

The Welfare Impact of Targeted Advertising

Veronica Marotta¹, Kaifu Zhang², and Alessandro Acquisti¹

¹Heinz College, Carnegie Mellon University

²Tepper School of Business, Carnegie Mellon University

WORKING PAPER

Abstract

We analyze how alternative consumer data handling regimes affect the welfare of consumers, advertising firms, and an intermediary Ad exchange in the context of targeted advertising. We find that the collection and use of consumer data for targeting purposes affect consumer welfare through three distinct, and possibly countervailing, effects: match improvement, offer discrimination, and supply expansion. Furthermore, we find that the economic interests of the three agents can be misaligned, depending on the degree of heterogeneity in consumer preferences. Finally, we find that a strategic intermediary may choose to share with advertising firms only a subset of consumer data, maximizing its profits at their cost. In situations where the intermediary has an incentive to reveal the information that maximizes its payoff, overlooking the other agents' interests, regulation of data collection and sharing may increase consumers' welfare.

1 Introduction

The debate over the economic value of personal data and the economic impact of privacy regulation has been lively in recent years. In March 2017, the US Senate voted to allow Internet Service Providers to share consumers' web-browsing data and other personal information with advertisers without first asking for consumer consent. The decision was not without reaction and has left

privacy advocates unsatisfied.¹ This is not surprising: the conflicting and disparate perspectives of policy makers, consumer advocates, and the online advertising industry have characterized the discussion on the collection and use of consumer information since its inception. Advocates of privacy protection have traditionally argued that revealing too much personal information may put consumers at peril (FTC 2012). Aside from malicious parties engaging in spam or identity theft, even reputable firms may capitalize on detailed personal information in manners that do not always benefit their consumers. An early, well known example is Amazon’s price experiment in 2001. Shoppers realized that Amazon was charging different users different prices for the same DVD² (in other words, it was experimenting with price discrimination). A more recent example is Staples, in 2012, charging users different prices based on their geographical location.³ On the other hand, the advertising industry has claimed that overly stringent protection of personal information hurts Internet ad revenues and, through that, reduces the availability of free content and free services online (ITIF 2010). Furthermore, personalized targeting—the practice of tailoring advertising based on increasingly detailed information about consumers—may become much less relevant if less consumer data is available to advertisers, and, as a result, consumers may be actually worse off when their privacy is protected (ITIF 2010). This paper proposes a new perspective on this debate. We present a framework for analyzing the economic impact of data sharing and personal data regulation in the context of targeted advertising. Online targeted advertising is one of the most common applications of the market for personal data (Tucker 2012). The online advertising industry has experienced remarkable growth in the past few years, reaching about \$60 billion in revenue in 2015 (IAB 2015). We study the online advertising market using a three-party model that includes consumers, advertising firms, and an intermediary Ad exchange. Our model is based on a widespread online advertising framework called Real-Time Bidding (RTB). RTB is a form of programmatic advertising through which advertisers can buy online display advertisement spaces in real time through Ad exchanges. RTB is forecast to take over the market for programmatic advertising with a compound annual growth rate of 24.78% during the period 2016 - 2020 (Research

¹Electronic Frontier Foundation, “Congress Sign Bill to Roll Back Privacy Rules into Law”, April 2017, <https://www.eff.org/deeplinks/2017/04/trump-signs-bill-roll-back-privacy-rules-law>.

²Elizabeth Dwoskin, “Why You Can’t Trust You’re Getting the Best Deal Online,” The Wall Street Journal, October 2014, <http://www.wsj.com/articles/why-you-cant-trust-youre-getting-the-best-deal-online-1414036862>.

³See previous footnote.

and Markets 2016).⁴ Our model analyzes various scenarios that differ in the type and amount of consumer data available to the intermediary Ad exchange and to the advertising firms that target ads to consumers. We use the model to compare the welfare implications of alternative consumer data-handling regimes.

The contribution of our analysis is threefold. First, we provide a formal taxonomy of personal information. We show that not all personal data affect consumer welfare in the same fashion. Information about the degree of compatibility between consumer preferences (which we name their “horizontal information”) and products’ attributes has vastly different welfare implications than information about consumer purchasing power (which we name their “vertical information”).

Second, we identify three effects through which the collection and the sharing of the two categories of information affect consumer welfare: a *match improvement* effect, an *offer discrimination* effect, and a *supply expansion* effect. The match improvement effect refers to the fact that greater knowledge about consumers’ preferences helps advertising firms choose their bids in an ad auction. This leads to a better matching of consumers with their preferred products/advertising firms. The offer discrimination effect refers to the fact that having consumers’ willingness-to-pay information allows advertising firms to offer customized deals that extract more consumer surplus. Typically, but not necessarily, this takes the form of price discrimination. Third, revealing more consumer information typically expands supply, since firms can customize their bids in the ad auction stage and their prices once a consumer clicks on an ad. This eliminates missed trading opportunities from asymmetric information.

While the existence of these effects is generally known, this paper shows formally how each effect is generated from the collection and the sharing of a specific category of information and investigates the overall effect, whose direction is not known a priori. Indeed, while the matching improvement and the supply expansion effects have a positive effect on consumer welfare, the offer discrimination effect can have a negative effect. As such, the overall effect can be positive or negative depending on the relative strength of the different effects.

Third, through formal equilibrium analysis, we analyze and compare the effect of different consumer data-handling regimes in online targeted advertising, on the economic outcomes of the

⁴The compound annual growth rate is the mean annual growth rate of an investment over a specified period of time longer than one year.

market agents considered in our model. The regimes differ in terms of the amount and type of consumer data that the Ad exchange is able to collect and share with the advertising firms for the purpose of targeted advertising: no information, only horizontal information, only vertical information, or both. This, in turn, allows us to investigate whether, when, and how a consumer protection agency should intervene and regulate personal data collection and utilization.

Our findings have various implications for business and public policy. One implication arises from the observation that the matching improvement and supply expansion effects have a positive effect on consumer welfare, whereas the offer discrimination effect can have negative effects. Entailed in this observation is that both sides in the privacy debate (the advertising side and the privacy advocates side) have identified some, but not all, of the economic forces at work in the targeted advertising process. The advertising industry has often presented online, and specifically targeted, advertising as an economic win-win: it ensures that ad placements display content that consumers are interested in rather than ads that are irrelevant and uninteresting; advertisers achieve higher brand awareness and a greater chance of selling the product; and publishers also win as being able to offer behavioral targeting increases the value of the ad placements and therefore their revenues (Unanimis Consulting Limited 2011). Our results, however, point at the importance of a more comprehensive framework to understand the overall effects of data sharing and data protection. Importantly, the idea that different pieces of information may have different effects on consumer welfare is not new. Varian (2009) argued that “consumers will rationally want certain kinds of information about themselves to be available to producers and want other kinds of information to be secret.” In this paper, however, we formalize the distinction between different categories of consumer data, and then identify the specific economic mechanisms that are associated to the collection of those different categories of information. In so doing, we find that the effects of different pieces of information on the consumer are more nuanced than previously suggested. A common argument is that consumers would like sellers to know what product they want but not how much they are willing to pay for it. In the context of our framework we find that this is not always the case: consumers’ preferences about which type of information to reveal depend mostly on the heterogeneity in consumers’ preferences. In some cases, consumers may even prefer companies to know how much they are willing to pay for a given product.

Furthermore, our analysis reveals a nuanced set of insights regarding the potential conflict of

economic interests among Ad exchanges, advertising firms, and consumers. At times, the preferred information regimes will coincide: for some model parameters, the intermediary’s profit-maximizing strategy also maximizes consumer welfare. In other cases, however, the economic incentives of the different players can vastly differ. In fact, a strategic intermediary able to collect both vertical and horizontal information, but also able to decide how much of it to pass onward to advertising firms for targeting, may choose an information-sharing strategy that maximizes its profits at the cost of advertisers’ or/and consumers’ welfare. In situations where such an intermediary has an incentive to strategically and selectively reveal certain consumer data but not others, regulation of data exchanges may actually increase consumers’ welfare. The misalignment of economic incentives arises because the intermediary receives revenues from the sales of ad inventory and can drive up equilibrium payments by intensifying competition among advertisers. As such, under certain conditions, a strategic intermediary may prefer to withhold horizontal information, so that each advertiser has to compete with the other advertiser(s) on all consumers. If horizontal information is revealed, each advertiser will bid more when there is a greater fit between the customer and the firm’s product, but less otherwise. In other words, disclosing horizontal information may improve consumer welfare (through better matching) but drive down the intermediary’s income. On the other hand, revealing vertical information expands supply but also facilitates price discrimination. This may lead to either higher or lower consumer welfare, depending on the parameters. However, this increases the advertisers’ willingness to bid, and usually benefits the intermediary. As such, the intermediary has incentives to target consumers using the vertical information rather than the horizontal, while this is not necessarily the optimal setup for the consumers. This implies that regulatory privacy interventions should have different focuses depending on the nature of the market. Even though our model treats information and targeting as discrete, the insights have practical implications and extend to a more continuous interpretation of targeting by suggesting the existence of the incentive, on the intermediary side, to implement a broader targeting rather than a narrow one (Levin and Milgrom 2010).

For regulators, our taxonomy of personal information suggests that not all personal data should be treated the same. While it may be beneficial to enforce protection of some personal information, the market may be trusted to handle the rest. Our argument focuses on economic considerations, not ethical and security ones. However, understanding which personal data should be protected

is benefited by a deeper understanding of the three economic roles played by information: the match improvement effect, the supply expansion effect, and the offer discrimination effect. While our equilibrium results depend on model specificity, the analytical framework provides a more structured approach to understand the role of personal data in an advertisement market.

2 Related Work

Our work connects different strands of literature: the economic literature on consumers' privacy; the IS literature on online targeting and online auctions; and the economic and marketing literature on advertising.

