Targeting Advertising Preferences

David Bounie[*], Antoine Dubus[†] and Martin Quinn[‡]

January 11, 2018

Abstract

This paper investigates how profiling technologies can help platforms to target consumers' advertising preferences. Profiling technologies are used by platforms to infer the ad-sensitivity of consumers. Platforms can tailor the level of ads to the ad-sensitivity of consumers. We show that such technologies have large implications on advertising intensity and welfare.

Key Words: Online advertising; Ad nuisance; Ad avoidance.
JEL Classification: L82; L86; M3.

1 Introduction

Advertising is the main mechanism to finance media content such as TV, radio or magazines. It also constitutes most of media platforms income on the Internet (Shiller et al., 2017). For instance, in the third quarter of 2016, U.S. advertisers invested \$17.6 billion in digital advertising according to the Internet Advertising Bureau,[1] a 20 percent increase over the same time period in 2015.

However, Internet users tolerate less and less online ads that degrade their online experience (Manchanda et al., 2006; Goldfarb and Tucker, 2011b), and as a response, they block ads using ad-avoidance technologies (AATs).[2] In 2016, 26.3% of US online users were using an ad-blocker (69.8 million Americans), a jump of 34.4% over last year.[3] AATs can be installed on web browser, and forthcoming versions of Chrome and Firefox will even directly integrate them by default.[4] The ad-blocking feature will filter out automatically ads such as pop-ups and auto-playing video.

The growing adoption of AATs by a significant part of online users confirms that advertising is perceived as a nuisance. However, an even larger proportion of consumers still continue to

[*]Telecom ParisTech (Economics and Social Sciences, 46 rue Barrault, 75634 Paris Cedex 13; Email: david.bounie@telecom-paristech.fr) and Center for the Study of Advertising Media; Corresponding Author.

[†]PhD student, Telecom ParisTech (Economics and Social Sciences, 46 rue Barrault, 75634 Paris Cedex 13; Email: antoine.dubus@telecom-paristech.fr and Chair Values and Policies of Personal Information

[‡]PhD student, Telecom ParisTech, Economics and Social Sciences, 46 rue Barrault, 75634 Paris Cedex 13; Email: martin.quinn@telecom-paristech.fr and Chair Values and Policies of Personal Information.

[1]IAB, Highest Third Quarter Spending on Record, December 28, 2016.

[2]Adblock Plus or uBlock are examples of popular AATs on the Internet.

[3]eMarketer, US Ad Blocking to Jump by Double Digits This Year, June 21, 2016.

[4]ArsTechnica, Report: Google will add an ad blocker to all versions of Chrome Web browser, April 19, 2017.

visit websites and click on ads that are related or targeted to their personal tastes and interests. Ad-sensitivity, or equivalently, ad preferences therefore vary across online users.

Online content providers may adopt various strategies to account for ad-sensitivity. For example, Tag (2009) analyzes a monopoly platform that offers either an ad-free pure subscription option or an ad-only program to discriminate consumers. The author shows that the subscription option induces a higher level of advertising for those remaining on the ad-only option. In the end, the aggregate consumer surplus falls, whereas the advertiser surplus rises through lower ad prices. Anderson and Gans (2011) extend the possibility for consumers to choose a costly ad-avoidance technology (such as TiVo) to strip out the ad nuisance in place of a subscription option. They confirm the results of Tag (2009), and also show that the adoption of AATs may reduce program quality.

The platform strategies analyzed in Tag (2009) and Anderson and Gans (2011) rely on the idea that ads are only perceived as a nuisance, and that consumers' advertising preferences are not directly observable (they are exogenous). By choosing a program, online users self-select and reveal afterward their ad-sensitivity. However, recent profiling technologies are allowing platforms to infer the ad-sensitivity of online users. On the basis of past behavior (click on ads, number of ads seen, etc.), or by associating consumers that have similar profiles, platforms can determine whether a user is more or less ad-sensitive, and tailor accordingly the level of advertising to prevent him adopting AATs. For example, digital platforms are using profiling technologies to limit the number of ads seen by consumer, the frequency of exposure of a consumer to a given ad, and ads to products a consumer has already purchased. As illustations, Facebook and Snapchat limit the number of ads users can see in newsfeeds[5] or to be appeared between friends' snaps and stories.[6] This technique is called "frequency capping" (Buchbinder et al., 2011).

This article analyzes the impact of the use of a profiling technology on platforms' strategies, the surplus of consumers and advertisers, and the volume of ads served on the market. We develop a model in line with Anderson and Coate (2005), Tag (2009) and Anderson and Gans (2011), where a monopoly platform infers consumers' ad preferences using a profiling technology. The model is based on three key features. Firstly, advertising is not necessarily perceived as a nuisance, as it may benefit Internet users to see some ads. Secondly, online users are heterogeneous in their ad-sensitivity. Two types of consumers are distinguished: there is a proportion of consumers who are *strongly ad-sensitive* ($\overline{\gamma}$) and a proportion of consumers who are *weakly ad-sensitive* ($\underline{\gamma}$). Thirdly, we allow the platform to classify users according to their

[5] TechCrunch, How Facebook News Feed Works, Sep 6, 2016.
[6] Adweek, Snapchat Is Letting More Brands Run Ads Between Friends' Stories, August 10, 2016.

ad-sensitivity thanks to a profiling technology. However, the profiling technology is not always efficient, i.e. it does not perfectly identify the users' ad-sensitivity. In other words, consumers $\overline{\gamma}$ and $\underline{\gamma}$ may respectively see a level of advertising unadapted to their advertising preferences, hence adopting AATs.

Our model shows that the use of a profiling technology by a platform have strategic implications. Firstly, we highlight that the profiling technology may not always be used by the platform. Indeed, the platform uses the profiling technology only if it generates greater revenue. This is the case when the technology is *efficient* such that the probability of correctly classifying the ad-sensitivity of users is sufficiently high. In this case, the technology allows the platform to tailor the levels of advertising to the ad-sensitivity of users, thus increasing its profits. Indeed, without profiling technology - or if the technology is not efficient enough to be used -, the platform cannot precisely infer the advertising preferences of each online user. It will therefore set a unique level of advertising for all users on the site, whether they are ad-sensitive or not. But this strategy is unlikely to be effective. If the platform sets a low level of advertising while users are not ad-sensitive, it could inflate the level of advertising to increase its profits. Similarly, if the platform sets a high level of advertising while users are very ad-sensitive, the latter will prefer not to see ads and will then adopt AATs. This result is not trivial because an efficient profiling technology is not always flawless. Indeed, a perfect profiling technology always perfectly classifies Internet users according to their type. However, an efficient but imperfect technology makes mistakes, even if it generate higher profits for the platform. In this case, technology may classify a consumer as being very ad-sensitive when he is actually not ad-sensitive, and *vice versa*. Hence, even if the use of the technology increases the platform's profits, it may not properly tailor the level of advertising to the ad-sensitivity of each type of user. This can have big implications on the platform profits. For example, a user who is wrongly classified as being very ad-sensitive will see a lower number of ads, thus reducing the profit opportunities for the platform. On the contrary, a user who is wrongly classified as being not ad-sensitive will see too many ads, and will adopt in reaction AATs.

Secondly, the use of an efficient profiling technology by the platform also changes the total number of ads served on the market. We find that in presence of a perfect profiling technology, the total number of ads served to Online users is always higher than without technology. This is due to the fact that the technology allows to serve more ads to weakly ad-sensitive users, while serving an appropriate level of ads to users who were avoiding ads before. However, when the technology is not flawless, the results change. We find that if the plaform faces a strongly ad-sensitive audience, more ads will be served with technology than without. Conversely, if the platform adresses a weakly ad-sensitive audience, the total number of ads served is higher with technology only if this technology is highly efficient.

Thirdly, we show that introducing a perfect profiling technology is always welfare increasing when the platform is facing and audience that is not very ad-sensitive, as the gains for advertisers and the platform offset the potential loss in Internet users surplus. However, the analysis become more complex when 1) the technology is perfect but the platform is facing a very ad-sensitive audience or 2) the technology is imperfect. In both case, the impact of the technology on welfare depends on Internet user surplus and advertisers profits. Firstly, the impact of a profiling technology on Internet users surplus depends on whether the less elastic fringe of users prefers to see many ads (being correctly classified) or few ads (being misclassified). Secondly, we show that advertisers may not benefit from the introduction of a profiling technology, while the platform does. This situation happens when the technology is sufficiently efficient to be implemented by the platform but not efficient enough to increase advertisers profits.

The paper is organized as follows. The next section describes the model. Section 3 solves the case where the platform does not use a profiling technology. Section 4 introduces the profiling technology. Section 5 analyzes the aggregate number of ads served on the market. Section 7 summarizes the key findings and strategic implications of the paper, and concludes.

2 Description of the Model

Our model deals with a monopolist platform (also called platform) that delivers content and displays ads on a website to online users (also named consumsers), and sells advertising space to advertisers. The platform therefore manages its website to attract online users on one side and advertisers on the other.

2.1 Online Users

Online users visit the website and receive utility:

$$U = \begin{cases} 1 + \gamma(a), \text{ if choosing to visit and see ads,} \\ \\ \theta, \text{ if choosing to visit without seeing ads,} \end{cases} \tag{1}$$

and 0 otherwise. A consumer receives therefore utility by viewing an editorial content of quality q, normalized here to 1. If she chooses to visit the website and see ads, she receives $\gamma(a)$, where a is the number of ads (or advertisers), and γ the sensitivity to advertising. $\gamma(a)$ is assumed to be continuous, differentiable, concave in a, with $\gamma(a = 0) = 0$. $\gamma(a)$ has two notable properties. First, when no ads is displayed on the website, the consumer only enjoys the quality of editorial content. Second, we hypothesize that a low number of ads is enjoyable and provides utility for consumers; targeted advertising for instance provides useful information about product firms. However, too many ads displayed on the website, even though they are targeted, degrade the user experience and provide disutility. Advertising is perceived in this case as a nuisance. This is crucial in generating demand for AATs. Indeed, when advertising is perceived as a nuisance, the consumer can choose to view content and earn utility θ for using AATs.

