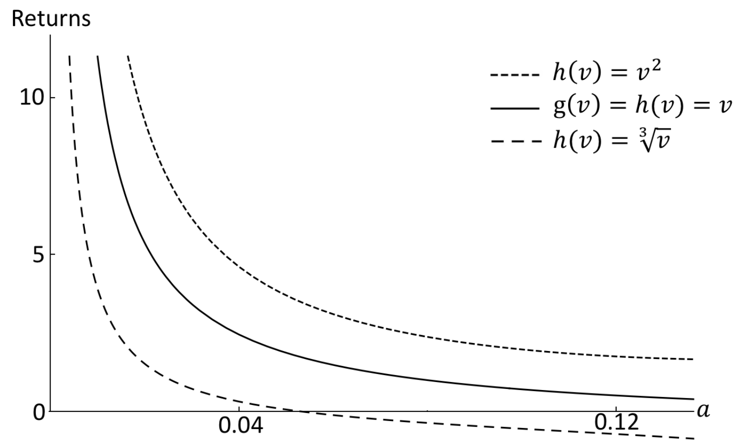


Figure 5: Returns to advertising for different $h(v)$ if $v \sim U[0,1]$ and $g(v) = v$.

Note: the range of a is chosen so as to make the Figure as clear as possible. Results are similar for other ranges of the parameter.

The second, more important, problem is that to overcome the selection bias, the researcher needs a B market where consumers are quite similar to those in the A market. This issue is compounded by the existence of the information function. Not only should consumers have a similar distribution of valuation but the information function should also be the same. Formally, suppose that on market A the information function is $g(v)$ but that on market B it is $h(v)$. Denote by p'_n the price on the B market. Then, returns are computed in the following way:

$$\frac{p_i \int_{\bar{v}^*}^{1} f(v)g(v)dv + p_i \int_{p_i}^{\bar{v}^*} f(v)dv - p'_n \int_{p'_n}^{1} f(v)h(v)dv}{a \int_{p_i}^{\bar{v}^*} f(v)dv - 0} - 1. \tag{24}$$

Unfortunately, it is not possible to know if this mistake leads to an over or an underestimation of the true returns. This mistake is illustrated in Figure 5 with two different $h(v)$ functions.

8 Discussion and extensions

We first discuss two assumptions informally and then analyse three extensions.

We have assumed that the monopolist knows the information function but it is not required to obtain our results. All we need is that the information the firm has be sufficiently fine-grained such that there are enough different groups of consumers with different levels of information. For instance, it may be enough to have information about two groups to obtain the results of the main model. Indeed, the incentive behind the results is still present: take advantage of consumer information to avoid sending costly ads to everyone. Many such examples can easily be built.[18] Because of discontinuities in the pricing and advertising strategies, there are two differences with the main model. First, the conditions under which information targeting arises are more stringent as a change in $\bar{v}$ implies that a *whole group* does not receive advertising. Second, all results involving strict inequalities now involve weak inequalities for a similar reason. Otherwise, all our main conclusions are qualitatively similar.

Another major assumption is the absence of competition: the monopolist knows that in case it does not advertise, the worst that can happen is that a few consumers do not buy. In a competitive setting, the implications of not advertising may be different and are highly dependent on the context and the modelling strategy. In particular, because in a monopoly setting the results hold under fairly general assumptions about the information function, we have been able to ignore exactly how it emerges without much damage. Under competition, the process behind the information function may be much more important. For instance, whether it informs consumers about only one product (consumers read a review) or about a product category (consumers read a detailed comparison article) should lead to different conclusions. A full characterization of a competitive setting is beyond the scope of this paper.

8.1 The firm's information and personalized pricing

The ability to price discriminate depends on the information the firm possesses about consumers and on the information that $g(v)$ provides to consumers. If the monopolist only knows the distribution of valuation, its only possibility to price discriminate is if it sends ads through an intermediary that can tailor the ads so that each consumer sees a different price. Advertising is much more advantageous than in the baseline model because the firm can extract all the surplus from consumers who see ads but can only set a uniform price for the others. This strategy only works so long as the uniform price is always higher than the personalized prices, which is always true. Indeed, because the information function is non-decreasing, it should always be (if any) consumers with a high valuation who do not receive ads and hence, the uniform price is higher than the personalized prices. Otherwise, it may be that consumers observe different prices through the ads and through the information function. We thus make the following assumptions:

[18]For instance, suppose that v is distributed according to pdf $f(v)$ and that the information function is $g(v)$ but that the firm only knows the following: in group 1, the consumer with the lowest valuation has $v = 1/3$ and his probability to be informed is 1/3. For the second group, these numbers are respectively 1/2 and 3/4. Each group has mass 1/2. Then, it can be shown that valuation and information targeting occur, respectively, if $0 \leq a < 0.0833$ and if $0.0833 \leq a < 0.0972$. If $0.0972 \leq a$, there is no advertising. In each case, the price is 1/3.