The privacy literature is broad and diverse. Acquisti et al. (2016) offer a comprehensive review of the theoretical and empirical research on the economics of privacy, with a specific focus on the trade-offs that are associated with the sharing and protection of personal data. Under this broad umbrella of research, our paper specifically relates to theoretical works that analyze the impact of the use and exchange of consumer information on market outcomes and social welfare (Varian 2009, Taylor 2004, Calzolari and Pavan 2006, Casadesus-Masanell and Hervas-Drane 2015). These works show how the disclosure and use of consumer information can have different impact on social welfare depending on the context: in some cases, information exchanges are beneficial; in others, the use of information may be harmful and decrease consumers' surplus.

Second, our work relates to the IS and economics studies that analyze the incentives of data intermediaries and online platforms to provide accurate information or act strategically in the context of online advertising. Hagiu and Jullien (2011) analyze the incentives of an intermediary to divert consumers' search by directing them first to the least preferred store. In other words, an intermediary that has information about the best match between consumers and firms and receives a fee for its service may manipulate the quality of the match for profit-maximizing purposes. Eliaz and Spiegler (2011) consider the incentive of a monopolistic search engine to degrade the quality of the search pool by setting a low price per click that encourages low relevance firms to enter and leads to higher prices in the pool. More recently, authors in the IS and marketing field have offered a more in-depth analysis of the targeted advertising ecosystem by taking into consideration the fact that the targeting process is an intermediated process (Zhang and Katona 2012), and

by considering the important role played by publishers in the targeting process (Chen and Stallaert 2014). For example, Zhang and Katona (2012) analyze how the existence of an independent, profit-maximizing intermediary that sells advertising space and implements the target technology impacts the market’s outcomes and targeting accuracy. They suggest that, when product market competition is low, the intermediary offers accurate targeting; when product market competition is high, the intermediary offers inaccurate target technology that decreases the ability of the advertisers to create informational differentiation. Finally, it is important to mention the literature on online auctions and search-advertising (Bapna et al. 2003, Pinker et al. 2003, Liu et al. 2010, Chen and He 2011). Broadly, these works study the features that characterize online auctions and focus on the equilibrium properties of the generalized second-price auction, commonly used to place search-advertisements.

Third, our work is related to the traditional literature on the economics of advertising and targeting (Soberman 2004, Iyer et al. 2005, Esteban and Hernandez 2007). Works in this area look at the impact of advertising on product information and pricing, analyzing the impact of targeting (generally defined, and not necessarily related to *online* targeting activities) on firms’ strategies and market outcomes. Goldfarb (2014) presents a complete review of the economics literature on online advertising.

Our paper differs from and expands the existing literature in various ways. While general categorizations of information have already been proposed, we provide a formal taxonomy and combine it with the identification of the specific economic effects through which the collection and sharing of specific categories of consumer information may affect market outcomes and consumer welfare. Differently from most existing works, we develop a three-party model that analyzes the interaction among: i) a monopoly intermediary that runs auctions for advertisement allocation and implements the targeting; ii) competing advertising firms that want to target consumers and participate in online auctions in order to buy advertisements; and iii) consumers who see advertisements and make purchase decisions. Furthermore, we allow consumers to differ along two dimensions: a horizontal dimension, that captures the consumers’ product preferences; and a vertical dimension, that captures differences in consumers’ purchasing power. Typically, models focus on only one dimension at a time. Furthermore, we expand the different strands of literature by analyzing the incentives of an online intermediary to share specific types of consumer information that are used for targeting

advertising purposes, and the effect that the use of different types of information has on consumers’ surplus. Finally, our paper focuses on Real-Time Bidding, a technology introduced in 2006 that, despite its success in the industry, has not received a lot of attention in the academic literature. Thus, our work draws attention to a technology that is expected to take over the market for the allocation of display advertising in the next years.

3 Institutional Details

Real-Time Bidding (RTB) is a paradigm of serving ads that aims at bringing more liquidity to the online advertising market. Specifically, RTB allows advertisers to buy online display advertisement spaces in real time through Ad exchanges. When a user visits the website of a publisher belonging to an Ad exchange network, a request is sent to the Ad exchange which subsequently broadcasts a bid request to bidders (advertising firms). The requests contain various attributes that describe the impression and from which the advertising firms may infer what we have referred to above as horizontal (preferences) and vertical (purchasing power) information. The attributes can include (but are not limited to): user’s browsing behavior and geographical location, inferred or predicted interests and preferences, as well as the identity of the publisher (the site that the user is currently visiting and on which the ad will be displayed), features of the ad space (such as size, position), third-party cookies data, and so on (Interactive Advertising Bureau, 2012). Bidders (the firms that want to advertise) analyze the impression and submit their bids. Traditionally, Ad exchanges design the format of their advertisement auctions—designs that can be quite intricate and complex. We abstract from the technicalities of the auction design and focus on the dominant online auction format for display advertisements, which is a second-price auction. We borrow the classic insight that each firm bids according to its true valuation of the consumer and apply it to the equilibrium analysis. The winning party is allowed to serve the advertisement to the user and pays the second-highest bid.

The structure of the online advertising ecosystem is also a complicated issue in and of itself, with a universe of diverse agents and intermediaries. The model developed in this paper abstracts away certain intermediaries (such as demand side or supply side platforms) and focuses on three types of agents: i) advertising firms—firms that wish to target their products to specific consumers

by buying advertisements through online auctions; ii) Ad exchange—an intermediary platform that implements the RTB process and the targeting of the advertisements (for the rest of the paper, we will refer to this agent as the intermediary); and iii) consumers—visitors of the website where the ad is displayed and to whom advertising firms are trying to target their products. Thus, advertising firms bid for a given consumer, represented by the collection of information that is shared during the auction for the advertisement. As noted, we abstract from other market agents, such as demand or supply side platforms. Another important class of market agents, which do not make a strategic decision in our model, are content publishers (e.g., websites) that acquire users’ behavior data and earn advertising revenues by displaying ads on their pages. In some cases, such as Google’s YouTube, the content publisher also runs the Ad exchange and holds impression auctions.

Our model considers alternative informational regimes that differ in the type and amount of consumer information available in the market: no information, only vertical information, only horizontal information, and complete information. The amount of consumer data available for targeting—that is, the informational regime—can, in principle, be determined by the actions of different entities exogenous or endogenous to the model. Policy makers may regulate consumer data collection and allow, or prohibit, the sharing and usage of certain types of data. (Some) consumers may employ privacy-enhancing strategies (such as deleting cookies, opting out of online tracking, or using anonymous browsing) to hide browsing histories or protect other personal information.⁵ Finally, and perhaps most importantly, Ad exchanges may be in the position to decide how much consumer data to collect and how much of it to share with other agents, such as advertising firms. Recent works that have introduced the idea that Ad exchanges may have an incentive to hide or reveal certain type of information (Abraham et al. 2013, Sun et al. 2016) focus on the auction mechanism and its design. In this paper, we first identify the economic mechanisms behind the intermediary motive and we subsequently analyze the consequences on the targeting outcome and on the payoff for the other agents.

Accordingly, our model considers both the case in which the realization of a given informational

⁵Even so, consumers are not always technology savvy, and may not be aware of the different forms of tracking that occur online. In addition, historically the data industry has kept developing more powerful tracking technologies that resist ordinary remedies. Examples include so-called “Zombie Cookies”, a more powerful version of cookies that cannot be deleted or managed routinely like ordinary cookies (Pierson and Heyman 2011), as well as browser fingerprinting, which refers to the ability of service providers to use information collected from a remote device (for instance, detailed information about its browser settings) in order to identify, and track, its owner (Boda et al. 2011).

regime is exogenous to the decisions of the Ad exchange (that is, the intermediary cannot choose how much information to collect about consumers and it shares with the advertising firms all the information collected); as well as the case in which the choice of the informational regime is endogenous to the Ad exchange (that is, the intermediary is able to collect both horizontal and vertical consumer data and is able to strategically choose which to share with the advertising firms).

In the next section (Section 4), we introduce the model setup. In Section 5, we consider the exogenous case for alternative informational regimes and analyze how the economic outcomes of the three types of agents—defined in terms of expected payoffs and expected surpluses—change as a function of that amount and type of information. In Section 6 we consider the endogenous scenario and analyze how a strategic intermediary will decide which of the consumer information to share with the advertising firms during the auction process. Additionally, we investigate under what conditions the intermediary will choose an information sharing strategy that is sub-optimal for consumer welfare, and we discuss what type of regulatory intervention may be used in these cases to increase consumer welfare.

4 The Model

We consider two firms, $j = 1, 2$, that produce two different products at a constant marginal cost (assumed to be zero without loss of generality). The market consists of a unit mass of consumers. Each consumer has a demand for at most one unit of the product. We assume that consumers differ along two dimensions: horizontal and vertical.

The horizontal dimension describes a consumer’s subjective preference for either of the brands. In other words, each firm has a segment of consumers who have a preference for its product (everything else being equal, their willingness to pay is higher for that product). Each consumer can take one of two horizontal positions. We assume that a proportion α_1 of consumers prefer Firm 1 and a proportion α_2 prefer Firm 2, with $\alpha_1 + \alpha_2 = 1$ and $0 \leq \alpha_j \leq 1$ with $j = 1, 2$. To simplify the analysis, we assume that $\alpha_1 = \alpha_2$ throughout the main analysis without loss of generality.

The vertical dimension broadly captures consumer differences in purchasing power. A consumer can be a low valuation consumer or a high valuation consumer. For a given product (therefore, for a given preference), a low valuation consumer has a lower willingness to pay (i.e., he or she can

afford to pay less) than a high valuation consumer. We assume that a proportion β of consumers are high valuation consumers and a proportion $1 - \beta$ are low valuation consumers, with $0 \leq \beta \leq 1$. Thus, in total, four segments of consumers exist, differing in terms of both their brand preferences as well as their purchasing power.⁶

We define a consumer's valuation for a product as z_{itj} , where $i = 1, 2$ captures the consumer's preference for one of the products and $t = L, H$ captures whether the consumer is low or high valuation. A consumer's valuation is a function of the product under consideration, which we identify with $j = 1, 2$.⁷ Specifically, we set $z_{itj} = v_t$ for $i = j$ and $z_{itj} = w_t$ for $i \neq j$, where $v_t > w_t$. In other words, we indicate with v_t the reservation price of a consumer of type it for her favorite product and with w_t the reservation price of a consumer of type it for a product that is not her favorite. For instance, $z_{2t2} = v_t$ will refer to the valuation of a consumer of type $2t$ for product 2, with $t = L, H$; differently, $z_{2t1} = w_t$ will refer to the valuation of a consumer of type $2t$ for product 1, with $t = L, H$. Thus, the four segments of consumers value the products at v_H, w_H, v_L and w_L respectively, depending on both fit and purchasing power. Importantly, we use symmetric valuations to make the algebraic expressions tractable and interpretable; the model can be extended to asymmetric valuations.