Consumers are heterogeneous in their sensitivity to advertising. To keep things as simple as possible, we consider two populations. First, there is a proportion β of consumers who are *strongly ad-sensitive*: $\overline{\gamma}$. By contrast, there is a proportion $1-\beta$ of consumers who are *weakly ad-sensitive*: $\underline{\gamma}$. By definition, consumers of type $\overline{\gamma}$ would prefer to see a lower number of ads compared to weakly ad-sensitive consumers $\underline{\gamma}$.[7]

To summarize, the respective utilities of consumers $\overline{\gamma}$ and $\underline{\gamma}$ are:

$$\overline{U} = \begin{cases} 1+\overline{\gamma}(a), & \text{if choosing to visit and see ads,} \\ \theta, & \text{if choosing to visit and not see ads,} \end{cases} \tag{2}$$

and,

$$\underline{U} = \begin{cases} 1+\underline{\gamma}(a), & \text{if choosing to visit and see ads,} \\ \theta, & \text{if choosing to visit and not see ads.} \end{cases} \tag{3}$$

From Eqs. 2 and 3, it is straightforward to see that when the cost to block ads is high, i.e. when $\theta < 0$, the choice of the consumer is, whatever his ad-sensitivity, either to consume with ads or not to visit the website. In the sequel, we will only consider the cases where the cost of using AATs is low enough, i.e. when $\theta > 0$. In other words, Internet users have only the choice between consuming and seeing ads or consuming without ads and using AATs.

This assumption is clearly in line with recent announces from Google and other large Internet companies. Google for example is about to roll out an ad-blocking version of his Chrome web browser (Wall Street Journal, Google Plans Ad-Blocking Feature in Popular Chrome Browser, April 19, 2017).[8] The ad-blocking feature, which could be switched on by default, would filter out automatically unacceptable ads such as pop-ups and auto-playing video, that degrade online user experience. As the feature of ad-blocking would be installed by default, the cost of blocking ads would be dramatically low.

[7] From the properties on $\gamma(a)$, we assume that $\underline{\gamma}(a) \geq \overline{\gamma}(a)$, $\forall a \geq 0$.

[8] In the U.S., Chrome has nearly 47.5% of the browser market across all platforms, according to online analytics provider StatCounter.

2.2 Advertisers

Following Anderson and Coate (2005) and Anderson and Gans (2011), we assume that a single ad suffices to reach all consumers on the platform, and so an advertiser decides to place an ad as long as the profit per consumer is no smaller than the price paid for an ad per consumer reached. We also rank advertisers from highest to lowest willingness to pay per consumer and derive the advertiser inverse demand curve $r(a)$. The corresponding total advertising revenue earned per individual, $R(a) = r(a)a$, where $r(a)$ is concave, and $r'(a) < 0$ when $r(a) > 0$, making $R(a)$ concave in a.[9]

2.3 Platform

The platform delivers content to Internet users and displays ads on his website. The platform is only financed by advertising (and not by subscription). His profit function corresponds to the revenue per Internet user $R(a)$ times the number N of Internet users:

$$\Pi = r(a)aN(a) = R(a)N(a).^{10} \tag{5}$$

As $R(a)$ is concave in a (because $r(a)$ is), the platform is interesting to get on board enough advertisers to be profitable, but up to a certain level, too many advertisers lower its revenues (at a certain point, $r(a)$ is too low as the further advertiser expects a lower impact of its ad).

3 Baseline Model Without Profiling Technology

In the baseline case, the platform has no profiling technology to find out the consumer ad-sensitivity. We introduce the profiling technology in Section 4.

The timing of the game in this baseline case is in two stages. In stage 1, the platform chooses the number of advertisers a (or equivalently, the number of ads to be displayed), and in stage 2, consumers choose to visit the website. In other words, when visiting the website, consumers discover the level of advertising and its possible nuisance, and choose to see or not ads using AATs. We solve the game by backward induction.

[9] These properties are similar to the ones in Anderson and Gans (2011).

[10] By assumption $1 + \underline{\gamma}(a) > 1 + \overline{\gamma}(a)\ \forall a$, and $N(a)$ can be written as:

$$N(a) = \begin{cases} 0, \text{ if } \theta > 1 + \underline{\gamma}(a), \\ 1 - \beta, \text{ if } 1 + \underline{\gamma}(a) > \theta > 1 + \overline{\gamma}(a), \\ 1, \text{ if } 1 + \overline{\gamma}(a) > \theta. \end{cases} \tag{5}$$

3.1 Stage 2

The users decide to view ads or to avoid them using AATs. Consumers of type $\overline{\gamma}$ and $\underline{\gamma}$ choose to visit the website depending of their respective sensitivity to advertising. Their respective utilities are defined in Eqs. (2) and (3). Let define $\overline{m}$ and $\underline{m}$ as the solutions for which $1 + \overline{\gamma}(\overline{m}) = \theta$ and $1 + \underline{\gamma}(\underline{m}) = \theta$. These respective solutions can be interpreted as the maximum levels of advertising that users $\overline{\gamma}$ and $\underline{\gamma}$ are willing to accept to visit a website.

If the cost of AATs decreases for example, consumers $\overline{\gamma}$ and $\underline{\gamma}$ are both less ready to see ads. However, consumers $\overline{\gamma}$ and $\underline{\gamma}$ respond differently to a change in AATs. For any given θ, consumers $\overline{\gamma}$ who are strongly sensitive to advertising, will accept to see less ads than consumers $\underline{\gamma}$, who are weakly sensitive to advertising. As consumers $\overline{\gamma}$ are more sensitive to the nuisance cost of ads than consumers $\underline{\gamma}$, the former will be more likely to use AATs when its utility from avoidance ads θ increases.

3.2 Stage 1

As the platform cannot discriminate consumers in the baseline case with a profiling technology, it can choose either to sell advertising spaces to a low number of advertisers, serving possibly consumers $\overline{\gamma}$ and $\underline{\gamma}$, or to sell spaces to a high number of advertisers and reach only consumers $\underline{\gamma}$ who have a high ad-sensitivity. The general optimization problem of the platform is:

$$\max_a \Pi = r(a)aN = R(a)N(a). \tag{6}$$

The strict concavity of the profit function guarantees the existence and uniqueness of an optimal level of advertising noted $\hat{a}$.[11] The latter may be lower than the maximum levels of advertising $\overline{m}$ and $\underline{m}$ that consumers $\overline{\gamma}$ and $\underline{\gamma}$ are willing to accept to visit the website. In this case, the ad-avoidance technology does not exert any competitive pressure on the platform, and there is an interior equilibrium where ads are served to both types of users. This is examined in Proposition 1.

Proposition 1: When the platform is not constrained ($\hat{a} < \overline{m}$) (which corresponds to $1 + \overline{\gamma}(\hat{a}) \geq \theta$), an interior equilibrium exists where ads are served to both types of users. The potential for ad avoidance does not impact platform behaviour, and hence the level of advertising.

Proposition 1 is original with respect to Tag (2009) and Anderson and Gans (2011). When Internet users are weakly ad-sensitive, ads are perceived as a benefit and not as a nuisance by both types of users; consumers derive a positive utility from seeing ads, and can be defined as

[11] The first order condition with respect to a is: $r(\hat{a}) = -\hat{a}\frac{\partial r(\hat{a})}{\partial a}$.

ad-lovers. In this case, the platform can choose to display few ads to both types of users to maximize its profits. The level of advertising does not therefore depend on AATs. This case is displayed in Figure 1: when setting $a^* = \hat{a}$, the platform earns $\Pi^* = R(a^* = \hat{a}) = \hat{R}$, and both types of consumers $\overline{\gamma}$ and $\underline{\gamma}$ visit the website ($N = 1$) as the optimal level of advertising is lower that the ones which would exclude both types of consumers from visiting the website, i.e. $\hat{a} < \overline{m} < \underline{m}$.

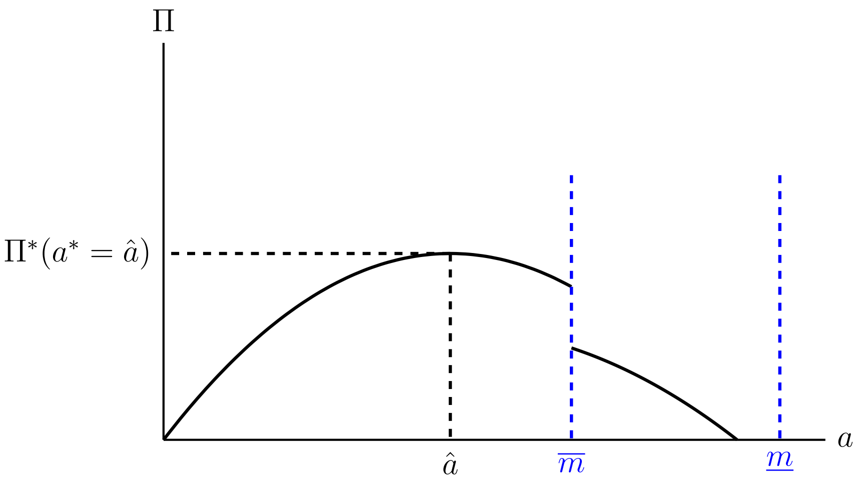

Figure 1: $1 + \overline{\gamma}(\hat{a}) \geq \theta$

However, Internet users may be more ad-sensitive. Indeed, $\hat{a}$ may be greater than the maximum levels of advertising $\overline{m}$ and/or $\underline{m}$ that consumers $\overline{\gamma}$ and $\underline{\gamma}$ are willing to accept to visit the website, forcing the platform to lower its advertising level to prevent Internet users from adopting AATs. For example, when $\min(\hat{a}, \overline{m}) = \overline{m}$ (which corresponds to $1 + \overline{\gamma}(\hat{a}) < \theta$), if the platform sets $a^* = \hat{a}$, it cannot reach consumers who are strongly ad-sensitive. Two main cases are thus possible.

In the first case, when $\hat{a} < \underline{m}$ (which corresponds to $\theta \leq 1 + \underline{\gamma}(\hat{a})$), the platform can only reach weakly ad-sensitive consumers when setting $a^* = \hat{a}$. As a consequence, the platform has the choice between setting $a^* = \overline{m}$ and serving all the users ($N = 1$), or setting $a^* = \hat{a}$ and reaching only weakly ad-sensitive consumers ($N = 1 - \beta$). We refer to this case in the sequel as a case where the platform is *weakly constrained* on its level of advertising. In this case, weakly ad-sensitive users remain *ad-lovers* as they always get positive utility from advertising, regardless the choice of the platform. Conversely, strongly ad-sensitive users are *ad-neutral* when $a^* = \overline{m}$, as they receive similar benefit from advertising than with AATs, and *ad-averse* when $a^* = \hat{a}$, as they all prefer AATs and not see ads. Figure 2 (left) illustrates the case where both types of consumers visit the website and the profits of the platform are higher for

$a^* = \overline{m}$. However, setting $a^* = \hat{a}$ as it is displayed in Figure 2 (right) would be preferable for the platform with the strongly ad-sensitive users on board.