1. Consumers who receive ads face personalized prices (first-degree price discrimination).

2. Consumers who do not receive ads face a uniform price p.

Given that perfectly extracting the surplus of a consumer costs a, there are two reasons not to send him an ad: (i) because $v < a$ and (ii) because he is highly likely to be informed. The profit-maximization problem is the following:

$$\max_p \Pi^{ppd} \quad = \quad \max_p \left[\int_a^p (v-a) f(v) dv + p \int_p^1 f(v) g(v) dv \right]. \tag{25}$$

The uniform price plays simultaneously the role of price and of targeting bound (thereby making $\bar{v}$ useless). The first term represents profit made on consumers who face perfect price discrimination while the second term is profit on consumers who do not receive ads and face a uniform price.

Denote by $\tilde{p}$ the price that solves this maximization problem. We define valuation targeting as $\tilde{p} = 1$, i.e. all consumers face perfect price discrimination.[19] If information targeting is used, then $\tilde{p} < 1$.

Not advertising to a consumer is more costly than in the model with a uniform price. Instead of losing the price, the monopolist loses the entire valuation (minus the advertising cost) of the consumer. Therefore, the conditions under which information targeting arises are more stringent than in the uniform price scenario. The first-order condition is:

$$\frac{\partial \Pi^{ppd}}{\partial p} = (\tilde{p} - a) f(\tilde{p}) + \int_{\tilde{p}}^1 f(v) g(v) dv - \tilde{p} f(\tilde{p}) g(\tilde{p}) = 0. \tag{26}$$

And the requirement to have $\tilde{p} < 1$ is therefore:

$$\left. \frac{\partial \Pi^{ppd}}{\partial p} \right|_{p=1} < 0 \Longleftrightarrow 1 - a < g(1). \tag{27}$$

This is indeed strictly more stringent than the requirement to have $\bar{v}^* < 1$ in Section 4, which is $g(1) > 1 - a/p^*$. If this condition is satisfied, the firm increases its profit by using information targeting.

Consumers also benefit unambiguously from information targeting. Nothing changes for consumers who face perfect price discrimination in both cases: their surplus is zero. However, consumers with $v \in [\tilde{p}, 1]$ are bound to gain, despite the fact that some of them may stop purchasing because of the lack of ads. Under valuation targeting, they face perfect price discrimination and their surplus is zero. Under information targeting on the other

[19] We do not examine the case of no advertising as it would only arise under extremely high advertising costs. It can be shown that in this case, information targeting increases the price and has ambiguous effects on consumer surplus (a higher price but less informed consumers) and welfare.

hand, they face a uniform price and consumer surplus has to be positive. The fact that some consumers face $\tilde{p}$ instead of v is the analogue of the $p_i < p_v$ result in Section 5. Formally, we have that:

$$CS_v^{ppd} = 0, \tag{28}$$

$$CS_i^{ppd} = \int_{\tilde{p}}^{1} (v - \tilde{p}) f(v) g(v) dv > 0. \tag{29}$$

Since both profit and consumer surplus increase, information targeting also increases welfare. By taking into account information, the monopolist wastes less resources on advertising for consumers who are likely to be informed. Consumers benefit because some face a uniform price and obtain positive surplus. The conclusions are therefore slightly stronger than those in the main model.

This is not the only possible form of price discrimination. If the monopolist knows the valuation of each consumer instead of the distribution only, it can potentially set a personalized price for each consumer. Then, the precise role of the information function is important. We have assumed that it provides information on the existence and the price of the product, and instructs each consumer of his valuation. It is unclear what the price information would be in a model where all consumers face personalized prices. Possibly, it would imply that consumers who are informed through the information function may know the price quoted to others. This may be conducive to arbitrage and should make price discrimination more difficult to enforce.