From our initial assumptions, it follows that $v_h \geq v_l$ and $w_h \geq w_l$. We do not make any assumption on the relationship between v_l and w_h . In other words, the price a high valuation consumer is willing to pay for a product that is not her favorite (w_h) can be higher or lower than the price a low valuation consumer is willing to pay for her favorite product (v_l). This makes our model flexible, enabling us to analyze what happens when different consumers' preferences and market configurations are considered.

We assume that consumers have preferences between products; however, without advertising, they will not know which company sells which product, and at what price. Advertising thus plays the traditional informative role, as it makes consumers aware of existing firms and their prices Bagwell (2007). The advertising targeting process is facilitated by the Ad exchange. During each

⁶The difference between having a preference for a product and the actual amount a consumer can afford to pay is crucial in online settings. For instance, a user with a passion for cars may often browse websites about sports cars; therefore, one can infer that the consumer has a preference for sports cars. Nevertheless, that does not mean that the consumer can actually afford to pay for a sports car. As a consequence, the separation between horizontal and vertical information is fundamental.

⁷We will use the subscript j to refer interchangeably to both firms and products.

impression auction, the Ad exchange may share with the advertising firms information about the consumers. The advertising firms, after observing the set of information about the consumer, decide how much to bid for the impression and which price to advertise. The highest bidder wins and pays the second-highest bid. As such, the bidding price, and therefore the intermediary's revenues from the auction, is affected by the information that is available about the consumer on the market.

The information regimes we consider range from a benchmark case where no information about the consumer is available, to availability on the market of at least one dimension of the information (either the horizontal or the vertical information), to full consumer information availability. Formally, we define the information regime as $\varphi = I \times T \rightarrow 2^{I \times T}$ where $I = \{1, 2\}$ and $T = \{H, L\}$. In other words, each consumer type is mapped into a partition of all consumer types. For instance, consider the following information regime: $\varphi(i, H) = \varphi(i, L) = \{(i, H), (i, L)\}$. Under this scenario, it is possible to identify a consumer's horizontal type, but it is not possible to distinguish whether the consumer is low valuation or high valuation. Consider, instead, an information regime such that $\varphi(i, t) = \{(i, t)\}$. Under this scenario, it is possible to identify both the consumer's horizontal and vertical position. This corresponds to availability of consumers' full information.

It follows that we have four possible informational regimes: 1) No Information: $\varphi(i, t) = \{(1, H), (1, L), (2, H), (2, L)\}$. Neither consumer's horizontal nor vertical information is available. 2) Horizontal Information: $\varphi(i, t) = \{(I, H), (I, L)\}$, for $I = \{1, 2\}$. Only the consumer's horizontal information is available. 3) Vertical Information: $\varphi(i, t) = \{(1, t), (2, t)\}$, for $T = \{H, L\}$. Only the consumer's vertical information is available. 4) Complete Information: $\varphi(i, t) = \{(i, t)\}$, for $I = \{1, 2\}$ and $T = \{H, L\}$. All the information about the consumer—that is, both horizontal and vertical—is available.

As noted earlier, advertising firms decide how much to bid for the advertisement and the pricing strategy for their product on the basis of the information available about consumers. Stated differently, prices and bids are affected by the information regime.

Formally, we define $p_j(\varphi(i, t))$ as the price set by firm j and $b_j(\varphi(i, t))$ as the bid set by the firm j , both as function of $\varphi(i, t)$. In other words, an advertising firm can only customize its price and bidding to the extent of $\varphi(i, t)$. The expected payoff for advertising firm j is equal to:

$$\Pi_j = \sum_{i,t} I((b_j(\varphi(i, t)) > b_{-j}(\varphi(i, t))) * (\alpha[I(z_{itj} \geq p_j(\varphi(i, t))) * \beta * p_j(\varphi(i, t))] + I(z_{itj} \geq p_j(\varphi(i, t))) * (1 - \beta) * p_j(\varphi(i, t)))) - b_{-j}(\varphi(i, t)))$$

where $I((b_j(\varphi(i, t)) > b_{-j}(\varphi(i, t))))$ is an indicator function that it is equal to 1 if firm j wins the auction and zero otherwise. The subscript $-j$ refers to the second-highest bid. $I(z_{itj} \geq p_j(\varphi(i, t)))$ is an indicator function that is equal to 1 if the consumer buys the product (that is, the consumer's reservation price for product j is not lower than product j 's price).

Finally, the intermediary's expected payoff is equal to the amount paid by the firm for the advertisement, which in turn is equal to the second-highest bid:

$$\Pi_I = b_{-j}(\varphi(i, t)) * I((b_j(\varphi(i, t)) > b_{-j}(\varphi(i, t)))).$$

5 Analysis

In this section, we solve for and compare the equilibrium solutions of the four information regimes (no information, only horizontal information, only vertical information, and complete information). We assume that the intermediary cannot choose the information regime. In other words, the Ad exchange cannot choose what type of consumer information to collect and share with the advertising firms: more or less information may be available in the market for reasons that are exogenous to the intermediary (see Section 3; for instance, regulation may prohibit the collection of certain types of consumer data, or a consumer may hide her browsing behavior through privacy enhancing technologies). The intermediary collects whatever information is available on the market about consumers and shares it with the advertising firms. (In Section 6 we consider a strategic intermediary that can collect both the horizontal and the vertical information but can decide which to share with the advertising firms.)

5.1 Sequence of Events

1. A consumer arrives at a website.
2. The intermediary receives a signal that the specific consumer is on a given website and forwards the signal to advertisers that may be interested in showing an advertisement to that consumer. The signal includes all the information about the consumer that is possible to collect on the market. As noted, different types of information about the consumer may be available: no information, only the horizontal information, only the vertical information, or

both.

3. A second-price auction is run to allocate the advertisement space. On the basis of the information available, advertising firms decide simultaneously how much to bid for the advertisement and set the price of the product.
4. The advertising firm that submits the highest bid wins the auction, pays the second-highest bid and shows the ad to the consumer.
5. The consumer sees the ad and decides whether or not to buy the product. The consumer buys as long as the price is lower than his or her reservation price.

5.2 Equilibrium Concept

In the game we described above, it is important to specify that while the two firms simultaneously set the bid for the advertisement and the price for the product, each firm's bidding strategy is sequential to its own pricing strategy: given its pricing strategy, each firm sets its bidding strategy.

To find the equilibrium for the advertisement auction and therefore the winning bid, we rely on existing results in auction theory for equilibria in second-price auctions (Vickrey 1961). Since the object being auctioned is a single, indivisible advertisement space, it is a weakly dominant strategy for firms to bid their true valuation. In our model, this corresponds to advertising firms bidding according to their true valuation of the consumer, that is how much each firm expects to gain from that particular consumer if it wins the auction and it obtains the right to show the advertisement for its product. Conditions for the uniqueness of the equilibrium can be maintained by introducing the infinitesimal probability that a random company decides to participate to the auction.

In deriving each firm's bidding and pricing strategy, we solve for the Nash Sub-game Perfect Equilibrium (Selten 1965, Kreps and Wilson 1982). With a finite horizon, such equilibrium is computed by backward induction. In our model, this requires that we first derive each firm's pricing strategy and then derive, backwards, the bidding strategy.

5.3 Equilibrium Solutions

In this section we analyze the equilibrium solutions for the four information scenarios described above. From a notational point of view, we will use p_{itj}^Ω and b_{itj}^Ω to refer to the price and bid set

by firm j under the informational scenario $\Omega \in \{N, Hor, Ver, Com\}$ where:

$$\{N, Hor, Ver, Com\} = \{No\ Information, Horizontal, Vertical, Complete\}$$

We start considering the benchmark case where no information about consumers is available. In this scenario, advertising firms have only the prior information on consumer distribution. Therefore, their bidding and pricing strategies factor in the uncertainty about the consumer's willingness to pay. In other words, if an advertiser obtains an impression, the consumer it who sees the firm's ad can be of any type with probabilities according to the prior. Targeted pricing is not possible. Each firm can set a higher price such that only the higher valuation consumers will buy, or a lower price and sell to more consumers. Given the distribution of consumer types, there are four possible price points v_h, v_l, w_h, w_l . Given the pricing strategy, each firm's bid is equivalent to the expected profit from showing the ad to the consumer.

The pricing and bidding strategies depend on the relative magnitude of v_h, v_l, w_h, w_l as well as β . In general, firms are more likely to sell to more consumers (that is, more consumers are likely to buy) when heterogeneity is small and β is small. For the rest of the analysis we assume $\alpha = 1/2$ for simplicity and without loss of generality. Since advertising firms behave symmetrically, the advertisers always bid the same amounts and matching is random. Lemma 1 summarizes the equilibrium pricing and bidding strategies. Extensive proofs are available in the appendix.

Lemma 1. No Information:

- Firm j 's bidding strategy is $b_{itj}^N = \max\{w_l, \frac{1}{2}v_l, \frac{1}{2}\beta v_h, \beta w_h + (1 - \beta)\frac{1}{2}w_h\}$ when $v_l \geq w_h$ and $b_{itj}^N = \max\{w_l, \beta v_l + (1 - \beta)\frac{1}{2}v_l, \frac{1}{2}\beta v_h, \beta w_h\}$ when $v_l < w_h$.
- Firm j 's pricing strategy is:
 - $p_{itj}^N = w_l$ when $b_{itj}^N = w_l$;
 - $p_{itj}^N = v_l$ when $b_{itj}^N = \frac{1}{2}v_l$ or $b_{itj}^N = \beta v_l + (1 - \beta)\frac{1}{2}v_l$;
 - $p_{itj}^N = w_h$ when $b_{itj}^N = \beta w_h + (1 - \beta)\frac{1}{2}w_h$ or $b_{itj}^N = \beta w_h$;
 - $p_{itj}^N = v_h$ when $b_{itj}^N = \frac{1}{2}\beta v_h$.