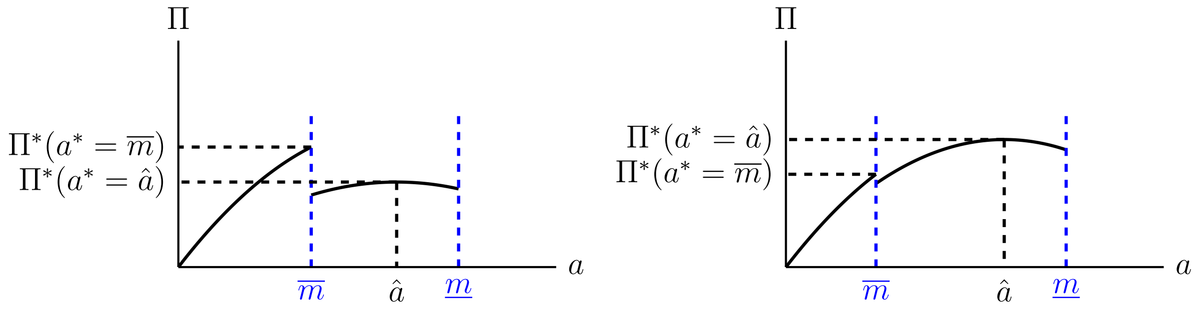

Figure 2: $1 + \overline{\gamma}(\hat{a}) < \theta \leq 1 + \underline{\gamma}(\hat{a})$

In the second case, when $\hat{a} > \underline{m}$ (which corresponds to $1 + \underline{\gamma}(\hat{a}) < \theta$), both types of consumers are not interested in visiting the website when setting $a^* = \hat{a}$. The platform has the choice between setting $a^* = \overline{m}$, and reach both types of consumers ($N = 1$), or setting $a^* = \underline{m}$ and serve again only weakly ad-sensitive consumers ($N = 1 - \beta$). By reference to the previous strategy, we will say that the platform is *strongly constrained* on its level of advertising. In this case, weakly ad-sensitive users are now ad-neutral as they are indifferent between seeing ads or using AATs. However, strongly ad-sensitive users are ad-averse as they all prefer AATs and not see ads. Moreover, if the strategy of the platform is $a^* = \overline{m}$, it follows the same logic as in the previous case: weakly ad-sensitive users are ad-lovers while strongly ad-sensitive users are ad-neutral. The choice faced by the platform is illustrated in Figure 3. In Figure 3 (left), setting a level of advertising $a^* = \underline{m}$ to target weakly ad-sensitive users is less profitable for the platform than setting $a^* = \overline{m}$, a lower number of advertisers allowing to serve both types of consumers.

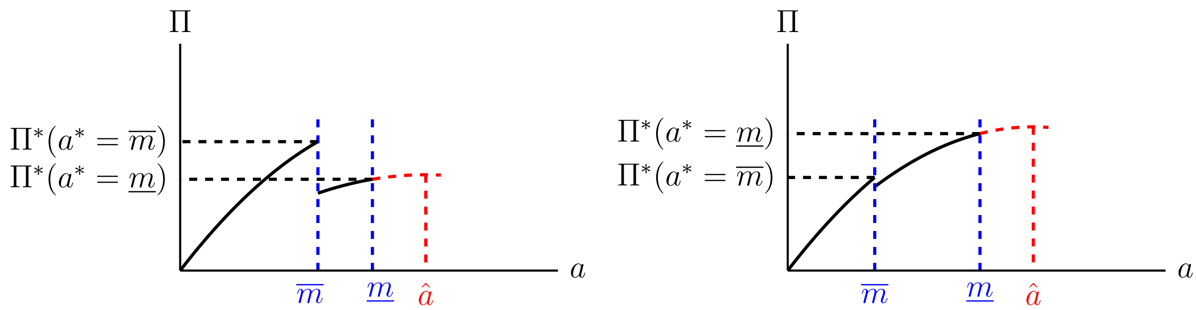

Figure 3: $1 + \underline{\gamma}(\hat{a}) < \theta$

To summarize, when the platform is weakly or strongly constrained on the level of advertising, it can choose to serve either both types of users or only weakly ad-sensitive users. The profits of the platform are in these cases:

$$\Pi^* = \begin{cases} \overline{R}, \text{ if both types of consumers visit the website,} \\ \\ (1-\beta)\underline{R}, \text{ if the only } \textit{weakly ad-sensitive} \text{ consumers visit the website,} \end{cases} \tag{7}$$

where $\overline{R} = R(a^* = \overline{m})$. $\underline{R}$ represents either the revenues from setting the optimal level of ads $\hat{a}$ (implying $\underline{R} = R(a^* = \hat{a})$) when the weakly ad-sensitive users have a higher utility when viewing ads rather than adopting AATs ($\hat{a} < \underline{m}$), or the revenues from setting the optimal level of ads $\underline{m}$ (implying $\underline{R} = R(a^* = \underline{m})$), when the weakly ad-sensitive users have a higher utility when adopting AATs rather than viewing advertising ($\hat{a} > \underline{m}$). The revenue per user when the platform chooses to focus only on weakly ad-sensitive consumers can therefore be written as $\underline{R} = R(a^* = \min(\hat{a}, \underline{m}))$. To keep the notations as simple as possible, we will use $\overline{R}$ and $\underline{R}$ in the rest of the paper.

Proposition 2 summarizes the different cases regarding the choices of advertising levels.

Proposition 2: When the platform is constrained ($\hat{a} > \overline{m}$), two cases are possible. First, when $\frac{\overline{R}}{\underline{R}} > 1-\beta$, it chooses the level of advertising $a^* = \overline{m}$, and displays ads to users $\overline{\gamma}$ and $\underline{\gamma}$. Second, when $\frac{\overline{R}}{\underline{R}} < 1-\beta$, it chooses $a^* = \min(\hat{a}, \underline{m})$ and only displays ads to consumers $\underline{\gamma}$. See Proof of Proposition 2 in Appendix A.1.

Proposition 2 states that when the proportion of strongly ad-sensitive users is high in the population ($\frac{\overline{R}}{\underline{R}} > 1-\beta$), the platform may find profitable to set the level of advertising so as to attract not only weakly ad-sensitive users but also strongly ad-sensitive ones. By contrast, when the proportion of strongly ad-sensitive users is low ($\frac{\overline{R}}{\underline{R}} < 1-\beta$), it is most profitable to only target weakly ad-sensitive users.

4 Equilibrium with Profiling Technology

In the baseline case, the platform does not have a profiling technology to find out the consumer preferences for advertising, and sets an equal level of advertising for both types of consumers. In this section, the platform can now use a profiling technology to tailor a level of advertising to each type of consumer $\overline{\gamma}$ and $\underline{\gamma}$, and prevent consumers from adopting AATs.

The profiling technology exploits programmatic techniques, which are methods for buying and selling online ads in real time. Real time means that the entire process takes only a few milliseconds to complete, before a web page is loaded by a consumer. Typically, a consumer visits a website. Once connected, the technology gathers personal information about users with the help of cookies, and produces a signal that can be of two types, $\underline{s}$ or $\overline{s}$.[12] A signal $\overline{s}$ (resp. $\underline{s}$) means that the technology correctly classifies a user of type $\overline{\gamma}$ (resp. $\underline{\gamma}$) with probability δ. Without loss of generality, we assume that $\delta \in [\frac{1}{2}, 1]$; $\delta = 1$ means that the technology is perfectly efficient and always classifies correctly Internet users, and $\delta = \frac{1}{2}$ indicates that the technology brings no information, and correctly classifies Internet users with probability $\frac{1}{2}$.

Once the classification has been made, the platform is able to display the content and the level of advertising adapted to the probable type of consumer: for example, the level of advertising $\overline{a}^T$ will be displayed to consumers classified as strongly ad-sensitive ($\overline{\gamma}$), and the level of advertising $\underline{a}^T$ to the ones classified as weakly ad-sensitive consumers ($\underline{\gamma}$). If the profiling technology is efficient, i.e. if $\delta = 1$, the platform perfectly discriminates the types of consumers, and tailors the level of advertising $\overline{a}^T$ and $\underline{a}^T$ to users $\overline{\gamma}$ and $\underline{\gamma}$, who do not use in turn AATs. On the other hand, if the technology does not perfectly identify the types of users $\delta \in [\frac{1}{2}, 1[$, consumers $\overline{\gamma}$ and $\underline{\gamma}$ may see a level of advertising unadapted to their advertising preferences, and adopt in reaction AATs.

The use of a profiling technology does not change the timing structure for the order of moves except that the platform chooses now the number of advertisers a in a context where the profiling technology is effective in stage 2. In stage 1, the platform sets the advertising levels $\overline{a}^T$ and $\underline{a}^T$ to be displayed to consumers classified as $\overline{\gamma}$ and $\underline{\gamma}$. In stage 2, consumers choose to visit the website, and decide to avoid ads if the level of advertising is not adapted. We solve the game by backward induction.

4.1 Stage 2

Following the baseline case, the users decide to view ads or to avoid them using AATs. Depending on their ad-sensitivity, the utilities can be written as:

$$\overline{U} = \begin{cases} 1 + \overline{\gamma}(\overline{a}^T), \text{ if choosing to visit with ad level } \overline{a}^T, \\ \\ 1 + \overline{\gamma}(\underline{a}^T), \text{ if choosing to visit with ad level } \underline{a}^T, \\ \\ \theta, \text{ if choosing to visit and not see ads,} \end{cases} \tag{8}$$

and,

[12] The profiling technology uses machine-learning methods to construct signals.

$$\underline{U} = \begin{cases} 1 + \underline{\gamma}(\overline{a}^T), \text{ if choosing to visit whith ad level } \overline{a}^T, \\ 1 + \underline{\gamma}(\underline{a}^T), \text{ if choosing to visit whith ad level } \underline{a}^T, \\ \theta, \text{ if choosing to visit and not see ads.} \end{cases} \tag{9}$$

4.2 Stage 1

4.2.1 Description of the profiling technology

Following the Bayes' rule, we can calculate the expected profits of the platform. The probability to receive the signal $\overline{s}$ knowing that the Internet user is of type $\overline{\gamma}$ is equal to δ, and can be written as $\mathbb{P}(\overline{s}|\overline{\gamma}) = \delta$. The signal $\overline{s}$ is received with probability $\mathbb{P}(\overline{s}) = \beta\delta + (1-\beta)(1-\delta)$ and $\underline{s}$ with probability $\mathbb{P}(\underline{s}) = \delta(1-\beta) + \beta(1-\delta)$. Upon receiving the signal s, the platform then knows that this signal is true with probability:

$$\mathbb{P}(\overline{\gamma}|\overline{s}) = \frac{\mathbb{P}(\overline{s}|\overline{\gamma})\mathbb{P}(\overline{\gamma})}{\mathbb{P}(\overline{s})} = \frac{\delta\beta}{\delta\beta + (1-\delta)(1-\beta)},$$

and

$$\mathbb{P}(\underline{\gamma}|\underline{s}) = \frac{\mathbb{P}(\underline{s}|\underline{\gamma})\mathbb{P}(\underline{\gamma})}{\mathbb{P}(\underline{s})} = \frac{\delta(1-\beta)}{\delta(1-\beta) + (1-\delta)\beta}.$$ [13]

Applying the Bayes rule, the expected profits of the platform are:

$$\begin{aligned} \max_{\overline{a}^T, \underline{a}^T} \mathbb{E}\left(\Pi^T\right) &= \mathbb{P}(\overline{s})\left(\mathbb{P}(\underline{\gamma}|\overline{s}) + \mathbb{P}(\overline{\gamma}|\overline{s})\right) R(\overline{a}^T) + \mathbb{P}(\underline{s})\left(\mathbb{P}(\underline{\gamma}|\underline{s}) + \mathbb{P}(\overline{\gamma}|\underline{s})\right) R(\underline{a}^T) \\ &= [\beta\delta + (1-\beta)(1-\delta)]R(\overline{a}^T) + [\beta(1-\delta) + \delta(1-\beta)]R(\underline{a}^T) \end{aligned} \tag{10}$$

4.2.2 Advertising choice of the platform with profiling technology

Similarly to the baseline case, when the platform is not constrained, that is when ads provide utility for both types of users ($\hat{a} < \overline{m}$), the optimal amount of ads targeted for consumers $\overline{\gamma}$ and $\underline{\gamma}$ is exactly the same. The profiling technology is not therefore useful in this case. This result is summarized in Proposition 3.