8.2 $g'(v) \leq 0$

Although the full formal discussion is provided in the Appendix D, let us discuss the main conclusions if there is a negative correlation between information and valuation.

First, in contrast to the positive case, it is optimal to have a lower-bound $\underline{v}$ such that consumers with $v \in [p, \underline{v}]$ do not receive ads, but having an upper bound $\bar{v}$ is not optimal any more. This is intuitive: it is now consumers with a low valuation who have the highest probability to be informed.

Regarding prices, the ordering is less ambiguous than in the positive case: $p_n < p_i < p_v$. That $p_i < p_v$ is due to the lower margin effect under information targeting. Now however, one must also account for a smaller demand effect. Indeed, the consumers with $v \in [p, \underline{v}]$ all buy under valuation targeting but not under information targeting. Yet, it can be shown that the lower margin effect prevails. That there is a clear ranking between the price under information targeting and that under no advertising can be explained by the monotonicity of p_i with regards to a. Indeed, in the positive case, we had obtained a non-monotonicity result because of the opposite signs of the direct and indirect effects. Here, however, there is no direct effect of a on p_i: the only way for the monopolist to influence the quantity of ads is though $\bar{v}$ because consumers with $v \in [p, \underline{v}]$ do not receive ads. Hence, increasing the price does not change the quantity of ads served. Only the indirect effect remains and it is negative: p_i is decreasing in a. Therefore, since we know that for a very high a, the monopolist always chooses not to advertise ($p_n = p_i$), it must be that $p_n < p_i$. Intuitively, the demand effect is similar in both cases but the margin effect is stronger under information targeting than under no advertising.

Consumer surplus (and welfare) is affected ambiguously by information targeting. Comparing it to that under valuation targeting, we have the same trade-off as in the positive case: a lower price but less informed consumers. However, the impact of consumers who stop purchasing under information targeting is now much smaller because it is consumers with a low valuation who do not receive ads. Therefore, it is much more likely than in the positive scenario that consumer surplus rises when the monopolist switches from valuation targeting to information targeting. The trade-off in the case of no advertising is that information targeting leads to a higher price but to more informed consumers.

8.3 The valuation of uninformed consumers

A major assumption of the model of Section 3 is that uninformed consumers have absolutely no information about the good, i.e. it is as if their valuation was $v = 0$. What happens if, instead, they know about the existence of the good but do not know their valuation, i.e. they have a common ex-ante valuation, and advertising/information reveals their true valuation? For instance, (uninformed) consumers may be aware of the distribution of valuation and may have an ex-ante valuation equal to $\mathbb{E}(v)$. Another example would be for uninformed consumers to hold all valuations as equally likely and thus to have an ex-ante valuation equal to $1/2$. As highlighted in Section 2, this setting is closer to the literature on information provision.

Formally, suppose that uninformed consumers have valuation $\tilde{v} \in [0, 1]$. If $\tilde{v} = 0$, we are back to the main case. The monopolist has two strategies to maximize profit.

First, it can set $p \leq \tilde{v}$ and nearly all consumers buy: only those with $v < p$ *and* who are informed do not purchase the good. The downside is that p is constrained to be low. Here, advertising has no role to play. Sending ads to consumers with $v < \tilde{v}$ would decrease demand and increase costs while sending ads to consumers with $v \geq \tilde{v}$ is useless: they are already buying and nothing more can be extracted out of them. In equilibrium, it is likely that the constraint is binding ($p = \tilde{v}$). Indeed, setting a lower price would mean a lost margin on all buyers while only attracting some of those with $v < \tilde{v}$ *and* who are uninformed.

Second, the firm can set $p \geq \tilde{v}$ and in that case, we are back to the general framework, with the additional price constraint. Indeed, without being informed or receiving ads, all consumers abstain from buying. The fact that $\tilde{v} > 0$ does not change anything to prices and surplus so long as the price constraint is not binding.[20]

Formally, profit is respectively:

$$\Pi \quad = p \int_0^p f(v)(1 - g(v))dv + p \int_p^1 f(v)dv \quad \text{if} \quad p < \tilde{v}, \tag{30}$$

$$\Pi \quad = (p - a) \int_p^{\bar{v}} f(v)dv + p \int_{\bar{v}}^1 f(v)g(v)dv \quad \text{if} \quad p \geq \tilde{v}. \tag{31}$$

The interesting impact of having $\tilde{v} > 0$ is thus that it increases the profitability of not advertising. So long as the monopolist chooses to advertise, the nature of our conclusions does not change.