- Firm j 's expected payoff is $\Pi_j^N = p_{it}^N - b_{it-j}^N$.
- The intermediary's expected payoff is $\Pi_I^N = \min(b_{itj}^N, b_{it-j}^N)$.

The no information regime represents a form of untargeted, random advertising technology: firms cannot target specific consumers. As a consequence, firms act in expectation but, by so doing, bid identically. The intermediary thus extracts all the surplus.

Next, we consider an information regime where only the consumer's horizontal information is available to advertising firms during the auction. Let us assume that, based on the information received, firms observe a consumer of type 1. Firm 1 can decide to adopt one of two strategies: it can choose to capture any consumer coming from its segment, by setting $p_1 = v_l$. In this case, the bid would be $b_1 = v_l$. Alternatively, if β is large enough, it can choose to capture, inside its segment, only the high valuation consumers by setting $p_1 = v_h$. In this case, the bidding strategy would be $b_1 = \beta v_h$. In summary, Firm 1's bidding strategy is $b_1 = \max(\beta v_h, v_l)$.

Note that Firm 2 also submits a positive bid when a type 1 consumer is observed, when the firms bid their true valuations.⁸ Indeed, since the consumer also has a positive valuation for the product, he or she will follow a bidding strategy similar to that of firm 1, by bidding $b_2 = \max(\beta w_h, w_l)$. Those results are summarized in Lemma 2. The proof is contained in Appendix 1.

Lemma 2. Horizontal Information:

- Firm j 's bidding strategy for consumer i is: $b_{itj}^{Hor} = \max\{\beta v_h, v_l\}$ for $j = i$, $i, j = 1, 2$ and $t = l, h$; $b_{itj}^{Hor} = \max\{\beta w_h, w_l\}$ for $j \neq i$, $i, j = 1, 2$ and $t = l, h$.
- Firm j 's pricing strategy for consumer i is: $p_{itj}^{Hor} = v_h$ when $\beta v_h \geq v_l$ and $p_{itj}^{Hor} = v_l$ otherwise, for $j = i$, $i, j = 1, 2$ and $t = l, h$; $p_{itj}^{Hor} = w_h$ when $\beta w_h \geq w_l$ and $p_{itj}^{Hor} = w_l$ otherwise, for $j \neq i$, $i, j = 1, 2$ and $t = l, h$.
- Firm j 's expected payoff is $\Pi_j^{Hor} = [\beta v_h I(\beta v_h \geq v_l) + v_l I(\beta v_h < v_l) - b_{it-j}]$, for $i = j$, $i, j = 1, 2$, $t = l, h$.

⁸Once again, the assumption that firms always bid their true valuations is both classic and non-restrictive. It may be tempting to assume that a firm will refrain from bidding altogether if it knows for certain that it cannot win an impression. This, however, is not the rational decision if there is a tiny bit of uncertainty regarding the behavior of other bidders. Bidding one's true valuation is always the rational strategy.

- The intermediary's expected payoff is $\Pi_I^{Hor} = \max(\beta w_h, w_l)$ for $i = 1, 2$.

When only the horizontal information is available, even though both companies submit a positive bid, firm j that observes a consumer of type $i = j$ always wins the auction. Consequently, consumers are shown the advertisement for their favorite product. Since the other firm also submits a positive bid, the intermediary's expected payoff is not zero.

The next scenario we consider is the regime in which only the vertical information is available to advertising firms during the auction: firms can distinguish between high valuation and low valuation consumers, but they do not know the consumer's horizontal preferences. Let us assume that the consumer is high valuation. Firms know that the consumer will be willing to pay v_h for his favorite product and w_h for the other. They also know the respective probability of getting a consumer from a specific segment, that is α_j . Hence, they can decide to adopt one of two strategies: i) firm j can decide to capture any consumer that is high valuation by setting a price equal to w_h ; or ii) it can decide to capture any consumer coming from its segment and that is high valuation, by setting a price equal to v_h . The same reasoning can be applied to the case when a consumer is observed to be of the low type. Importantly, in this case, it is not immediately clear which company is going to win the auction: the final outcome depends on the values of the parameters. For example, when $\alpha_j = 0.5$ both submit the same bid and the consumer is randomly assigned. The same reasoning extends to a low valuation consumer.

Lemma 3 summarizes the result. The proof is provided in the Appendix.

Lemma 3. Vertical Information:

- Firm j 's bidding strategy is $b_{itj}^{Ver} = \max\{\frac{1}{2}v_h, w_h\}$, for $t = h$ and $i, j = 1, 2$; $b_{itj}^{Ver} = \max\{\frac{1}{2}v_l, w_l\}$, for $t = l$ and $i, j = 1, 2$.
- Firm j 's pricing strategy is $p_{itj}^{Ver} = v_h$ when $\frac{1}{2}v_h \geq w_h$ and $p_{itj}^{Ver} = w_h$ otherwise, for $t = h$, $i = 1, 2$ and $j = 1, 2$; $p_{itj}^{Ver} = v_l$ when $\frac{1}{2}v_l \geq w_l$ and $p_{itj}^{Ver} = w_l$ otherwise, for $t = l$, $i = 1, 2$ and $j = 1, 2$.
- Firm j 's expected payoff is $\Pi_j^{Ver} = \beta[\frac{1}{2}v_h I(\frac{1}{2}v_h \geq w_h) + w_h I(\frac{1}{2}v_h < w_h) - b_{ih-j}] + (1 - \beta)[\frac{1}{2}v_l I(\frac{1}{2}v_l \geq w_l) + w_l I(\frac{1}{2}v_l < w_l) - b_{il-j}]$, for $i = 1, 2$, $j = 1, 2$.

- *Intermediary's expected payoff is*

$$\Pi_I^{Ver} = \beta[\min\{\max\{\frac{1}{2}v_h, w_h\}, \max\{\frac{1}{2}v_h, w_h\}\}] + (1 - \beta)[\min\{\max\{\frac{1}{2}v_l, w_l\}, \max\{\frac{1}{2}v_l, w_l\}\}],$$

for $i = 1, 2, j = 1, 2$.

Finally, in the last informational regime that we analyze, both horizontal and vertical information about the consumer are available. In this case, both firms are able to perfectly personalize bid and price at the individual level. The equilibrium strategies are described in Lemma 4.

Lemma 4. Complete Information:

- *Firm j 's bidding strategy is $b_{itj}^{Com} = v_h$ for $j = i$ and $t = H$; $b_{itj}^{Com} = w_h$ for $j \neq i$ and $t = H$; $b_{itj}^{Com} = v_l$ for $j = i$ and $t = L$; $b_{itj}^{Com} = w_l$ for $j \neq i$ and $t = L$.*
- *Firm j 's pricing strategy is $p_{itj}^{Com} = v_h$ for $j = i$ and $t = H$; $p_{itj}^{Com} = w_h$ for $j \neq i$ and $t = H$; $p_{itj}^{Com} = v_l$ for $j = i$ and $t = L$; $p_{itj}^{Com} = w_l$ for $j \neq i$ and $t = L$.*
- *Firm j 's expected payoff is $\Pi_j^{Com} = \beta(v_h - w_h) + (1 - \beta)(v_l - w_l)$.*
- *The intermediary's expected payoff is $\Pi_I^{Com} = \beta w_h + (1 - \beta)w_l$.*

In this last scenario, firms are able to observe both a consumer's product preference and his purchasing power. Consequently, if firm 1 observes, for instance, a high valuation consumer of type 1, it can set the product's price to exactly equal the consumer's reservation price, v_h , and it also submits a bid equal to v_h (the firms expects that the consumer will buy its product at that price). Similarly, firm 2 knows that the consumer is only willing to pay w_h for its product; hence, it sets the price of the product and the bid for the advertisement accordingly. Since by assumption we have that $v_h \geq w_h$, firm 1 that bids v_h always wins the auction for the consumer. The same reasoning applies to a consumer that is low valuation.

5.4 Welfare Analysis

The previous section analyzed how the outcome of the advertisement's auction, in terms of advertisers' expected payoff and intermediary's payoff, changes under the four different information regimes. As a reminder, the four information regimes are: i) a benchmark case where no information

about consumers can be collected and used; ii) a scenario where only the horizontal information about consumers can be collected and used; iii) a scenario where only the vertical information about consumers can be collected and used; and iv) a complete information scenario where both types of information can be collected and used.

In this section we compare the four scenarios and we analyze which information regime maximizes welfare from the perspectives of the consumers, the advertisers, and the intermediary. This allows us to answer critical questions regarding the potentially differing interests of the parties.

The reader remembers that we use $\Omega \in \{N, Hor, Ver, Com\}$ to refer to the different information scenarios. Let us further define $z_{it}^\Omega = \sum_j z_{itj} I(b_{itj}^\Omega > b_{it-j}^\Omega)$ as consumer i 's valuation of the product whose ad she actually sees in equilibrium. Similarly, let us define $p_{it}^\Omega = \sum_j p_{itj} I(b_{itj}^\Omega > b_{it-j}^\Omega)$ as the price which consumer i actually pays in equilibrium. Taken together, the consumer surplus function for a consumer of type it and informational scenario Ω is therefore:

$$W_{it}^\Omega = z_{it}^\Omega - p_{it}^\Omega$$

We use a traditional definition of consumer surplus that does not capture elements such as possible annoyance from advertisements or an explicit taste for privacy. Stated differently, we use a “conservative” definition of consumer surplus.

In the following propositions we analyze how the availability of different pieces of information about consumers affects consumer surplus, and through which mechanisms.