Proposition 3: When the platform is not contrained- that is when $\hat{a} < \overline{m}$- the platform always sets its interior level of advertising which is equal for all users, i.e. $\overline{a}^{T*} = \underline{a}^{T*} = \hat{a}$. The profiling technology is not useful in this case.

[13] By definition, $\mathbb{P}(\underline{\gamma}|\overline{s}) \equiv 1 - \mathbb{P}(\overline{\gamma}|\overline{s})$, and $\mathbb{P}(\overline{\gamma}|\underline{s}) \equiv 1 - \mathbb{P}(\underline{\gamma}|\underline{s})$.

Proposition 3 is equivalent to Proposition 1, and the same intuitions apply. A more interesting case arises when the platform is constrained, as the profiling technology can now be helpful in discriminating consumers in setting two distinct levels of advertising.

Focusing on the case where the publisher is contrained $\hat{a} > \overline{m}$, we notice that expected profits in Eq. (10) are made of four terms: two terms related to the probability of receiving the signal $\overline{s}$, and two terms related to the probability of receiving the signal $\underline{s}$. First, upon receiving the signal $\overline{s}$, the platform will set the level of advertising $\overline{a}^{T*} = \overline{m}$ to attract users $\overline{\gamma}$. The resulting profits are also composed of two terms. The first term, $\beta\delta R(\overline{a}^{T^*})$, is the revenue related to the successful classification of users $\overline{\gamma}$, and the second term, $(1-\beta)(1-\delta)R(\overline{a}^{T*})$, is the revenue from the wrong classification of users $\underline{\gamma}$. In both cases, users $\overline{\gamma}$ and $\underline{\gamma}$ will choose to see ads without AATs.

Second, upon receiving the signal $\underline{s}$, the platform will set the level of advertising $\underline{a}^{T*} = \min(\hat{a}, \underline{m})$ to attract users $\underline{\gamma}$. The resulting profits are also composed of the proportion of successfully classified users $\underline{\gamma}$, $\delta(1-\beta)R(\underline{a}^{T*})$, and the wrong classification of users $\overline{\gamma}$, $\beta(1-\delta)R(\underline{a}^{T*})$. In this case, only users $\underline{\gamma}$ will choose to see ads without AATs. Indeed, users $\overline{\gamma}$ receive a higher utility when using AATs than from seeing ads.

We summarize the strategies of the platform in the Proposition 4:

Proposition 4: When the platform is constrained ($\hat{a} > \overline{m}$), the use of the profiling technology allows the platform to discriminate consumers $\overline{\gamma}$ and $\underline{\gamma}$: a proportion $\delta\beta + (1-\delta)(1-\beta)$ of users classified as $\overline{\gamma}$ sees a level of advertising $\overline{a}^{T*} = \overline{m}$, while a proportion $\delta(1-\beta)+(1-\delta)\beta$ classified as $\underline{\gamma}$ sees a level of advertising $\underline{a}^{T*} = \min(\hat{a}, \underline{m})$.

In setting the optimal level of advertising, the strategy of the platform will depend on the probability δ to correctly classify users $\overline{\gamma}$ and $\underline{\gamma}$. The expected profits of the platform at the equilibrium can be written as:

$$\mathbb{E}\left(\Pi^T\right)^* = \delta\beta\overline{R} + (1-\beta)(1-\delta)\overline{R} + \delta(1-\beta)\underline{R}.^{14} \tag{11}$$

4.2.3 Technological choice of the platform

Given the efficiency δ of the profiling technology, the platform may or may not adopt it. We therefore compare the platform's profits with and without profiling technology to underline the condition of adoption of the profiling technology by the platform.

When $\hat{a} < \overline{m}$, Proposition 3 already demonstrates that the platform has no use for the profiling technology

[14] We recall that $\overline{R} = R(\overline{m})$ and $\underline{R} = R(\min(\hat{a}, \underline{m}))$.

Conversely, when $\hat{a} > \overline{m}$, comparing profits in Eq. (11) to Eq. (7) allows us to obtain the minimum efficiency value δ for which it is always profitable for the platform to use the profiling technology.

Proposition 5: When the platform is constrained ($\hat{a} > \overline{m}$), it always uses the technology if $\frac{\underline{R}}{\overline{R}} \geq 1-\beta$ and $\delta \in [\overline{\delta}, 1]$, or $\frac{\underline{R}}{\overline{R}} \leq 1-\beta$ and $\delta \in [\underline{\delta}, 1]$, with $\overline{\delta} = \frac{\beta\overline{R}}{\overline{R}\beta+(1-\beta)(\underline{R}-\overline{R})}$ and $\underline{\delta} = \frac{(1-\beta)(\underline{R}-\overline{R})}{\overline{R}\beta+(1-\beta)(\underline{R}-\overline{R})}$. See Proof of Proposition 5 in Appendix A.2.

Proposition 5 states that the platform will use the profiling technology when the probability of correctly classifying the types of users is sufficiently high, i.e. when $\delta > \max(\overline{\delta}, \underline{\delta})$.[15] Indeed, when the profiling technology is deficient ($\delta < \max(\overline{\delta}, \underline{\delta})$), the platform does not properly classify a part of Internet users, resulting into lower profits. For example, when users $\overline{\gamma}$ are classified as $\underline{\gamma}$, their utility decreases and they choose to avoid ads by adopting AATs, resulting in a drop of profits.

5 Some considerations about the volume of ads in equilibrium

The use of a profiling technology affects the level of ads served to consumers with respect to a situation where the platform does not use a profiling technology. In this section, we determine more precisely how the number of ads served at equilibrium changes when the platform uses the profiling technology. Do consumers see more ads when a platform uses a profiling technology?

We restrict the analysis to the cases where the platform may use the technology.[16] Proposition 6 presents two different cases.

Proposition 6: When the platform is constrained ($\hat{a} > \overline{m}$) and chooses to use the technology ($\delta > \max(\underline{\delta}, \overline{\delta})$), it affects Internet users in two ways.

- Firstly, when $\frac{\underline{R}}{\overline{R}} \geq 1-\beta$, a proportion $\delta\beta + (1-\delta)(1-\beta)$ of consumers sees the same level of ads with or without technology, whereas a proportion $\delta(1-\beta) + (1-\delta)\beta$ of consumers sees more ads. The level of ads increases so much for the proportion $(1-\delta)\beta$ of consumers that they choose to adopt AATs when they would choose to see ads without technology.

[15] When $\frac{\underline{R}}{\overline{R}} \leq 1-\beta$ we have $\underline{\delta} \leq \frac{1}{2} \leq \overline{\delta} \leq 1$, and when $\frac{\underline{R}}{\overline{R}} \geq 1-\beta$ we have $\overline{\delta} \leq \frac{1}{2} \leq \underline{\delta} \leq 1$. Therefore, we know from Proposition 5 that the platform will use the profiling technology if $\delta > \max(\overline{\delta}, \underline{\delta})$.

[16] We indeed established that when all Internet users are not ad-sensitive ($\min(\hat{a}, \overline{m}) = \hat{a}$), the platform does not use the technology.

- Secondly, when $\frac{\overline{R}}{\underline{R}} < 1-\beta$, a proportion $\delta(1-\beta)+(1-\delta)\beta$ of consumers sees the same level of ads as without technology, whereas a proportion $\delta\beta+(1-\delta)(1-\beta)$ sees less ads. The level of ads decreases so much for the proportion $\delta\beta$ of consumers that they prefer to see ads when they would prefer to avoid them without technology.

Proposition 6 shows that introducing a profiling technology greatly changes the level of ads consumers may see with respect to a situation without profiling technology.

Firstly, Proposition 2 shows that if the platform does not use a profiling technology and $\frac{\overline{R}}{\underline{R}} \geq 1-\beta$, it will set a low level of ads $a^* = \overline{m}$ and will attract all Internet users. However, when using the profiling technology, the platform sets two levels of advertising $\overline{a}^{T*} = \overline{m}$ and $\underline{a}^{T*} = \min(\hat{a}, \underline{m})$ when receiving the signals $\overline{s}$ and $\underline{s}$. The profiling technology does not however always correctly classify consumers, which may encourage some to adopt AATs. This is exactly the case when the platform tailors a level of ads $\overline{a}^{T*}$ for users $\overline{\gamma}$ but classifies some of them as $\underline{\gamma}$. In this latter case, it is preferable for the proportion $(1-\delta)\beta$ of these users to adopt AATs. The number of consumers who chooses to see ads therefore decreases with respect to the baseline case.

Second, Proposition 2 also shows that if the platform does not use a profiling technology and $\frac{\overline{R}}{\underline{R}} < 1-\beta$, it may want to set a higher level of advertising $a^* = \min(\hat{a}, \underline{m})$ to get higher profits in attracting less Internet users $\underline{\gamma}$. In the case with profiling technology, both levels of advertising $\overline{a}^T$ and $\underline{a}^T$ are still available. As the platform sets $\overline{a}^{T*} = \overline{m}$ when receiving $\overline{s}$, the level of advertising decreases for respective proportions $\delta\beta$ and $(1-\delta)(1-\beta)$ of users correctly and incorrectly classified as $\overline{\gamma}$. In this case, the proportion of users $\delta\beta$ correctly identified as $\overline{\gamma}$ chooses to see ads when they would choose to adopt AATs without profiling technology. The use of a profiling technology allows the platform in this case to increase the number of consumers who choose to see ads with respect to the baseline case.