In the framework of Johnson and Myatt (2006), the advertising we consider is always a demand shifter because it is perfect: the firm never advertises to a consumer who may purchase if he is uninformed but does not if he is informed. This is a hint that big data and the ability to target consumers more precisely may tilt the impact of advertising towards demand shifts rather than demand rotations.

9 Conclusion

Big data is a new fact of economic life and has many ramifications. We have analysed one in detail: the ability to relate valuation with information and to target ads accordingly. This new type of advertising leads to counter-intuitive consequences that sometimes go against well-established results, the most important being that better targeting often reduces prices. Information advertising also has important ramifications for empirical researchers who, in particular, should be especially wary of different information functions in different markets.

These consequences show that both theoretical and empirical works can not assume information away. The simple information function examined in this article is just a start and more complex information structures should be explored. In particular, the process behind the information function should be explicitly modelled.

The relevance of this work for real-world practice should be of particular interest. For instance, why do we rarely see ads when we search the name of a brand on a search engine? Is it because of the mechanisms we have highlighted or are there other reasons? How do estimates of returns to advertising change if we take information into account? Are firms fooled by the current measures? Under which practical conditions should they target

[20] If the price constraint is binding but the monopolist still chooses to advertise, profit is reduced but consumer surplus (and welfare) is affected ambiguously. Indeed, while consumers suffer from the higher price, the constraint on p also has an impact on $\bar{v}$ and it is a priori unclear how these two effects interact.

information? All those questions require empirical work.

We hope this study will only be a first step in the study of the interplay between advertising and information.

10 References

Abraham, M. (2008): "The Off-Line Impact of Online Ads," *Harvard Business Review.*

Ackerberg, D. A. (2001): "Empirically distinguishing informative and prestige effects of advertising," *RAND Journal of Economics*, 316–333.

Athey, S. and J. S. Gans (2010): "The Impact of Targeting Technology on Advertising Markets and Media Competition," *The American Economic Review*, 608–613.

Bagnoli, M. and T. Bergstrom (2005): "Log-concave probability and its applications," *Economic theory*, 26, 445–469.

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

Bar-Isaac, H., G. Caruana, and V. Cuñat (2010): "Information gathering and marketing," *Journal of Economics & Management Strategy*, 19, 375–401.

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

Blake, T., C. Nosko, and S. Tadelis (2015): "Consumer heterogeneity and paid search effectiveness: a large-scale field experiment," *Econometrica*, 83, 155–174.

Brahim, N. B. E.-B., R. Lahmandi-Ayed, and D. Laussel (2011): "Is targeted advertising always beneficial?" *International Journal of Industrial Organization*, 29, 678–689.

Esteban, L., A. Gil, and J. M. Hernandez (2001): "Informative advertising and optimal targeting in a monopoly," *The Journal of Industrial Economics*, 49, 161–180.

Esteves, R.-B. and J. Resende (2016): "Competitive targeted advertising with price discrimination," *Marketing Science*, 35, 576–587.

Hernandez-Garcia, J. M. (1997): "Informative advertising, imperfect targeting and welfare," *Economics Letters*, 55, 131–137.

Iyer, G., D. Soberman, and J. M. Villas-Boas (2005): "The targeting of advertising," *Marketing Science*, 24, 461–476.

Johnson, J. P. and D. P. Myatt (2006): "On the simple economics of advertising, marketing, and product design," *The American Economic Review*, 96, 756–784.

Jones, M. (2009): "Kumaraswamy's distribution: A beta-type distribution with some tractability advantages," *Statistical Methodology*, 6, 70–81.

Lewis, R. and J. Rao (2015): "The Unfavorable Economics of Measuring the Returns to Advertising," *The Quarterly Journal of Economics*, 130, 1941–1973.

Lewis, T. and D. Sappington (1994): "Supplying Information to Facilitate Price Discrimination," *International Economic Review*, 35, 309–327.

Meurer, M. and D. O. Stahl (1994): "Informative advertising and product match," *International Journal of Industrial Organization*, 12, 1–19.

Ottaviani, M. and A. Prat (2001): "The value of public information in monopoly," *Econometrica*, 69, 1673–1683.

Renault, R. (2016): "Advertising in Markets," *Handbook of Media Economics, vol 1A*, 121–204.