PROPOSITION 1. For given values of the parameters, revealing the Horizontal Information has the following effects:

- *Horizontal Information improves matching.* $\forall i, t, z_{it}^{Hor} \geq z_{it}^N$ given that $z_{it}^N > 0$. In expectation, each consumer will be matched to a (weakly) better product.
- *Horizontal Information expands supply.* $S^{Hor} = \{i, t | z_{it}^{Hor} - p_{it}^{Hor} \geq 0\}$ and $S^N = \{i, t | z_{it}^N - p_{it}^N \geq 0\}$ with $S^N \subseteq S^{Hor}$. (Weakly) more consumers will buy a product.
- *Horizontal Information allows personalized pricing.* Given the matching with any product, $p_{it}^{Hor} \geq p_{it}^N$.

The first expression captures what we call the *Matching Improvement Effect*. When the horizontal information is revealed, the consumer's surplus increases because he will be matched with his favorite product. The second expression captures the *Supply Expansion Effect*. More consumers will be exposed to a product they like, therefore increasing the proportion of consumers that will buy a product. Finally, the last point illustrates the *Price Discrimination Effect*. When the horizontal information is revealed, and therefore product preferences are revealed, the firm is not only able to target the advertisement and the product, it is also able to target the price.

PROPOSITION 2. Revealing the Vertical Information has the following effects:

- *Vertical Information does not improve matching.* $\forall i, z_{it}^{Ver} = z_{it}^N$. In expectation, consumers are not matched to a better product.
- *Vertical Information expands supply.* For $S^N = \{i | z_{it}^N - p_{it}^N \geq 0\}$ and $S^{Ver} = \{i | z_{it}^{Ver} - p_{it}^{Ver} \geq 0\}$: $S^N \subseteq S^{Ver}$ when $w_h/v_h \geq \alpha$. (Weakly) more consumers will buy a product.
- *Vertical Information allows personalized pricing.* $p_{it}^{Ver} \geq p_{it}^N$. Price is higher when the Vertical Information is revealed.

When the Vertical Information is revealed, there is no Matching Improvement. As in the No Information case, the consumer may be either matched with his favorite product or with a product he does not like. In addition, when the Vertical Information is revealed, and consequently purchasing power is revealed, the firm is able to distinguish between high valuation and low valuation consumers and, therefore, is able to target a different price for each group. Finally, when the market is characterized by a low degree of horizontal differentiation, revealing the vertical information expands the supply as more consumers are served.

The different effects that the different types of information have on market outcomes lead to a correlation between the underlying consumers' market structure and outcome in terms of consumer surplus. In other words, which information leads to an increase in consumer surplus depends on the market configuration, captured by the parameters that characterize the market. Proposition 3, below, identifies the conditions under which different informational regimes maximize consumer surplus. Figure 1 offers a visual representation of the conditions.

PROPOSITION 3. For consumers, the optimal information regime is as follows:

- When $\beta v_h \geq v_l$, $\alpha v_h \geq w_h$ and $w_h/v_h \leq \frac{\alpha\beta}{\alpha+\beta-\alpha\beta}$ or $v_l/w_h \leq \frac{\alpha\beta}{\alpha+\beta-\alpha\beta}$, consumers are indifferent between all four information policies.
- When $\beta v_h \geq v_l$, $\alpha v_h \leq w_h$ and $w_h/v_l > \frac{\alpha+\beta-\alpha\beta}{\beta}$, revealing the Vertical Information is optimal for the consumers.
- When $\beta v_h \leq v_l$, $\alpha v_h \leq w_h$, $\beta \geq (1 - \alpha)$ and $w_l/w_h \geq \alpha + \beta - \alpha\beta$ or $w_l/v_l \geq \alpha + \beta - \alpha\beta$, revealing No Information is optimal for the consumers.
- When $\beta v_h \leq v_l$, $\alpha v_h \geq w_h$ and $w_h/v_l \leq \frac{\alpha}{\alpha+\beta-\alpha\beta}$, revealing the Horizontal Information is optimal for the consumers.
- Otherwise, there is no clear dominance.

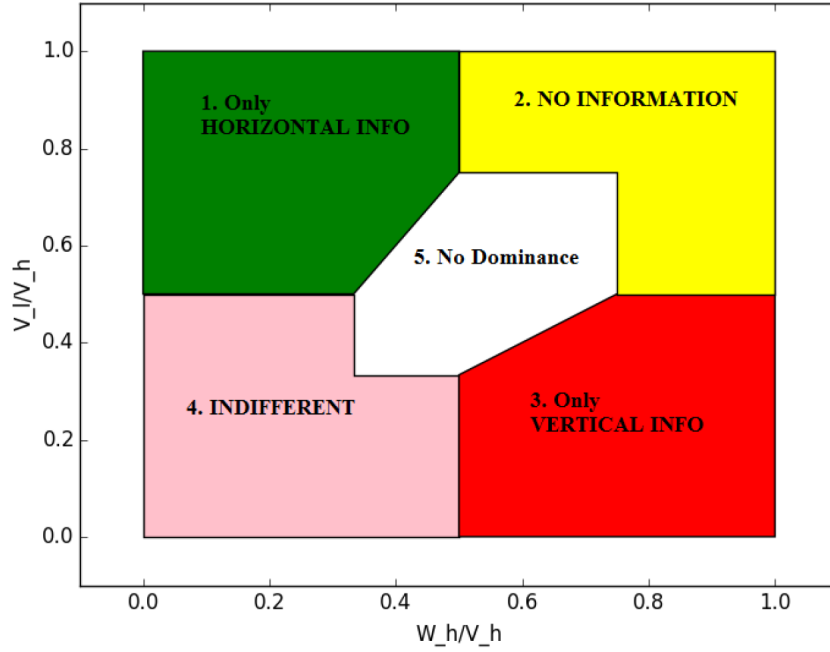


Figure 1: Preferred Information Scenarios for Consumers

Proposition 3 delineates the conditions under which different informational regimes maximize consumer surplus, as illustrated in Figure 1. The horizontal axis of Figure 1 captures the degree of consumer horizontal differentiation, represented by the ratio w_h/v_h . When the ratio is close to zero,

consumers have extreme product preferences; in other words, the reservation price for a consumer's favorite product tends to be much higher than the reservation price for the other product. On the other hand, when the ratio tends to one, consumers tend to have less strict preferences, implying that the reservation prices for the two products are quite similar. The vertical axis captures the degree of consumers' vertical differentiation, represented by the ratio v_l/v_h . Again, when the ratio is close to zero, the consumers' market is characterized by a high degree of vertical differentiation, meaning that the reservation price for a given product of a high valuation consumer is much higher than the reservation price of a low valuation consumer. When the ratio tends to one, the difference between high and low valuation consumers tends to shrink and reservation prices are similar.

Figure 1 identifies five different regions that correspond to different combinations of parameters. Note that the figure assumes $\alpha_j = 0.5$ and $\beta = 0.5$, but the results can be generalized (we explain below how the figure changes if we change the values).

Region 1, in green, is characterized by a low degree of vertical differentiation and a high degree of horizontal differentiation. While high valuation consumers are not very different from low valuation consumers, brand preferences are very well defined and almost extreme, making matching very important. Therefore, in this region, consumer surplus is maximized when the horizontal information is revealed. Indeed, as specified in Proposition 1, horizontal information improves product matching and it will ensure that consumers are targeted with advertisements for the products they like the most. This dominates the personalized pricing effect, which may reduce consumer welfare.

Region 2, in yellow, is characterized by a low degree of horizontal and vertical information; in other words, consumers' reservation prices and preferences are very similar. Thus, revealing any type of information would simply lead to personalized pricing without substantially improving matching or expanding demand. Therefore, consumer surplus is maximized when no additional information is collected and shared and consumers are not targeted.

Region 3, in red, is characterized by a high degree of vertical differentiation and a low degree of horizontal differentiation. In this case, revealing the vertical information is beneficial for consumers because this ensures that low valuation consumers are served. In other words, revealing the vertical information in that region leads to a demand expansion, as explained in Proposition 2. Although revealing the vertical information also leads to personalized pricing, which decreases consumer welfare, the demand expansion effect dominates. In this region, since horizontal differentiation is

small, the matching improvement effect is not as important, and consumers become worse off due to the price discrimination effect.

Region 4, in pink, identifies the combination of parameters where consumers are indifferent about whether and what information is being collected and shared. The region is characterized by a high degree of horizontal and vertical differentiation, and all the different informational regimes would lead to high prices and expected consumer surplus of zero. Although revealing more information will expand the demand, firms always manage to charge each consumer segment their expected valuation.

Finally, Region 5, in white, delineates the combination of parameters under which there is no obvious informational regime that would make every consumer better off. It is important to note that we are being conservative in the analysis of this region; indeed, while it is possible to identify, for each point in the region, which informational regimes maximize consumer surplus, such results likely depend on the specific functional form which we consider. Furthermore, it is not possible to find a given informational regime that makes every consumer better off.

As mentioned at the beginning, Figure 1 assumes that $\beta = 0.5$. A lower value for the parameter would imply a lower proportion of high valuation consumers in the market. This would imply an expansion of Region 1 and Region 2 and a shrinking of Regions 3, 4 and 5. Differently, a higher value for the parameter would imply a higher proportion of high valuation consumers in the market and would correspond to an expansion of Regions 3 and 4 (and a consequent reduction in Regions 1 and 2).

We next look at the advertising firms' preference, summarized in Proposition 4 below.

PROPOSITION 4. (Advertiser's Preference). In Equilibrium, the advertiser's expected payoff is greater when there is complete information about the consumer.

The result contained in Proposition 4 suggests that firms' expected payoffs are highest when complete information about consumers is available. Intuitively, when firms can perfectly target consumers, any uncertainty is removed. When advertising firms know not only the consumer preference but also how much a consumer is willing to pay for a given product, they can make a perfectly informed decision about which consumers to target and how to price the product. The

horizontal differentiation structure ensures that each firm is left with a positive surplus after paying the advertising fee.

6 Strategic Intermediary

In this section, we consider the case of an intermediary that can access both horizontal and vertical information but can also strategically choose which information to share with advertising firms during the auction process. We investigate which information scenario maximizes the intermediary's expected payoff and, therefore, which choice a strategic intermediary would make in equilibrium. This allows us to answer a series of interesting questions such as: When would the intermediary choose an information regime that also maximizes consumer welfare? When would a policy maker interested in maximizing consumer welfare want to intervene and correct the information regime preferred by the intermediary?