To evaluate in the end whether the levels of ads with profiling technology $\overline{a}^{T*}$ and $\underline{a}^{T*}$ differ from the one without technology $\overline{a}^{T*}$, we need to calculate how many ads are served to consumers. Indeed, we cannot simply compare the levels of ads with or without technology, as different proportions of consumers can be targeted by two distinct levels of advertising. Basically, as an illustration, we have to compare a case without profiling technology where the platform sells 10 ads to reach 100 consumers to a case where the platform can use a profiling technology and sell on the one hand 8 ads to reach 75 consumers $\overline{\gamma}$, and 10 ads to reach 25 consumers $\underline{\gamma}$ on the other.

We denote by $V = Na^*$ and $V^T = N_{|s=\overline{s}}\overline{a}^{T*} + N_{|s=\underline{s}}\underline{a}^{T*}$, the total number of ads served by the platform without and with profiling technology. We have:

$$V = \begin{cases} \hat{a} \text{ if } \hat{a} < \overline{m}, \\ \\ \overline{m} \text{ if } \hat{a} > \overline{m} \text{ and } \frac{\overline{R}}{\underline{R}} \geq 1-\beta, \\ \\ (1-\beta)\min(\hat{a}, \underline{m}) \text{ if } \hat{a} > \overline{m} \text{ and } \frac{\overline{R}}{\underline{R}} < 1-\beta, \end{cases}$$

and,

$$V^T = \begin{cases} V \text{ if } \hat{a} < \overline{m} \text{ or } \hat{a} > \overline{m} \text{ and } \delta < \max(\underline{\delta}, \overline{\delta}), \\ \\ \overline{m}(\delta\beta + (1-\beta)(1-\delta)) + \delta(1-\beta)\min(\hat{a}, \underline{m}) \text{ if } \hat{a} > \overline{m} \text{ and } \delta > \max(\underline{\delta}, \overline{\delta}). \end{cases}$$

When online users are not very ad-sensitive, the platform will not use the technology as all users still visit the website when setting its favorite level of advertising $a^* = \hat{a}$. When at least a proportion of users is strongly ad-sensitive ($\min(\hat{a}, \overline{m}) = \overline{m}$), it is clear that when the profiling technology does not correctly classify users, i.e. when the technology is deficient ($\delta < \max(\underline{\delta}, \overline{\delta})$), the platform does not use it, and the situation remains unchanged: $V^T = V$. However, when the technology is efficient ($\delta > \max(\overline{\delta}, \underline{\delta})$), the platform uses the profiling technology to tailor the level of ads to both types of consumers. This situation has to be compared with the two possible strategies available for the platform when it does not have a profiling technology, i.e. when $V = \overline{m}$ or $V = (1-\beta)\min(\hat{a}, \underline{m})$. We establish the following proposition:

Proposition 7: When the platform is constrained ($\hat{a} > \overline{m}$) and chooses to use the technology ($\delta > \max(\underline{\delta}, \overline{\delta})$), the total number of ads served by the platform to Internet users is higher $V^T > V$ when $\frac{\overline{R}}{\underline{R}} \geq 1-\beta$. If $\frac{\overline{R}}{\underline{R}} < 1-\beta$, $V^T > V$ if $\delta \in [\underline{\delta_V}, 1]$, with $\underline{\delta_V} = \frac{(1-\beta)(\min(\hat{a},\underline{m})-\overline{m})}{(1-\beta)(\min(\hat{a},\underline{m})-\overline{m})+\beta\overline{m}}$.

See Proof of Proposition 7 in Appendix A.3.

Proposition 7 shows that the platform will serve more ads in total when $\frac{\overline{R}}{\underline{R}} \geq 1-\beta$, while it is not always the case when $\frac{\overline{R}}{\underline{R}} < 1-\beta$.

Indeed, according to Proposition 6, introducing a technology when $\frac{\overline{R}}{\underline{R}} \geq 1-\beta$ decreases the number of Internet users seeing as. On the one hand, the platform attracts all Internet users without technology, while it misclassifies a proportion of strongly ad-sensitive users when using one. On the other hand, the technology allows to set a higher level of ads to a proportion of correctly classified weakly ad-sensitive users. All in all, we find that the higher level of ads shown to correctly classified weakly ad-sensitive users always offsets the decrease in the number of Internet users seeing advertising.

Conversely, according to Proposition 6, introducing a technology when $\frac{\overline{R}}{\underline{R}} < 1-\beta$ allows to attract more users on the platform. On the one hand, correctly classified strongly ad-sensitive users choose not to avoid advertising and see a low level of ads. On the other hand, misclassified weakly ad-sensitive users see less ads than before. All in all, we show that the higher number of Internet users seeing ads offsets the lower level of ads seen by misclassified weakly ad-sensitive user only if the profiling technology does not misclassify too much.

6 Winners and Losers of the Profiling Technology

We analyze how the introduction of a profiling technology may affect consumers, advertisers, and the platform. In doing that, we define total welfare as the sum of the platform profits (Π), the advertisers surplus (S_a) and the users surplus (S_u). We compute the total welfare without technology (W) and the total welfare with the profiling technology (W^T) when it is good enough to be used by the platform (i.e. when $\delta > \max(\overline{\delta}, \underline{\delta})$).

6.1 Internet user surplus

Two cases are interesting to analyze when $\hat{a} > \overline{m}$.[17]

Proposition 8: When the platform is constrained ($\hat{a} > \overline{m}$) and chooses to use the technology ($\delta > \max(\underline{\delta}, \overline{\delta})$):

- If $\frac{\overline{R}}{\underline{R}} > (1-\beta)$, $\mathrm{S}_u^T > \mathrm{S}_u$ if $\underline{\gamma}(\min(\hat{a}, \underline{m})) > \underline{\gamma}(\overline{m})$.
- If $\frac{\overline{R}}{\underline{R}} < (1-\beta)$, $\mathrm{S}_u^T > \mathrm{S}_u$ if $\delta \neq 1$ and $\underline{\gamma}(\min(\hat{a}, \underline{m})) < \underline{\gamma}(\overline{m})$. See Proof of Proposition 8 in Appendix A.4

Proposition 8 underlines two important results. Firstly, the impact of the profiling technology on users surplus largely depends on the utilities associated to a correct or a misclassification. For example, an Internet users could tolerate a high advertising intensity, but may prefer a lower one, hence prefering to be misclassfied by the technology. Conversely, a user may want to receive more ads on a specific subject, hence generating a higher utility from being correctly classifed. Secondly, it also depends on the structure of the market, as the impact of introducing a technology depends on the advertising level set by the platform without technology.

Indeed, when $\frac{\overline{R}}{\underline{R}} > (1-\beta)$, introducing a profiling technology will increase consumer surplus only if weakly ad-sensitive users prefer to be correctly classified ($\underline{\gamma}(\min(\hat{a}, \underline{m})) >$

[17]The consumer surplus will not be affected by the introduction of technology when $\hat{a} = \hat{a} < \overline{m}$. Indeed, Proposition 3 shows that the technology will not be used by the platform in this case.

$\underline{\gamma}(\overline{m})$). [18] This is due to the fact that the platform attract all Internet users with a low level of advertising, hence showing few ads to users weakly ad-sensitive. Hence, using the technology allows to show more ads to weakly ad-sensitive users. The impact of the technology on user surplus therefore depends on whether weakly ad-sensitive users prefer to see more ads (being correctly classified) or fewer ads (being misclassified).

The inverse situation arises when $\frac{\overline{R}}{\underline{R}} < (1-\beta)$ as introducing an imperfect profiling technology increases consumer surplus only if weakly ad-sensitive users prefer to be misclassified ($\underline{\gamma}(\overline{m}) > \underline{\gamma}(\min(\hat{a}, \underline{m}))$). In this situation, the platform chooses to strip out weakly ad-sensitive users of their utility while excluding strongly sensitive users when he has no profiling technology. Hence, introducing a technology would reduce the level of ads shown to misclassified weakly ad-sensitive users. Consequently, user surplus increase with a profiling technology if weakly ad-sensitive users prefers a lower level of ads (being misclassified) than a high level of ads (being correctly classified).

6.2 Advertiser surplus

Like the consumer surplus analysis, we only analyze the situation encountered when the platform is constrained $\hat{a} > \overline{m}$. More generally, introducing a perfect profiling technology will increase advertisers surplus. However, an imperfect technology may have different effect as the platform have different practice when it has no profiling technology.

Proposition 9: When the platform is constrained ($\hat{a} > \overline{m}$) and chooses to use the technology ($\delta > \max(\underline{\delta}, \overline{\delta})$), advertisers always benefit from it $\mathrm{S}_a^T > \mathrm{S}_a$ if $\frac{\overline{R}}{\underline{R}} \geq 1-\beta$. If $\frac{\overline{R}}{\underline{R}} \leq 1-\beta$, advertisers benefit from the technology if $\delta \in [\underline{\delta_{s_a}}, 1]$, with $\underline{\delta_{s_a}} = \frac{(1-\beta)(\underline{s_a}-\overline{s_a})}{\overline{s_a}\beta+(1-\beta)(\underline{s_a}-\overline{s_a})}$. See Proof of Proposition 9 in Appendix A.5.

Proposition 9 highlights important results for advertisers. When $\frac{\overline{R}}{\underline{R}} \geq (1-\beta)$, the advertisers will gain from the profiling technology as long as it it is beneficial for the platform to introduce it only. Conversely, if $\frac{\overline{R}}{\underline{R}} < (1-\beta)$, there is an area on δ where introducing the technology is profitable for the platform but lower the advertisers surplus.

In this case, the platform may choose to adopt the technology even if it lowers advertisers surplus, hence making their situation worse than without technology. This is due to the fact that an efficient profiling technology allows a better detection of strongly ad-sensitive users, which require a lower level of ad, hence decreasing advertisers surplus. This result exerts a strong

[18] As $\underline{\gamma}(\overline{m}) > \underline{\gamma}(\underline{m})$, the profiling technology can only increase consumer surplus when the platform is weakly constrained $\min(\hat{a}, \underline{m}) = \hat{a}$) and $\underline{\gamma}(\overline{m}) < \underline{\gamma}(\hat{a})$. This is intuitive as a correct identification in this case generates higher utility than a misclassification for Internet users weakly ad-sensitive.

impact on the computation of total welfare.

6.3 Profits of the platform

The analysis follows the same mechanisms as in Proposition 5. The platform choose to use the technology only if it generates high enough profits, i.e if the technology is good enough. Therefore the condition for the technology to be profit increasing are the same than in Proposition 5.

6.4 Total welfare

We simplify the analysis in writing the industry profits as $\underline{G} = \underline{R} + \underline{s_a}$ and $\overline{G} = \overline{R} + \overline{s_a}$ and the incentives to be correctly identified for weakly ad-sensitive consumers as $I^{\underline{\gamma}} = \underline{\gamma}(\min(\hat{a}, \underline{m})) - \underline{\gamma}(\overline{m})$.