Saak, A. E. (2008): "The value of private information in monopoly," *The Journal of Industrial Economics*, 56, 169–189.

Shapiro, C. (1980): "Advertising and welfare: Comment," *The Bell Journal of Economics*, 749–752.

Tirole, J. (1988): *The theory of industrial organization*, The MIT Press.

Xu, L., J. Chen, and A. Whinston (2012): "Effects of the presence of organic listing in search advertising," *Information Systems Research*, 23, 1284–1302.

Appendix

A A lower targeting bound

Having an additional bound $\underline{v} \in (p, \bar{v})$ is never a profit-maximizing strategy. If it is not optimal to send an ad to consumer v then it must be, because $g'(v) \geq 0$ that it is not optimal to send ads to consumers with $v' > v$. To see this formally, write the profit-maximization problem as:

$$\max_{p,\underline{v},\bar{v}} \Pi = \max_{p,\underline{v},\bar{v}} \left[p \int_{p}^{\underline{v}} f(v)g(v)dv + (p-a) \int_{\underline{v}}^{\bar{v}} f(v)dv + p \int_{\bar{v}}^{1} f(v)g(v)dv \right]. \tag{32}$$

The first-order conditions of this problem with regards to $\underline{v}$ and $\bar{v}$ are:

$$\frac{\partial \Pi}{\partial \underline{v}} = pf(\underline{v})g(\underline{v}) - f(\underline{v})(p-a) = 0, \tag{33}$$

$$\frac{\partial \Pi}{\partial \bar{v}} = f(\bar{v})(p-a) - pf(\bar{v})g(\bar{v}) = 0. \tag{34}$$

If both first-order conditions are satisfied, then we have:

$$\underline{v}^* = \bar{v}^* = g^{-1}\left(1 - \frac{a}{p^*}\right). \tag{35}$$

In that case, there is no advertising. Therefore, for advertising to be possible, we need a corner solution for at least one of the variables. If $\underline{v}^* = p^*$, we are back to our mainstream case and the point is proven. We show that the other corner solution, $\bar{v}^* = 1$ can not arise. If $\bar{v}^*$ is a corner solution, then, from Equation (34):

$$g(1) < 1 - \frac{a}{p^*}. \tag{36}$$

But from Equation (33), we also have that:

$$g(\underline{v}^*) = 1 - \frac{a}{p^*}. \tag{37}$$

This implies that $g(1) < g(\underline{v}^*)$. Given that $g'(v) \geq 0$, this is a contradiction and the point is proven.

B Proof of Proposition 2

For the first part of the result ($p_i < p_v$), we simply compare the first-order conditions (Equations 9 and 10). We have that $p_i < p_v$ if:

$$\int_{\bar{v}^*}^{1} f(v)g(v)dv + \int_{p}^{\bar{v}^*} f(v)dv - (p-a)f(p) < \int_{p}^{1} f(v)dv - (p-a)f(p), \tag{38}$$

$$\iff \int_{\bar{v}^*}^{1} f(v)g(v)dv + \int_{p}^{\bar{v}^*} f(v)dv < \int_{p}^{1} f(v)dv, \tag{39}$$

$$\iff 0 < \int_{\bar{v}^*}^{1} f(v)(1-g(v))dv. \tag{40}$$

This always true because there is always at least one $v \in [\bar{v}^*, 1]$ such that $g(v) < 1$:

$$g(\bar{v}^*) = g\left(g^{-1}\left(1 - \frac{a}{p_i}\right)\right) = 1 - \frac{a}{p_i} < 1. \tag{41}$$

For the second and third part of the results, let us start by showing that p_i is not necessarily monotonically increasing in a. By the implicit function theorem, we find that:

$$\frac{dp_i}{da} = \frac{\partial p_i}{\partial a} + \frac{\partial p_i}{\partial \bar{v}^*}\frac{\partial \bar{v}^*}{\partial a} = \frac{-f(p_i)}{\frac{\partial^2 \Pi}{\partial p_i^2}} + \frac{f(\bar{v}^*)(1-g(\bar{v}^*))}{\frac{\partial^2 \Pi}{\partial p_i^2}}\frac{1}{pg'(\bar{v}^*)}. \tag{42}$$

On the one hand, $\frac{\partial p_i}{\partial a} > 0$ because both the numerator and the denominator are negative, but on the other hand, $\frac{\partial p_i}{\partial \bar{v}^*}\frac{\partial \bar{v}^*}{\partial a} < 0$ because $\frac{\partial^2 \Pi}{\partial p_i^2} < 0$. p_i is decreasing in a if:

$$f(p_i)p_i < f(\bar{v}^*)(1-g(\bar{v}^*))\frac{1}{g'(\bar{v}^*)}. \tag{43}$$

That $p_i < p_n$ if p_i is strictly monotonically increasing in a is derived from the fact that (i) p_n is constant in a and (ii) if a is sufficiently large (e.g. $a = 1$) then $p_i = p_n$. Hence, if p_i is strictly monotonically increasing in a, it must be that $p_i < p_n$.