The sequence of events for this variant of the model is similar to what is described in Section 4.2. Step 2 changes: the intermediary collects all the information about the consumer and strategically decide which information to disclose to the advertising firms during the auction process. While the results discussed in the previous Lemmas and Propositions do not change, the focus is now on the intermediary's choice—that is, what is the information regime preferred by the Ad exchange if it has control over what information is shared with the advertising firms.

PROPOSITION 5. (Intermediary Choice). In Equilibrium, an Intermediary that can strategically decide which information to disclose will always have the incentive to disclose the Vertical Information.

Proposition 5 reveals an intriguing finding. While consumers may prefer different information regimes under different parameter values, as shown in the previous section, the intermediary always prefers revealing the vertical information. The intuition for this result relies crucially on the interaction between personal data and the ad auction. Revealing horizontal information, as we have discussed above, improves the matching between the consumers and the products. While this improves consumer welfare and increases willingness to pay on the part of the advertisers, it reduces

the competition in the auction for the advertisement. This is because each advertising firm would now bid higher for the more relevant segment of users and bid lower for the less relevant segment. This *lowers* the intermediary's payoff due to the structure of the second-price auction.

On the other hand, revealing vertical information facilitates price discrimination and increases the advertisers' willingness to bid. In contrast to the case of horizontal information, the vertical information simultaneously increases the willingness of both advertisers to bid for high valuation consumers. This intensifies competition in the second-price auction and increases the intermediary income from the high valuation segment. It is worth noting that, in most cases, revealing the vertical information also increases the intermediary's income from the low valuation consumers, who would otherwise be priced out of the market and worth little to the advertisers. These effects, when put together, lead to an unambiguously positive effect on the intermediary's overall revenue.

The results reported in Propositions 3, 4 and 5 highlight the conflict of interests among the different agents. Firms will be better off if they are able to perfectly target consumers, but the results of our model suggest that the intermediary has an incentive to allow this targeting imperfectly, by using only the vertical information, rather than both. Consumers may prefer some targeting, but never perfect targeting, depending on the parameters.

Compared with the advertising firms, consumers always prefer *less* information to be disclosed. Interestingly, there exist situations where consumers' interest is aligned with intermediary's interest: the results in Proposition 2 suggest that there are situations where consumer surplus is highest when the vertical information is observed. When this is the case, the intermediary can be trusted to choose the socially optimal outcome—while being entirely motivated by profit concern. In other cases, however, the intermediary's choice may not coincide with the information paradigm that is optimal for consumers. Interestingly, this may either mean that the intermediary reveals *too much* information, or that the intermediary reveals the *wrong type* of information. Corollary 1 identifies and formalizes the different situations by looking at the loss in consumer surplus that is caused by the intermediary's equilibrium choice.

COROLLARY 1. Let us define the following consumer surplus loss function:

$$L_i(r) = W_{it}^*(r) - W_{it}(r)$$

where $W_i^*(r)$ is the maximum surplus the consumers can obtain in region r , for $r = 1, 2, 3, 4$; and $W_i(r)$ is the surplus the consumers obtain in region r , given the Intermediary's equilibrium choice. Stated differently, the function captures the loss in consumer surplus due to the fact that in equilibrium the Intermediary may choose to reveal information different from the information that would maximize consumer surplus. We have the following situations:

- $L_i(1) = W_{it}^*(1) - W_{it}(3) = (1 - \alpha_j)\beta(v_h - v_l)$. *The intermediary reveals the wrong type of information.*

Consumers suffer a positive surplus loss because the informational regime chosen by the Intermediary diverges from the regime that maximizes consumer surplus. Specifically, in that region consumers would be better off if the Horizontal Information was revealed. Indeed, revealing the Horizontal Information would ensure a better product matching and a lower price.

- $L_i(2) = W_{it}^*(2) - W_{it}(3) = w_h - w_l$. *The Intermediary reveals too much information.*

Consumers suffer a positive surplus loss because the informational regime chosen by the Intermediary diverges from the regime that maximizes consumer surplus. Specifically, in that region, consumers would be better off if no information was revealed because more consumers would enjoy a lower price.

- $L_i(3) = W_{it}^*(3) - W_{it}(3) = 0$. *The information regime chosen by the Intermediary maximizes consumer surplus.*

Consumers do not suffer any surplus loss because the informational regime chosen by the Intermediary coincides with the regime that maximizes consumer surplus.

- $L_i(4) = W_{it}^*(4) - W_{it}(3) = 0$. *Consumers are Indifferent.*

Consumers do not suffer any surplus loss because they are indifferent with respect to which informational regime is chosen by the Intermediary.

Corollary 1 provides a set of guidelines to policy makers whose goal is to increase consumer welfare by intervening on the information regime. First, the results of the model suggest that there are situations where consumers are indifferent with respect to which information is available and used in the market. This happens in markets where consumers are highly differentiated, both in terms

of product preference and purchase power. In such situations, a policy maker may not need to intervene as all of the different informational regimes will lead to the same results for consumers. Second, in markets where consumers' degree of differentiation (both horizontal and vertical) is low, policy intervention may be needed, as consumers have a strong preference for their information being protected. Indeed, in situations where consumers are more homogeneous, revealing any type of information would lead to an unwanted product and price personalization that is not going to benefit consumers. These are the markets where consumers may benefit from a more stringent regulation of data collection and usage. Third, there are situations where consumers only want a certain category of information to be collected and used. This happens: i) when consumers are highly vertically differentiated but the degree of horizontal differentiation is low; and ii) when consumers are highly horizontally differentiated but the degree of vertical differentiation is low. In the first situation, consumers would like advertisers to know their vertical position (that is, their purchase power) because the match improvement effect is not as important. In the second situation, consumers would like advertisers to know only their horizontal position but not the vertical one so to ensure a better product matching and a low price.

These results are significant as they correct and expand a conventional wisdom of privacy economics whereby consumers would rationally want firms to know which product they like (so as to get relevant offers) but not *how much* they are willing to pay for it (so as to avoid perfect price discrimination: Varian (2009)). According to our results, there exist situations where the reverse happens: consumers may want to share their purchase power, but not their brand preferences. These findings suggest that a policy intervention may be needed in markets that display such features, but the degree of intervention needs to be adequate and targeted to specific categories of information. A too strict regulation or a regulation that targets the wrong type of data may end up hurting rather than helping consumers.

7 The Value of Information

The analysis presented in the previous sections assumed that consumer information is freely available in the market. In reality, the collection and acquisition of information is costly to the intermediary and to advertising firms themselves. The findings of our model suggest that different types

of information will be valued differently depending on the underlying market structure. Therefore, the amount that players are willing to invest in information acquisition varies accordingly.

Let us consider the intermediary. The collection or acquisition of different types of information (which in our model have been classified in two main categories, horizontal and vertical) affects the degree of precision with which consumers can be targeted and, as we have seen in the analysis, affects the incentives and the willingness to pay of the advertisers. In this section, we analyze in more detail what is the *additional value* generated by a specific type of information and, therefore, what is the corresponding amount that a data intermediary would have incentive to invest in information collection. In other words, if information is too costly, there may be situations where an intermediary may not have incentive to invest in the acquisition of additional information and, consequently, may not have incentive to improve the accuracy of targeting.⁹ Proposition 6 formalizes the results.

PROPOSITION 6. The Value of the Information.

- When $\beta v_h \geq v_l$, $\alpha v_h \geq w_h$ and $w_h/v_h \leq \frac{\alpha\beta}{\alpha+\beta-\alpha\beta}$ or $v_l/w_h \leq \frac{\alpha\beta}{\alpha+\beta-\alpha\beta}$, the Intermediary has incentive to invest at most $(1-\beta)\alpha v_l$ in the acquisition of the Vertical Information. If the Horizontal Information is already collected, the Intermediary is willing to invest at most $(1-\beta)w_l$.
- When $\beta v_h \geq v_l$, $\alpha v_h \leq w_h$ and $w_h/v_l > \frac{\alpha+\beta-\alpha\beta}{\beta}$, the Intermediary has incentive to invest at most $(1-\beta)w_l$ in the acquisition of the Vertical Information, regardless of whether the Horizontal Information is already collected.
- When $\beta v_h \leq v_l$, $\alpha v_h \leq w_h$, $\beta \geq (1-\alpha)$ and $w_l/w_h \geq \alpha + \beta - \alpha\beta$ or $w_l/v_l \geq \alpha + \beta - \alpha\beta$, the Intermediary should invest at most $\beta(w_h - w_l)$ in the acquisition of the Vertical Information, regardless of whether the Horizontal Information is already collected.
- When $\beta v_h \leq v_l$, $\alpha v_h \geq w_h$ and $w_h/v_l \leq \frac{\alpha}{\alpha+\beta-\alpha\beta}$, the Intermediary has incentive to invest at most $\beta(\alpha v_h - w_h) + (1-\beta)\alpha(v_l - w_h)$ in the acquisition of the Vertical Information.

⁹Here we are implicitly assuming that the intermediary has no incentive for collecting information that he is not planning on using during the targeting process. This makes sense specifically in cases where collecting (and then storing) information is costly.

If the Horizontal Information is collected, the Intermediary has incentive to invest at most $\beta(w_h - w_l)$ in the acquisition of the Vertical Information.