Proposition 10: When the platform is constrained ($\hat{a} > \overline{m}$) and chooses to use the technology ($\delta > \max(\underline{\delta}, \overline{\delta})$):

- If $\frac{\overline{R}}{\underline{R}} > (1-\beta)$, $W^T > W$ if $\delta(\beta\overline{G} + (1-\beta)(\underline{G} - \overline{G}) + (1-\beta)I^{\underline{\gamma}}) > \beta\overline{G}$
- If $\frac{\overline{R}}{\underline{R}} < (1-\beta)$, $W^T > W$ if $\delta(\beta\overline{G} + (1-\beta)(\underline{G} - \overline{G}) + (1-\beta)I^{\underline{\gamma}}) > (1-\beta)(\underline{G} - \overline{G}) + (1-\beta)I^{\underline{\gamma}}$. See Proof of Proposition 10 in Appendix A.6

Proposition 10 underline different situations. We analyze only the cases where the technology is efficient enough and have an impact on market equilibrias (that is when the platform is constrained by the nuisance of Internet users when setting its level of advertising).

A first situation arises when $\frac{\overline{R}}{\underline{R}} > (1-\beta)$. In that case, the industry profits - that is the platform profits and the advertisers surplus - always benefits from the introduction of an efficient technology. Two cases can arise. Firstly, if weakly ad-sensitive users generate more utility from being correctly identified ($\underline{\gamma}(\min(\hat{a}, \underline{m})) > \underline{\gamma}(\overline{m})$), the introduction of technology will be welfare increasing as it improves the situation of all agents in the market. Secondly, if weakly ad-sensitive Internet users generate more utility from being misclassified as strongly ad-sensitive ($\underline{\gamma}(\min(\hat{a}, \underline{m})) < \underline{\gamma}(\overline{m})$), the introduction of technology will be welfare increasing if the opportunity cost of not being misclassifed is not too high ($\underline{\gamma}(\min(\hat{a}, \underline{m})$ close to $\underline{\gamma}(\overline{m})$).

A second situation arises when $\frac{\overline{R}}{\underline{R}} < (1-\beta)$. In this case, the industry profit may not increase with the introduction of the technology, as platform may use the technology while it decreases advertisers surplus. more precisely, three different situations may occur. Firstly, if the introduction of an efficient technology decreases industry surplus [19] and weakly ad-sensitive Internet users earn higher utility from being correctly classified ($\underline{\gamma}(\min(\hat{a}, \underline{m})) > \underline{\gamma}(\overline{m})$), it will be welfare decreasing as both industry profits and Internet users surplus are lower with technology than without. Secondly, if the introduction of an efficient technology increases industry surplus and weakly ad-sensitive Internet users earn higher utility from being misclassified ($\underline{\gamma}(\hat{a}, \underline{m})) < \underline{\gamma}(\overline{m})$), it everyone benefit from the technology. Finally, if the technology decreases industry surplus and increases Internet user surplus or increase industry surplus and decrease Internet user surplus, the introduction of such technology may have an ambiguous effect on welfare. Results are sumed up in Figure 4.

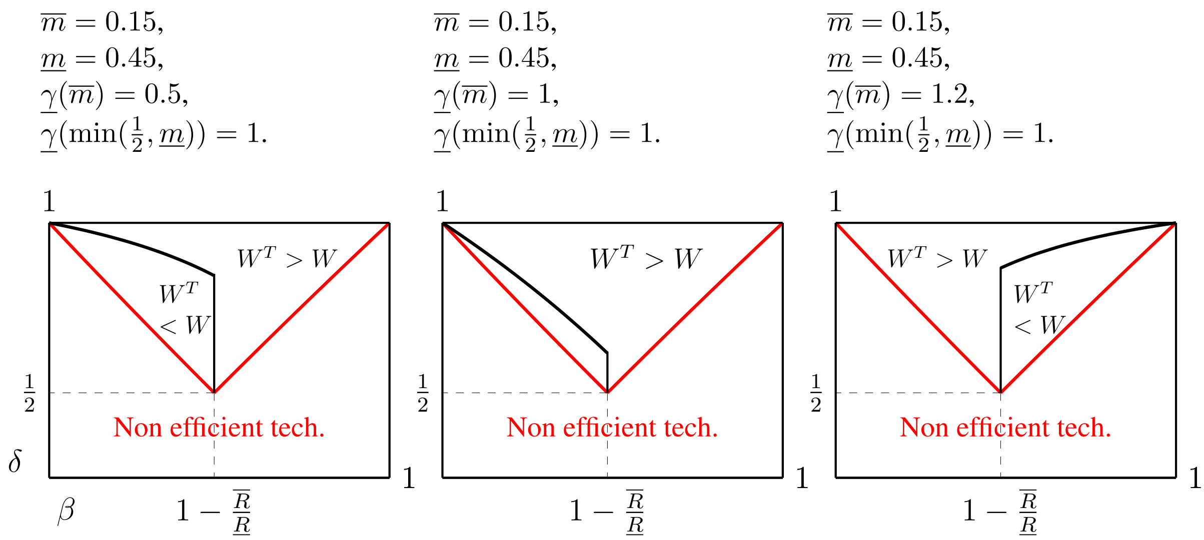

Figure 4: Welfare impact of technology with $r(a) = 1 - a$

7 Strategic implications and conclusion

In this part, we summarize the key results and implications of the model, and concludes.

We observe that the main findings depends on whether attracting users who are strongly ad-sensitive is profitable or not for the platform. We recall from Proposition 2 that when there is no profiling technology, the publisher will choose to set a low level of advertising and attract all Internet users if $\frac{\overline{R}}{\underline{R}} > 1 - \beta$ or set a high level of advertising and only attract the least sensitive Internet users if $\frac{\overline{R}}{\underline{R}} < 1 - \beta$.

Let assume for simplicity that the marginal revenue function $R(a)$ of the publisher is fixed, as well as the maximum advertising level users of each population are willing tolerate $\overline{m}$ and

[19]The technology is efficient enough to be used by the platform but sharply decreases advertisers surplus ($\underline{\delta_{s_a}} > \delta > \underline{\delta}$)

$\underline{m}$. [20] In this case, the incentives of the platform to attract or not the entire audience depends on the proportion β of strongly sensitive Internet users visiting its website. Firstly, we note that when there is a dominant proportion of strongly sensitive Internet user (high β), the publisher is likely to attract all Internet users with a low level of advertising. When this is the case, we consider that the platform is facing a strongly elastic audience that creates incentives to lower its advertising level. Conversely, when there is a dominant proportion of weakly sensitive Internet users (low β), the publisher will be more likely to set a high level of advertising and attract only weakly ad-sensitive Internet users. In that case, the platform is facing an audience we define as weakly elastic to advertising.

From above, we consider that the platform can encounter two types of audience: strongly elastic or weakly elastic to advertising level. For example, illegale streaming or downloading websites such as The Pirate Bay or DpStream generally face an audience weakly elastic to advertising, and choose to set a high level of advertising. Conversely, News websites such as The New York Times, The Washington Post or The Financial Times adress a larger audience which is significantly more elastic to advertising levels, hence fostering them to lower the level of advertising.

The platform sets its advertising level depending on the type of audience it attracts, knowing its marginal revenue function $R(a)$. Considering this typology of audiences, we are able to derive key strategic implications:

Key Result 1:

- On a platform facing an audience weakly elastic to advertising, the introduction of a profiling technology reduces AAT demand, and therefore increases the number Internet users seeing advertising.
- Conversely, on a platform facing an audience strongly elastic to advertising, the introduction of a profiling technology increases AAT demand, and therefore decreases the number of Internet users watching ads.

This result directly stems from Proposition 6. Without profiling technology, a platform facing weakly elastic audience choose to focus on attracting only Internet users who are the less ad-sensitive. Hence, introducing a profiling technology allows the platform to attract strongly ad-sensitive users in offering a better Internet user experience. For illustration, introducing a profiling technology on DpStream would improve the user experience of a part of the audience.

[20] which defines $\overline{R}$ and $\underline{R}$ as $R(a^* = \overline{m}) = \overline{R}$ and $R(a^* = \underline{m}) = \underline{R}$

Conversely, without profiling technology, a platform facing a strongly elastic audience chooses to attract the entire audience of Internet users. Therefore, introducing a profiling technology may degrade the experience of strongly ad-sensitive users who are misclassified, and therefore foster them to use AATs instead of watching ads. As an example, introducing a profiling technology on The New York Times would deteriorate the user experience of a part of the audience.

A subsequent analysis can be drawn on the impact of the profiling technology on welfare. Propositions 10 determines the conditions for which the introduction of such technologies is welfare increasing. Mapping the propositions to our analysis allows to underline the following key results:

Key Result 2:

- On a platform facing an audience weakly elastic to advertising, the introduction of an efficient profiling technology is welfare increasing if Internet users benefits more from being misclassified than correctly classified and advertisers benefit from the technology.
- Conversely, on a platform facing an audience strongly elastic to advertising, the introduction of a profiling technology is always welfare increasing if Internet users benefits more from being correctly classified than misclassified.

Key result 2 highlight that the impact of the introduction of a profiling technologies depends on the audience of the website that adopts it. From the analysis on Proposition 10, we understand that if the profiling technlogy is highly efficient, its introduction is welfare enhancing no matter the type of websites. However, if the profiling technology is less efficient, the introduction of a profiling technology impacts welfare differently depending on the structure of its audience.

For example, introducing a profiling technology on a platform facing a strongly elastic audience such as The Washington Post is welfare increasing if this very audience is composed of internet users that enjoy being correctly classified (or don't care about their classification). Indeed, on the one hand, both advertisers and the platform benefit from the introduction of a profiling technology. On the other hand, the less elastic fringe of the audience is seeing more ads with profiling technology than without. Therefore, the impact of the profiling technology on total welfare obviously depends on whether the audience prefers to be correctly classified (seeing more ads) or not (seeing less ads).

Conversely, introducing the profiling technology on platform facing a weakly elastic demand such as The Pirate Bay may have more ambiguous impact on total welfare. Firstly, the advertisers do not always enjoy the introduction of such technology when facing that type of audience, as the platform practice lower level of ads. Secondly, the less elastic fringe of the audience is seeing less ads with profiling technology than without in this case. All in all, the impact of the profiling technology on total welfare depends on whether the audience prefers to be correctly classified (seeing more ads) or not (seeing less ads) and if the advertisers extract benefit from the profiling technology. Therefore, introducing a profiling technology on a platform facing a weakly elastic demand such as DpStream is less likely to be welfare improving

The efficiency of the profiling technology relies on two important points. On the one hand, the profiling technology must use a good enough classification technology (such as machine learning), which may represent an investment for the platform. On the other hand, the profiling technology uses Internet user personal information to classify as averse or not to advertising. Thus, the fact that more and more Internet users tend to protect themselves using Privacy Enhancing Tools (PETs) may lower quality of the profiling technology. This last point is crucial for regulators. For example, the European commission is regulating the use of personal information for business purpose, giving more market power to Internet user. For this different reasons, one of the limit of the paper is to consider the efficiency of the profiling technology as exogenous, while it is endogenous to the choice of Internet users and depends on the regulation. Future research may therefore consider the link between the willingness to share information of Internet users, the privacy regulation operating on the market and the efficiency of profiling technologies used by firms to serve advertising.