The third part of the result is proven with the example from Figure 2b. The PDF of the Kumaraswamy $(2, 5)$ distribution is $f(v) = 10v(1-v^2)^4$. Assuming that the conditions of Proposition 1 are fulfilled, profit and the first-order conditions are:

$$\Pi = (p-a)\int_p^{\bar{v}} 10v(1-v^2)^4 dv + p\int_{\bar{v}}^1 10^2(1-v^2)^4 dv, \tag{44}$$

$$\frac{\partial \Pi}{\partial p} = \int_{\bar{v}}^1 10v^2(1-v^2)^4 dv + \int_p^{\bar{v}} 10v(1-v^2)^4 dv - (p-a)10p(1-p^2)^4 = 0, \tag{45}$$

$$\frac{\partial \Pi}{\partial \bar{v}} = p - a - p\bar{v} = 0. \tag{46}$$

Hence, we have $\bar{v}^* = 1 - a/p^*$. The expression of p^* is too complex to be written down. Targeted advertising arises so long as $0 < a < 0.2288$. Up to $a \simeq 0.1836$, p^* is increasing in a. If $a > 0.1836$, it is decreasing.

C Second-order conditions

Assuming that $p^* = p_i$ and hence, that $p^* < \bar{v}^* < 1$[21] , to satisfy the second-order conditions of the profit-maximization problem exposed in Equation 1, we require the Hessian matrix

[21] See footnote 12 for the two other cases, which are more standard.

to be negative definite, which implies that the following inequalities should be satisfied:

$$\frac{\partial^2 \Pi}{\partial p^2} = -2f(p^*) - (p^* - a)f'(p^*) \leq 0, \tag{47}$$

$$\frac{\partial^2 \Pi}{\partial (\bar{v}^*)^2} = -p^* f(\bar{v}^*) g'(\bar{v}^*) \leq 0, \tag{48}$$

$$\frac{\partial^2 \Pi}{\partial p^2} \frac{\partial^2 \Pi}{\partial (\bar{v}^*)^2} - \left(\frac{\partial^2 \Pi}{\partial p^* \partial \bar{v}^*} \right)^2 \geq 0. \tag{49}$$

Let us start with Inequality 47. From, the first-order condition (with regards to p, Equation 9) of the profit-maximization problem, we know that

$$p^* - a = \frac{\int_{\bar{v}^*}^1 f(v)g(v)dv + \int_{p^*}^{\bar{v}^*} f(v)dv}{f(p^*)} \tag{50}$$

and therefore, after some algebra, Inequality 47 can be rewritten as

$$f'(p^*) \geq \frac{-2f(p^*)^2}{\int_{\bar{v}^*}^1 f(v)g(v)dv + \int_{p^*}^{\bar{v}^*} f(v)dv}. \tag{51}$$

We have assumed that f is log-concave and hence, it has a monotonically increasing hazard rate, which implies that

$$f'(p^*)(1 - F(p^*)) > -f(p^*)^2 \iff f'(p^*) \geq \frac{-f(p^*)^2}{1 - F(p^*)}. \tag{52}$$

Because $1 - F(p^*) > \int_{\bar{v}^*}^1 f(v)g(v)dv + \int_{p^*}^{\bar{v}^*} f(v)dv$, we therefore have that

$$f'(p^*) > \frac{-f(p^*)^2}{1 - F(p^*)} > \frac{-2f(p^*)^2}{\int_{\bar{v}^*}^1 f(v)g(v)dv + \int_{p^*}^{\bar{v}^*} f(v)dv}. \tag{53}$$

The log-concavity of f implies that Inequality 47 is always respected. Inequality 48 is always respected because $g' \geq 0$.