As the reader recalls, Proposition 4 in Section 5 highlighted that a profit-maximizing intermediary that has power over which information is being shared with the advertising firms during the advertisement auction will have the incentive to reveal the vertical information. Intuitively, Proposition 6 then suggests that the intermediary will have a stronger incentive to invest in the acquisition of the vertical information, regardless of whether it begins with no information or with horizontal information. Differently, the intermediary will have a weaker incentive to invest in horizontal information, in either case. While this result may seem puzzling, it in fact has important empirical validations and it shows support for the argument that increasingly precise targeting (derived from collection of increasingly precise information) tends to create thin markets where, possibly, only a single advertising firm has high willingness to pay (Celis et al. 2011). These situations put into question the use of the auction mechanisms that are commonly utilized for advertisement allocation because they reduce the competition among the bidding firms, making it harder for the intermediary to extract the surplus (as shown in our previous analysis). To solve this problem, Levin and Milgrom (2010) have introduced the idea of conflation where "similar but distinct products are treated as identical in order to make markets thick". In the online advertising market, this idea would translate into implementing a *broad* targeting rather than an *exact* matching, making it harder for experienced advertisers to cherry-pick impressions and hurt intermediaries' revenues. The results of our paper provide formal support for this argument by showing that, indeed, while the collection and use of precise information about the consumers during the targeting process is beneficial for advertisers, it can hurt the intermediary's revenue. As such, the intermediary in our model has incentive to use only the vertical information about the consumer (and not the horizontal) to reduce the precision of the targeting and of the matching process between advertisers and consumers.

8 Discussion and Conclusions

In this paper, we develop an analytical framework with the goal of providing a structured approach to understanding the role of personal data in a market for targeted advertisement. The model

focuses on the interaction among three types of agents: advertising firms (who compete with each other for consumer attention), consumers (the online users, who visit websites, are shown targeted ads, and purchase products online), and an intermediary (the Ad exchange platform). We assume that consumers are differentiated along both a horizontal dimension (e.g., brand preference) and a vertical dimension (e.g., purchasing power). Consumer information may indicate either a consumer’s vertical or horizontal location. Advertisers buy advertisements by participating in real-time auctions run by the intermediary.

We consider four informational regimes that differ in the type and amount of information that is available on the market and that advertising firms have available during the bidding process: i) a regime where only the horizontal information, that is which product a consumer prefers, is available; ii) a regime where only the vertical information (whether the consumer is high valuation or low valuation) is available; iii) a regime where both the horizontal and the vertical information about consumers is available; and iv) a regime where no information about consumers is available, corresponding to a benchmark case of complete data protection. For each of the four regimes, we derived the advertisers’ bidding strategy and pricing strategy; we then determined the winner of the auction and the final outcome of the game in terms of the advertisers’ payoff, the Ad exchange’s payoff and the consumer welfare. Our findings suggest that consumer welfare is higher when only specific type of information is collected and shared and, under certain conditions, when less information is collected and shared. Furthermore, there exist situations in which the incentives of the intermediary are misaligned with respect to the consumer interest; stated differently, the intermediary that acts as a profit-maximizing agent may decide to adopt strategies that increase its expected payoff to the detriment of the other agents involved in the process.

The proposed model and findings are not without limitations. The model takes into consideration the interaction between three types of agents: advertisers, intermediary and consumers. The online advertising ecosystem is more complex and includes additional subjects. For instance, companies that want to participate in auctions for online advertising usually rely on a Demand Side Platform (DSP) that serves advertisers or ad agencies by bidding for their campaigns in multiple ad networks automatically. On the other side, Supply Side Platforms (SSPs) serve publishers by registering their inventories (ad space) in different ad networks and accepting the most beneficial automatically. Publishers themselves play a fundamental role in determining how much of their

inventory is sold through programmatic advertising. Similarly, the Ad exchange is modeled as a monopolist intermediary that acts undisturbed. This modeling approach captures the empirical reality that the advertising market is highly concentrated, with few ad-selling companies (such as Google and Facebook) that account for about 75% of the total revenues generated in the advertising market (IAB 2015). However, a possible extension of the model would allow some degree of competition among intermediaries and consider the possibility that companies decide to enter different ad networks. Additionally, the consumer behavior is simplified to make the analysis tractable. The model assumes that the consumer sees one ad at the time and that she buys as long as the reservation price is lower than the price for the advertised product, abstracting away searching behavior.

Despite the highlighted limitations, the model provides subtle insights on the mechanisms through which the collection and disclosure of personal data can affect consumers' welfare. By illustrating how different types of consumer data tracking and sharing can differentially affect the welfare of data holders and data subjects, these findings can contribute to the ongoing industry and regulatory debate over the economic and social implications of the adoption of tracking and advertising systems. Notably, these results do not imply that the collection of consumer information should be prohibited—rather, they suggest that there is information which is beneficial for consumers to share, and other information which, instead, could be used by others to the consumer's detriment. Furthermore, what emerges from the model is a scenario alternative to the economic “win-win” often heralded as the likely and desirable outcome of the increased collection and trade of consumers' data (Unanimis Consulting Limited 2011). We find, instead, that different market agents that operate in the ecosystem may have contrasting interests. In turn, such interests may create incentives for practices that are not transparent. For instance, an intermediary that has the power to control which type of information to highlight to other parties (such as advertising firms) during the auction process, may have the incentive to act strategically, by revealing the information that ensure him the highest expected return. According to our analysis, this may be against consumers' best interest.

APPENDIX

PROOF FOR THE LEMMAS. Each firm's pricing strategy and bidding strategy in the four informational scenarios are derived by backward induction: we first derive the pricing strategy; given the price, we derive the bidding strategy. In general, firm j chooses the price to maximize its expected profit, that is:

$$\Pi_j = \sum_{i,t} I(b_j(\varphi(i,t)) > b_{-j}(\varphi(i,t))) * \alpha [I(z_{itj} \geq p_j(\varphi(i,t))) * \beta * p_j(\varphi(i,t)) + I(z_{itj} \geq p_j(\varphi(i,t))) * (1 - \beta) * p_j(\varphi(i,t))] - b_{-j}(\varphi(i,t))$$

where p_j is the price set by company j . Whether or not the consumer buys product j depends on two things: i) firm j has to win the auction for the advertisement, $I(b_j > b_{-j})$; and ii) conditional on seeing the ad, the consumer's reservation price for the product must not be lower than the price offered by the company, $I(z_{itj} \geq p_j)$.¹⁰

The basic reasoning for the derivation of the results is the same for all the four Lemmas we presented in the Analysis. As such, we present complete proof for Lemma 2. The proof can be extended to the other results as well.

In the scenario summarized by Lemma 2, we assume that advertising firm observes a consumer of type i (the horizontal information is available), with $i = 1, 2$. Remember, in this scenario firms have no information on the consumer's vertical position (that is, t).

Let us start from the pricing strategy and let us consider the sub-game where firms try to capture their own segment of loyal consumers. Company j can set two different prices: v_h or v_l . If it sets $p_j = v_l$, all the consumers in the company's segment will buy the product and, therefore, expected revenue will be $\Pi_j = \sum_{i,t} [\alpha v_l - b_{-j}]$. If the company sets $p_j = v_h$, only the high valuation consumers will buy the product. As a consequence, company's expected revenue becomes: $\Pi_j = \sum_{i,t} [\alpha \beta v_h - b_{-j}]$. In this case, not all the consumers are going to buy the product but only the proportion β of high valuation consumers. Given the equations specified above, firm j sets $p_j = v_h$ when $\beta v_h \geq v_l$ and it sets $p_j = v_l$ otherwise. The same reasoning can be applied in deriving pricing and bidding strategy for firm $-j$. It follows that: $p_{-j} = w_h$ when $\beta w_h \geq w_l$ and $p_{-j} = w_l$ otherwise.

¹⁰We drop out the $\varphi(i,t)$ for easy of exposition.

Next, let us consider the bidding strategy. Our results are based on the fact that in second-price auctions, truthful bidding is a dominant strategy. For the result to hold in this case, it is sufficient to introduce an infinitesimal probability that companies do not know who they are competing with. Holding this condition, company j 's strategy will be to bid its truthful valuation for consumer j , that is equal to the revenue the company expects to gain if that consumer buys the product. When company j sets a price equal to v_l , the expected revenue is also v_l ; when it sets a price equal to v_h , its expected revenue is equal to βv_h . Consequently, company j 's bidding strategy is to bid $b_j = \max\{\beta v_h, v_l\}$. We can similarly derive the bidding strategy for firm $-j$ as being equal to $b_{-j} = \max\{\beta w_h, w_l\}$.

From the bidding strategies, it can be easily seen that when firms observe consumers' horizontal information, in equilibrium firms will win their loyal consumers; that is, firm j wins the auction for consumer i , for $i = j$. To see why, simply consider the following four cases. Let us assume $b_j = \beta v_h$ and $b_{-j} = \beta w_h$. Then $b_j \geq b_{-j}$ as, by assumption, $v_h \geq w_h$. The same conclusion holds if $b_{-j} = w_l$. Indeed, if $b_j = \beta v_h$, it means that $\beta v_h \geq v_l$; since, by assumption, $v_l \geq w_l$, then it must also be that $\beta v_h \geq w_l$. Hence, $b_j \geq b_{-j}$. Next, let us assume that $b_j = v_l$ and $b_{-j} = \beta w_h$. We know that, by assumption, $v_l \geq w_h$. Since $\beta \leq 1$, we also have that $v_l \geq \beta w_h$. The same result holds if $b_{-j} = w_l$ as again, by assumption, $v_l \geq w_l$.

While a sub-game where advertising firms try to win both types of consumers is theoretically possible, it can be easily shown that such a sub-game will not occur in equilibrium because of the incentives to truthfully bid in a second-price auction.

The proofs for Lemmas 1, 3 and 4 follow the same logic.

PROOF FOR THE PROPOSITIONS

We present below the full proof for Proposition 3. Propositions 1 and 2 are derived in the process.

1) When $\beta v_h \geq v_l$, $\alpha v_h \geq w_h$ and $w_h/v_h \leq \frac{\alpha\beta}{\alpha+\beta-\alpha\beta}$ or $v_l/w_h \leq \frac{\alpha\beta}{\alpha+\beta-\alpha\beta}$, consumers are indifferent.

The region is characterized by the following conditions: $\beta v_h \geq v_l$ and $\alpha v_h \geq w_h$. v_l can be lower or greater than w_h . This implies the following outcomes in the different scenarios:

- a) Horizontal Information: $b_j^* = \beta v_h$, $p^* = v_h$, $W^{Hor} = \beta(v_h - v_h) = 0$.
- b) Vertical Information: $b_j^* = \alpha v_h$, αv_l , $p^* = \{v_h, v_l\}$, $W^{Ver} = \alpha\beta(v_h - v_h) + \alpha(1 - \beta)(v_l - v_l) = 0$.
- c) No Information: We have two different maximization problems depending on whether $v_l \geq w_h$ or $v_l < w_h$.