References

Abrams, Zoë and Erik Vee, "Personalized Ad Delivery When Ads Fatigue: An Approximation Algorithm," in Xiaotie Deng and Fan Chung Graham, eds., *Internet and Network Economics: Third International Workshop, WINE 2007, San Diego, CA, USA, December 12-14, 2007. Proceedings*, Springer Berlin Heidelberg, 2007, pp. 535–540.

Agarwal, Ashish, Kartik Hosanagar, and Michael D Smith, "Location, Location, Location: An Analysis of Profitability of Position in Online Advertising Markets," *Journal of Marketing Research*, 2011, *48* (6), 1057–1073.

Anderson, Simon and Bruno Jullien, "The advertising-financed business model in two-sided media markets," Technical Report TSE 632, Toulouse School of Economics 2016.

— **and Jean Gabszewicz**, "Chapter 18 The Media and Advertising: A Tale of Two-Sided Markets," in Victor A. Ginsburg and David Throsby, eds., *Handbook of the Economics of Art and Culture*, Vol. 1, Elsevier, 2006, pp. 567 – 614.

Anderson, Simon P. and Joshua S. Gans, "Platform Siphoning: Ad-Avoidance and Media Content," *American Economic Journal: Microeconomics*, November 2011, *3* (4), 1–34.

— **and Stephen Coate**, "Market Provision of Broadcasting: A Welfare Analysis," *The Review of Economic Studies*, 2005, *72* (4), 947–972.

Andrews, Michelle, Xueming Luo, Zheng Fang, and Anindya Ghose, "Mobile Ad Effectiveness: Hyper-Contextual Targeting with Crowdedness," *Marketing Science*, 2016, *35* (2), 218–233.

Armstrong, Mark, "Competition in two-sided markets," *The RAND Journal of Economics*, 2006, *37* (3), 668–691.

Athey, Susan and Joshua S Gans, "The impact of targeting technology on advertising markets and media competition," *Available at SSRN 1535325*, 2010.

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

Blake, Thomas, Chris Nosko, and Steven Tadelis, "Consumer Heterogeneity and Paid Search Effectiveness: A Large-Scale Field Experiment," *Econometrica*, 2015, *83* (1), 155–174.

Bleier, Alexander and Maik Eisenbeiss, "Personalized Online Advertising Effectiveness: The Interplay of What, When, and Where," *Marketing Science*, 2015, *34* (5), 669–688.

Buchbinder, Niv, Moran Feldman, Arpita Ghosh, and Joseph (Seffi) Naor, "Frequency Capping in Online Advertising," in Frank Dehne, John Iacono, and Jörg-Rüdiger Sack, eds., *Algorithms and Data Structures: 12th International Symposium, WADS 2011, New York, NY, USA, August 15-17, 2011. Proceedings*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 147–158.

Caillaud, Bernard and Bruno Jullien, "Competing Cybermediaries," *European Economic Review*, 2001, *45*, 797–808.

— **and** — , "Chicken and Egg: Competition among Intermediation Service Providers," *The RAND Journal of Economics*, 2003, *34* (2), 309–328.

Davis, H., *Google Advertising Tools: Cashing in with AdSense, AdWords, and the Google APIs* O'Reilly Series, O'Reilly Media, 2006.

de Cornière, Alexandre, "Search Advertising," *American Economic Journal: Microeconomics*, 2016, *8* (3), 156–188.

Dukes, Anthony and Esther Gal-Or, "Negotiations and Exclusivity Contracts for Advertising," *Marketing Science*, 2003, *22* (2), 222–245.

Estrada-Jiménez, José, Javier Parra-Arnau, Ana Rodríguez-Hoyos, and Jordi Forné, "Online advertising: Analysis of privacy threats and protection approaches," *Computer Communications*, 2017, *100*, 32 – 51.

Evans, David S., "The Online Advertising Industry: Economics, Evolution, and Privacy," *Journal of Economic Perspectives*, September 2009, *23* (3), 37–60.

— , "The online advertising industry: Economics, evolution, and privacy," *The journal of economic perspectives*, 2009.

Falkinger, Josef, "Limited Attention as a Scarce Resource in Information-Rich Economies," *The Economic Journal*, 2008, *118* (532), 1596–1620.

Flosi, Stephanie, Gian Fulgoni, and Andrea Vollman, "If an Advertisement Runs Online And No One Sees It, Is It Still an Ad?," *Journal of Advertising Research*, June 2013, *53* (2), 192–199.

Ghose, Anindya and Sha Yang, "An Empirical Analysis of Search Engine Advertising: Sponsored Search in Electronic Markets," *Management Science*, 2009, *55* (10), 1605–1622.

Ghosh, Bikram and Axel Stock, "Advertising Effectiveness, Digital Video Recorders, and Product Market Competition," *Marketing Science*, 2010, *29* (4), 639–649.

Goldfarb, Avi and Catherine Tucker, "Advertising bans and the substitutability of online and offline advertising," *Journal of Marketing Research*, 2011, *48* (2), 207–227.

_ and _ , "Online Display Advertising: Targeting and Obtrusiveness," *Marketing Science*, 2011, *30* (3), 389–404.

_ and _ , "Standardization and the Effectiveness of Online Advertising," *Management Science*, 2015, *61* (11), 2707–2719.

Grossman, Gene M. and Carl Shapiro, "Informative Advertising with Differentiated Products," *The Review of Economic Studies*, 1984, *51* (1), 63–81.

Hagiu, Andrei and Hanna Halaburda, "Information and Two-Sided Platform Profits," SSRN Scholarly Paper ID 2360263, Social Science Research Network, Rochester, NY April 2014.

Hermalin, Benjamin E and Michael L Katz, "Privacy, property rights and efficiency: The economics of privacy as secrecy," *Quantitative Marketing and Economics*, 2006, *4* (3), 209–239.

Johnson, Justin P, "Targeted advertising and advertising avoidance," *The RAND Journal of Economics*, 2013, *44* (1), 128–144.

Katz, Michael L. and Carl Shapiro, "Network Externalities, Competition, and Compatibility," *The American Economic Review*, 1985, *75* (3), 424–440.

Lambrecht, Anja and Catherine Tucker, "When Does Retargeting Work? Information Specificity in Online Advertising," *Journal of Marketing Research*, 2013, *50* (5), 561–576.

Manchanda, Puneet, Jean-Pierre Dubé, Khim Yong Goh, and Pradeep K. Chintagunta, "The Effect of Banner Advertising on Internet Purchasing," *Journal of Marketing Research*, 2006, *43* (1), 98–108.

Peitz, Martin and Markus Reisinger, "The Economics of Internet Media," in "Handbook of Media Economics, vol. 1A," Elsevier, 2016.

Pepelyshev, Andrey, Yuri Staroselskiy, and Anatoly Zhigljavsky, "Adaptive Targeting for Online Advertisement," in Panos Pardalos, Mario Pavone, Giovanni Maria Farinella, and Vincenzo Cutello, eds., *Machine Learning, Optimization, and Big Data: First International Workshop, MOD 2015, Taormina, Sicily, Italy, July 21-23, 2015, Revised Selected Papers*, Cham: Springer International Publishing, 2015, pp. 240–251.

Rochet, Jean-Charles and Jean Tirole, "Cooperation among Competitors: Some Economics of Payment Card Associations," *The RAND Journal of Economics*, 2002, *33* (4), 549–570.

Rutz, Oliver J and Randolph E Bucklin, "From Generic to Branded: A Model of Spillover in Paid Search Advertising," *Journal of Marketing Research*, 2011, *48* (1), 87–102.

Sajjacholapunt, Pitch and Linden J. Bal, "The influence of banner advertisements on attention and memory: human faces with averted gaze can enhance advertising effectiveness," *Frontiers in Psychology*, 2014, *5*, 166.

Sanghvi, Nikul, "Cost Per Hour. Using a Time-Based Currency for Digital Advertising," Technical Report, Analytics Consultant 2015.

Shiller, Ben, Joel Waldfogel, and Johnny Ryan, "Will Ad Blocking Break the Internet?," Working Paper 23058, National Bureau of Economic Research January 2017.

Sun, Monic and Feng Zhu, "Ad Revenue and Content Commercialization: Evidence from Blogs," *Management Science*, 2013, *59* (10), 2314–2331.

Tag, J., "Paying to Remove Advertisements," *Information Economics and Policy*, 2009, *21* (4), 245–252.

Toubiana, Vincent, Arvind Narayanan, Dan Boneh, Helen Nissenbaum, and Solon Barocas, "Adnostic: Privacy Preserving Targeted Advertising," in "Proceedings of the Network and Distributed System Security Symposium, NDSS 2010, San Diego, California, USA, 28th February - 3rd March 2010" 2010.

Tucker, Catherine E., "The economics of advertising and privacy," *International Journal of Industrial Organization*, 2012, *30* (3), 326–329.

Turow, Joseph, Jennifer King, Chris Jay Hoofnagle, Amy Bleakley, and Michael Hennessy, "Americans Reject Tailored Advertising and Three Activities that Enable It," Working Paper 1478214, SSRN September 2009.

Vratonjic, Nevena, Mohammad Hossein Manshaei, Jens Grossklags, and Jean-Pierre Hubaux, "Ad-blocking games: Monetizing online content under the threat of ad avoidance," in "The Economics of Information Security and Privacy," Springer, 2013, pp. 49–73.

Wang, Chong, Achir Kalra, Cristian Borcea, and Yi Chen, "Viewability Prediction for Online Display Ads," in "Proceedings of the 24th ACM International on Conference on Information and Knowledge Management" CIKM '15 ACM New York, NY, USA 2015, pp. 413–422.

Zhang, Kaifu and Miklos Sarvary, "Differentiation with User-Generated Content," *Management Science*, 2015, *61* (4), 898–914.

Zhang, Weinan, Ye Pan, Tianxiong Zhou, and Jun Wang, "An Empirical Study on Display Ad Impression Viewability Measurements," *arXiv:1505.05788 [cs]*, May 2015. arXiv: 1505.05788.