The last inequality implies that we should have:

$$\left[-2f(p^*) - (p^* - a)f'(p^*)\right] \left[-p^* f(\bar{v}^*) g'(\bar{v}^*)\right] - \left[1 - g(\bar{v}^*)\right] \geq 0 \tag{54}$$

$$\iff \left[-2f(p^*) - (p^* - a)f'(p^*)\right] \left[-p^* f(\bar{v}^*) g'(\bar{v}^*)\right] \geq \left(\frac{a}{p}\right)^2 \tag{55}$$

This is an additional assumption that must be imposed on the information function and the distribution of valuation. We strongly suspect that it is implied by log-concavity but have not been able to prove it yet. It is verified for all the examples of the paper.

D A negative correlation between information and valuation

Suppose that instead of a positive correlation, there is a negative or no correlation between valuation and information: $g'(v) \leq 0$. The role of the targeting bound, which we now call $\underline{v}$, is inverted: consumers with $v < \underline{v}$ are not targeted because they are likely informed.[22] The profit-maximization problem and the first-order conditions[23] are the following:

$$\max_{p,\underline{v}} \Pi = \max_{p,\underline{v}} \left[p \int_p^{\underline{v}} f(v)g(v)dv + (p-a) \int_{\underline{v}}^1 f(v)dv \right], \tag{56}$$

$$\frac{\partial \Pi}{\partial p} = \int_p^{\underline{v}} f(v)g(v)dv - pf(p)g(p) + \int_{\underline{v}}^1 f(v)dv = 0, \tag{57}$$

$$\frac{\partial \Pi}{\partial \underline{v}} = pf(\underline{v})g(\underline{v}) - (p-a)f(\underline{v}) = 0 \Longleftrightarrow \underline{v}^* = g^{-1}\left(1 - \frac{a}{p}\right). \tag{58}$$

Information targeting is optimal if $p^* < \underline{v}^* < 1$. The definitions of the two corner solutions are different from the main case. Now, if $p^* = 1$ we have no advertising and if $p^* = \underline{v}^*$ targeting is on valuation.

We have the same three conditions as in the main case:

$$\left.\frac{\partial \Pi}{\partial \bar{v}}\right|_{\bar{v}=1} < 0, \tag{59}$$

$$\left.\frac{\partial \Pi}{\partial p}\right|_{p=\bar{v}^*} < 0, \tag{60}$$

$$\left.\frac{\partial \Pi}{\partial \bar{v}}\right|_{\bar{v}=p^*} > 0. \tag{61}$$

Because the interpretation of the corner solutions differs from the main case, the interpretation of these conditions also differs. The first states that information targeting is preferred to valuation targeting while the second and the third state that information targeting is preferred to no advertising.

Using exactly the same techniques as in the main case, we obtain a natural analogue to Proposition 1. If:

1. $g(1) < 1 - \dfrac{a}{p^*} < g(p^*)$.

[22]Setting another targeting bound is never optimal for reasons analogous to those of the positive correlation case. See Appendix A for more details.

[23]From log-concavity and using the same reasoning as for the main case, we know that $\frac{\partial^2 \Pi}{\partial p^2} \leq 0$ and $\frac{\partial^2 \Pi}{\partial \underline{v}^2} \leq 0$. An assumption similar to that of the main case must also be imposed.

2. $0 < a < a' \equiv p^* \left(1 - \frac{\int_{\underline{v}^*}^1 f(v)dv}{f(\underline{v})\underline{v}}\right)$.

Then, we have that $p^* < \underline{v}^* < 1$. Otherwise, the optimal price is a corner solution.

Suppose that $p^* < \underline{v}^* < 1$. We denote p^* by p_i for "information targeting". If we set $\underline{v} = p^*$, we denote p^* by p_v and finally, if we set $\underline{v} = 1$, we denote p^* by p_n. The first-order conditions in these respective cases are:

$$\begin{aligned}
\frac{\partial \Pi}{\partial p} &= \int_{p_i}^{\underline{v}} f(v)g(v)dv - p_i f(p_i)g(p_i) + \int_{\underline{v}}^1 f(v)dv = 0, && (62)\\
\left.\frac{\partial \Pi}{\partial p}\right|_{\underline{v}=p^*} &= \int_{p_v}^1 f(v)dv - (p_v - a)f(p_v) = 0, && (63)\\
\left.\frac{\partial \Pi}{\partial p}\right|_{\underline{v}=1} &= \int_{p_n}^1 f(v)g(v)dv - p_n f(p_n)g(p_n) = 0. && (64)
\end{aligned}$$

A straightforward[24] comparison of the first-order conditions shows that $p_n < p_i < p_v$ if $0 < a < a'$. This differs slightly from Proposition 2 but we retain the main conclusion that an increased ability to target does not necessarily increases price.