A Appendix

A.1 Proof of Proposition 2

We compare the profits of the platform when it serves consumers $\overline{\gamma}$ and $\underline{\gamma}$ to the ones when it serves only consumers $\underline{\gamma}$:

- $\theta \leq 1 + \underline{\gamma}(\hat{a}(v))$: $\Pi(a^* = \hat{a}) = (1 - \beta)R(a^* = \hat{a})$
- $\theta \geq 1 + \underline{\gamma}(\hat{a}(v))$: $\Pi(a^* = \underline{m}) = (1 - \beta)R(a^* = \underline{m})$

The threshold $\frac{\overline{R}}{\underline{R}} > 1 - \beta$ results directly from the comparison.

A.2 Proof of Proposition 5

Two cases have to be analyzed.

- First case: $\frac{\overline{R}}{\underline{R}} > 1 - \beta$.

  Two strategies are available, which gives the following profits:

$$\Pi^* = \begin{cases} \mathbb{P}(\overline{s})\overline{R} + \mathbb{P}(\underline{s})\mathbb{P}(\underline{\gamma}|\underline{s})\underline{R} \text{ if the platform follows the signal,} \\ \\ \overline{R} \text{ else.} \end{cases}$$

  Comparing both profits leads directly to the rule that following the signal is more profitable for $\delta \in [\frac{\overline{R}\beta}{\overline{R}\beta+(1-\beta)(\underline{R}-\overline{R})}, 1]$. Else, ignoring it is more profitable.

- Second case: $\frac{\overline{R}}{\underline{R}} \leq 1 - \beta$.

  Again two strategies are available, either following the signal or not:

$$\Pi^* = \begin{cases} \mathbb{P}(\overline{s})\overline{R} + \mathbb{P}(\underline{s})\mathbb{P}(\underline{\gamma}|\underline{s})\underline{R} \text{ if the platform follows the signal,} \\ \\ (1-\beta)\underline{R} \text{ else.} \end{cases}$$

  Comparing both profits leads directly to the rule that following the signal is more profitable for $\delta \in [\frac{(\underline{R}-\overline{R})(1-\beta)}{(\underline{R}-\overline{R})(1-\beta)+\overline{R}\beta}, 1]$. Else, ignoring it is more profitable.

It is immediate to see that both values are below 1, and a direct resolution allows to verify that they are above $\frac{1}{2}$.

A.3 Proof of Proposition 7

- When $\frac{\overline{R}}{\underline{R}} \geq 1 - \beta$, $V^T > V$ if $\delta \in [\overline{\delta_V}, 1]$.
- When $\frac{\overline{R}}{\underline{R}} < 1 - \beta$, $V^T > V$ if $\delta \in [\underline{\delta_V}, 1]$.

With $\overline{\delta_V} = \frac{\beta\overline{m}}{(1-\beta)(\min(\hat{a},\underline{m})-\overline{m})+\beta\overline{m}}$ and $\underline{\delta_V} = \frac{(1-\beta)(\min(\hat{a},\underline{m})-\overline{m})}{(1-\beta)(\min(\hat{a},\underline{m})-\overline{m})+\beta\overline{m}}$.

However, we know from Proposition 5 that for the technology to be used by the platform, it has to be efficient enough such that $\delta > \max(\underline{\delta}, \overline{\delta})$.

We therefore compare the conditions for which the platform would use the technology and the volume of ads may increase with profiling technology. We find that as $r(a)$ is decreasing in a, $r(\overline{m}) > r(\min(\hat{a}, \underline{m}))$ and therefore $\overline{\delta_V} < \overline{\delta}$ and $\underline{\delta_V} > \underline{\delta}$.

Hence, we are able to write that :

- When $\frac{\overline{R}}{\underline{R}} \geq 1-\beta$, $V^T > V$
- When $\frac{\overline{R}}{\underline{R}} < 1-\beta$, $V^T > V$ if $\delta \in [\underline{\delta_V}, 1]$.

with $\underline{\delta_V} = \frac{(1-\beta)(\min(\hat{a},\underline{m})-\overline{m})}{(1-\beta)(\min(\hat{a},\underline{m})-\overline{m})+\beta\overline{m}}$.

A.4 Proof of Proposition 8

$$\mathrm{S}_u = \overline{NU} + \underline{NU} = \begin{cases} \beta(1+\overline{\gamma}(\hat{a})) + (1-\beta)(1+\underline{\gamma}(\hat{a})) \text{ if } \hat{a} < \overline{m} \\ \beta\theta + (1-\beta)(1+\underline{\gamma}(\overline{m})) \text{ if } \hat{a} > \overline{m} \text{ and } \frac{\overline{R}}{\underline{R}} > (1-\beta), \\ \beta\theta + (1-\beta)(1+\underline{\gamma}(\min(\hat{a},\underline{m}))) \text{ if } \hat{a} > \overline{m} \text{ and } \frac{\overline{R}}{\underline{R}} < (1-\beta), \end{cases} \tag{12}$$

$$\mathrm{S}_u^T = \begin{cases} \beta(1+\overline{\gamma}(\hat{a})) + (1-\beta)(1+\underline{\gamma}(\hat{a})) \text{ if } \hat{a} < \overline{m} \\ \beta\theta + (1-\delta)(1-\beta)(1+\underline{\gamma}(\overline{m})) + \\ \delta(1-\beta)(1+\underline{\gamma}(\min(\hat{a},\underline{m})) \text{ if } \hat{a} > \overline{m} \end{cases} \tag{13}$$

A.5 Proof of Proposition 9

$$\mathrm{S}_a = N\int_0^{a^*} (r(a)-r(a^*))\mathrm{d}a = \begin{cases} s_a(\hat{a}) \text{ if } \hat{a} < \overline{m} \\ s_a(\overline{m}) \equiv \overline{s_a} \text{ if } \hat{a} > \overline{m} \text{ and } \frac{\overline{R}}{\underline{R}} > (1-\beta), \\ (1-\beta)s_a(\min(\hat{a},\underline{m})) \equiv (1-\beta)\underline{s_a} \text{ if } \hat{a} > \overline{m} \text{ and } \frac{\overline{R}}{\underline{R}} < (1-\beta), \end{cases} \tag{14}$$

$$\mathrm{S}_a^T = \begin{cases} s_a(\hat{a}) \text{ if } \hat{a} < \overline{m} \\ (\delta\beta + (1-\delta)(1-\beta))\overline{s_a} + \delta(1-\beta)\underline{s_a} \text{ if } \hat{a} > \overline{m} \end{cases} \tag{15}$$

When the platform is constrained ($\hat{a} > \overline{m}$), the advertisers benefit from the technology $\mathrm{S}_a^T > \mathrm{S}_a$ if $\frac{\overline{R}}{\underline{R}} \geq 1-\beta$ and $\delta \in [\overline{\delta_{s_a}}, 1]$, or $\frac{\overline{R}}{\underline{R}} \leq 1-\beta$ and $\delta \in [\underline{\delta_{s_a}}, 1]$, with $\overline{\delta_{s_a}} = \frac{\beta\overline{s_a}}{\overline{s_a}\beta+(1-\beta)(\underline{s_a}-\overline{s_a})}$ and $\underline{\delta_{s_a}} = \frac{(1-\beta)(\underline{s_a}-\overline{s_a})}{\overline{s_a}\beta+(1-\beta)(\underline{s_a}-\overline{s_a})}$.

However, we know from Proposition 5 that for the technology to be used by the platform, it has to be efficient enough such that $\delta > \max(\underline{\delta}, \overline{\delta})$.

We compare the conditions that insure benefits from the technology for advertisers and the platform. We find that $\frac{\overline{R}}{\underline{R}} \geq 1-\beta$ (resp. $\frac{\overline{R}}{\underline{R}} < 1-\beta$) the condition for the platform to benefit from the profiling technology overlooks the one for advertisers i.e $\overline{\delta} > \overline{\delta_{s_a}}$ (resp. $\underline{\delta} > \underline{\delta_{s_a}}$) if $\frac{\overline{R}}{\underline{R}} > \frac{\overline{s_a}}{\underline{s_a}}$ (resp. $\frac{\overline{R}}{\underline{R}} < \frac{\overline{s_a}}{\underline{s_a}}$).

We call $G(a)$ the antiderivative of $r(a)$. We therefore can rewrite the above condition in $\frac{\overline{R}}{\underline{R}} > \frac{\overline{G}-\overline{R}}{\underline{G}-\underline{R}}$ (resp. $<$), which can be simplified to $\frac{\overline{R}}{\underline{R}} > \frac{\overline{G}}{\underline{G}}$. Computing the derivative of G and R

with respect to a, we find that $G' = r$ and $R' = ar' + r$. As $r' < 0$, $G' > R'$, which prove the condition $\frac{\overline{R}}{\underline{R}} > \frac{\overline{G}}{\underline{G}}$ to be always true.

We can therefore wright that if $\frac{\overline{R}}{\underline{R}} \geq 1 - \beta$, $\mathrm{S}_a^T > \mathrm{S}_a$ and if $\frac{\overline{R}}{\underline{R}} < 1 - \beta$, $\mathrm{S}_a^T > \mathrm{S}_a$ only if $\delta \in [\underline{\delta_{s_a}}, 1]$ with $\underline{\delta_{s_a}} = \frac{(1-\beta)(\underline{s_a} - \overline{s_a})}{\overline{s_a}\beta + (1-\beta)(\underline{s_a} - \overline{s_a})}$.

A.6 Proof of Proposition 10

$$W = \begin{cases} R(\hat{a}) + s_a(\hat{a}) + 1 + \beta\overline{\gamma}(\hat{a}) + (1-\beta)\underline{\gamma}(\hat{a}) \text{ if } \hat{a} < \overline{m} \\ \\ \overline{R} + \overline{s_a} + \beta\theta + (1-\beta)(1 + \underline{\gamma}(\overline{m})) \text{ if } \hat{a} > \overline{m} \text{ and } \frac{\overline{R}}{\underline{R}} > (1-\beta), \\ \\ (1-\beta)(\underline{R} + \underline{s_a}) + \beta\theta + \\ (1-\beta)(1 + \underline{\gamma}(\min(\hat{a}, \underline{m}) \text{ if } \hat{a} > \overline{m} \text{ and } \frac{\overline{R}}{\underline{R}} < (1-\beta), \end{cases} \tag{16}$$

$$W^T = \begin{cases} R(\hat{a}) + s_a(\hat{a}) + 1 + \beta\overline{\gamma}(\hat{a}) + (1-\beta)\underline{\gamma}(\hat{a}) \text{ if } \hat{a} < \overline{m} \\ \\ (\delta\beta + (1-\delta)(1-\beta))(\overline{R} + \overline{s_a}) + \delta(1-\beta)(\underline{R} + \underline{s_a}) \\ +\beta\theta + (1-\delta)(1-\beta)(1 + \underline{\gamma}(\overline{m})) + (1-\beta)(1 + \underline{\gamma}(\min(\hat{a}, \underline{m})) \text{ if } \hat{a} > \overline{m} \end{cases} \tag{17}$$