The non-monotonicity result on price does not hold any more because there is a near-dichotomy between the pricing and the advertising decisions: a only enters the first-order condition with regard to price through $\bar{v}$. Interestingly however, an increase in a always leads to a *decrease* in price. Formally,

$$\begin{aligned}
\frac{dp_i}{da} &= \frac{\partial p_i}{\partial a} + \frac{\partial p_i}{\partial \bar{v}^*}\frac{\partial \bar{v}^*}{\partial a}, && (65)\\
&= \frac{\partial p_i}{\partial \bar{v}^*}\frac{\partial \bar{v}^*}{\partial a} = -\frac{f(v)(g(v)-1)}{\frac{\partial^2 \Pi}{\partial p^2}}\left(\frac{-1}{p}\right) g'(\underline{v}) < 0 && (66)
\end{aligned}$$

Information targeting also affects consumer surplus, which is the following:

$$\begin{aligned}
CS_i &= \int_{p_i}^{\underline{v}^*} f(v)g(v)(v - p_i)dv + \int_{\underline{v}^*}^1 f(v)(v - p_i)dv, && (67)\\
CS_v &= \int_{p_v}^1 f(v)(v - p_v)dv, && (68)\\
CS_n &= \int_{p_n}^1 f(v)g(v)(v - p_n)dv. && (69)
\end{aligned}$$

Compared to no advertising, the effect of information targeting on consumer surplus is ambiguous. On the one hand, we know that $p_n < p_i$ but on the other hand, more consumers

[24]See footnote 13 for details.

are informed under information targeting. Compared to valuation targeting, in theory we also have that information targeting has ambiguous effects. On the positive side $p_i < p_v$ but on the negative side, it may be that less consumers buy because of $\underline{v}^* > p_i$. In practice, this is very unlikely. Indeed, even if this effect is negative, it should be small because it only concerns consumers with a low valuation ($v \in [p_i, \bar{v}^*]$). This is quite different from the main case where consumer surplus was much more likely affected because the consumers who did not receive ads were those with $v \in [\bar{v}^*, 1]$. The same reasoning applies to welfare.

Interestingly, we still obtain the non-monotonicity result regarding consumer surplus. Due to the interaction between the pricing and the advertising strategies – a higher a leads to a lower p_i but also to a higher number of uninformed consumers (a higher $\underline{v}^*$)– consumer surplus can increase or decrease in a. The same is true of welfare.

E The Kumaraswamy (2,5) distribution

The Kumaraswamy (a,b) distribution has the following probability and cumulative distribution functions (PDF and CDF):

$$f(x) = abx^{a-1}(1-x^a)^{b-1} \quad \text{if} \quad x \in [0,1] \quad \text{and} \quad 0 \quad \text{otherwise} \tag{70}$$

$$F(x) = 1-(1-x^a)^b \quad \text{if} \quad x \in [0,1] \quad \text{and} \quad 0 \quad \text{otherwise} \tag{71}$$

The PDF and CDF produced in Figure 6 can be produced by setting $a = 2$ and $b = 5$. The main advantage of this distribution is that it has a simple closed form which allows us to have relatively complicated information functions and still be able to solve the model.

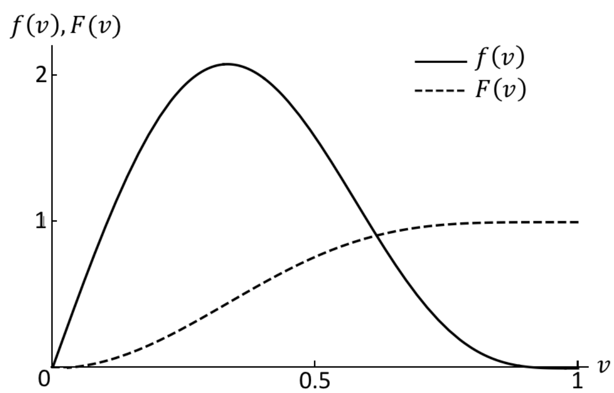


Figure 6: CDF and PDF of the Kumaraswamy (2,5) distribution.