- For $v_l \geq w_h$, $b_j^* = \max\{w_l, \alpha\beta v_h, \alpha v_l, \beta w_h + (1 - \beta)\alpha w_h\}$. From the initial conditions, we know that $\alpha v_h \geq w_h$ implies that $w_l \leq \alpha v_l$, therefore w_l cannot be the maximum. Also, $\beta v_h \geq v_l$ implies that $\alpha v_l \leq \alpha\beta v_h$; as a consequence, αv_l cannot be the maximum. Therefore, the maximum is either $\alpha\beta v_h$ or $\beta w_h + (1 - \beta)\alpha w_h$.

When $w_h/v_h \leq \frac{\alpha}{\alpha+\beta-\alpha\beta}$, $b_j^* = \alpha\beta v_h$, $p^* = v_h$, and $CSn = \alpha\beta(v_h - v_h) = 0$.

Therefore, when $w_h/v_h \leq \frac{\alpha}{\alpha+\beta-\alpha\beta}$: $W^{Hor} = W^{Ver} = W^N = W^{Com} = 0$; $S^N = \{i, t | z_{it}^N - p_{it}^N > 0\}$, $S^{Hor} = \{i, t | z_{it}^{Hor} - p_{it}^{Hor} > 0\}$ and $S^{Ver} = \{i, t | z_{it}^{Ver} - p_{it}^{Ver} > 0\}$ with $S^{Ver} = S^{Hor} = S^N = \emptyset$.

- For $v_l < w_h$, $b_j^* = \max\{w_l, \alpha\beta v_h, \alpha v_l + \beta(1 - \alpha)v_l, \beta w_h\}$. From the initial conditions, we know that $\alpha v_h \geq w_h$ implies that $\beta w_h \leq \alpha\beta v_h$, therefore w_h cannot be the maximum. Also, $\beta v_h \geq v_l$ implies that $w_l \leq \beta w_h$; as a consequence, αv_l cannot be the maximum. Therefore, the maximum is either $\alpha\beta v_h$ or $\alpha v_l + \beta(1 - \alpha)v_l$.

When $v_l/v_h \leq \frac{\alpha\beta}{\alpha+\beta-\alpha\beta}$, $b_j^* = \alpha\beta v_h$, $p^* = v_h$, and $CSn = \alpha\beta(v_h - v_h) = 0$.

Therefore, when $v_l/v_h \leq \frac{\alpha\beta}{\alpha+\beta-\alpha\beta}$: $W^{Hor} = W^{Ver} = W^N = W^{Com} = 0$; $S^N = \{i, t | z_{it}^N - p_{it}^N > 0\}$, $S^{Hor} = \{i, t | z_{it}^{Hor} - p_{it}^{Hor} > 0\}$ and $S^{Ver} = \{i, t | z_{it}^{Ver} - p_{it}^{Ver} > 0\}$ with $S^{Ver} = S^{Hor} = S^N = \emptyset$.

2) When $\beta v_h \geq v_l$, $\alpha v_h \leq w_h$ and $v_l/w_h \leq \frac{\beta}{\alpha+\beta-\alpha\beta}$, revealing the Vertical Information is optimal for the consumers.

The region is characterized by the following relationships: $\beta v_h \geq v_l$ and $\alpha v_h \leq w_h$. Also, $v_l < w_h$. This implies the following outcomes in the different scenarios:

- a) Horizontal Information: $b_j^* = \beta v_h$, $p^* = v_h$, $W^{Hor} = \beta(v_h - v_h) = 0$.

b) Vertical Information: $b_j^* = \{w_h, w_l\}$, $p^* = \{w_h, w_l\}$, $W^{Ver} = \alpha\beta(v_h - w_h) + \alpha(1 - \beta)(v_l - w_l) > 0$.
c) No Information: $b_j^* = \max\{w_l, \alpha\beta v_h, \alpha v_l + \beta(1 - \alpha)v_l, \beta w_h\}$. From previous conditions we know that $\beta v_h \geq v_l$ implies that $\beta w_h \geq w_l$, therefore the maximum cannot be w_l . Also, $\alpha v_h \leq w_h$ implies that $\beta w_h \geq \alpha\beta v_h$ and $\alpha\beta v_h$ cannot be the maximum. Therefore, the maximum can be either $\alpha v_l + \beta(1 - \alpha)v_l$ or βw_h .

- When $w_h/v_l > \frac{\alpha + \beta - \alpha\beta}{\beta}$, $b_j^* = \beta w_h$, $p^* = w_h$, $W^N = \alpha\beta(v_h - w_h) > 0$.
- When $w_h/v_l \leq \frac{\alpha + \beta - \alpha\beta}{\beta}$, $b_j^* = \alpha v_l + \beta(1 - \alpha)v_l$, $p^* = v_l$ and $W^N = \alpha\beta(v_h - v_l) + (1 - \alpha)\beta(w_h - v_l) > 0$.

Therefore, when $w_h/v_l > \frac{\alpha + \beta - \alpha\beta}{\beta}$: $W^{Ver} > W^N > W^{Hor} = W^{Com} = 0$; $S^N = \{i, t | z_{it}^N - p_{it}^N > 0\}$ and $S^{Ver} = \{i, t | z_{it}^{Ver} - p_{it}^{Ver} > 0\}$ with $S^N \subset S^{Ver}$.

3) When $\beta v_h \leq v_l$, $\alpha v_h \leq w_h$, $\beta \geq (1 - \alpha)$ and $w_l/w_h \geq \alpha + \beta - \alpha\beta$ or $w_l/v_l \geq \alpha + \beta - \alpha\beta$, revealing no additional information is optimal for the consumers.

The region is characterized by the following relationships: $\beta v_h \leq v_l$ and $\alpha v_h \leq w_h$. We can have $v_l \geq w_h$ or $v_l < w_h$. This implies the following outcomes in the different scenarios:

a) Horizontal Information: $b_j^* = v_l$, $p^* = v_l$, $W^{Hor} = \beta(v_h - v_l) > 0$. b) Vertical Information: $b_j^* = \{w_h, w_l\}$, $p^* = \{w_h, w_l\}$, $W^{Ver} = \alpha\beta(v_h - w_h) + \alpha(1 - \beta)(v_l - w_l) > 0$. c) No Information: We have two different maximization problems depending on whether $v_l \geq w_h$ or $v_l < w_h$.

- For $v_l \geq w_h$, $b_j^* = \max\{w_l, \alpha\beta v_h, \alpha v_l, \beta w_h + (1 - \beta)\alpha w_h\}$. From the initial conditions, we know that $\alpha v_h \leq w_h$ implies that $w_l \geq \alpha v_l$, therefore αv_l cannot be the maximum. Also, $\beta v_h \leq v_l$ implies that $\alpha v_l \geq \alpha\beta v_h$, implying again that $w_l \geq \alpha\beta v_h$; as a consequence, $\alpha\beta v_h$ cannot be the maximum. Therefore, the maximum is either w_l or $\beta w_h + (1 - \beta)\alpha w_h$.
- When $w_l/w_h \geq \alpha + \beta - \alpha\beta$, $b_j^* = w_l$, $p^* = w_l$, and $W^N = \alpha\beta(v_h - w_l) + \alpha(1 - \beta)(v_l - w_l) + (1 - \alpha)\beta(w_h - w_l) > 0$.

Therefore, when $w_l/w_h \geq \alpha + \beta - \alpha\beta$, $W^N > W^{Hor} > W^{Com}$ and $W^N > W^{Ver} > W^{Com}$; $S^N = \{i, t | z_{it}^N - p_{it}^N > 0\}$, $S^{Hor} = \{i, t | z_{it}^{Hor} - p_{it}^{Hor} > 0\}$ and $S^{Ver} = \{i, t | z_{it}^{Ver} - p_{it}^{Ver} > 0\}$ with $S^{Ver}, S^{Hor} \subset S^N$.

- For $v_l < w_h$, $b_j^* = \max\{w_l, \alpha\beta v_h, \alpha v_l + \beta(1-\alpha)v_l, \beta w_h\}$. From the initial conditions, we know that $\beta v_h \leq v_l$ implies that $w_l \geq \beta w_h$, therefore βw_h cannot be the maximum. Also, $\alpha v_h \leq w_h$ implies that $\beta w_h \geq \alpha\beta v_h$, further implying that $w_l \geq \alpha\beta v_h$; it follows that $\alpha\beta v_h$ cannot be the maximum. Therefore, the maximum is either w_l or $\alpha v_l + \beta(1-\alpha)v_l$. When $w_l/v_l \geq \alpha + \beta - \alpha\beta$, $b_j^* = w_l$, $p^* = w_l$, and $W^N = \alpha\beta(v_h - w_l) + \alpha(1-\beta)(v_l - w_l) + (1-\alpha)\beta(w_h - w_l) > 0$. Therefore, when $w_l/v_l \geq \alpha + \beta - \alpha\beta$, $W^N > W^{Hor} > W^{Com}$ and $W^N > W^{Ver} > W^{Com}$; $S^N = \{i, t | z_{it}^N - p_{it}^N > 0\}$, $S^{Hor} = \{i, t | z_{it}^{Hor} - p_{it}^{Hor} > 0\}$ and $S^{Ver} = \{i, t | z_{it}^{Ver} - p_{it}^{Ver} > 0\}$ with $S^{Ver}, S^{Hor} \subset S^N$.

4) When $\beta v_h \leq v_l$, $\alpha v_h \geq w_h$ and $w_h/v_l \leq \frac{\alpha}{\alpha + \beta - \alpha\beta}$, revealing the Horizontal Information is optimal for the consumers.

The region is characterized by the following relationships: $\beta v_h \leq v_l$; $\alpha v_h \geq w_h$; and $v_l \geq w_h$.

This implies the following outcomes in the different scenarios:

- Horizontal Information: $b_j^* = v_l$, $p^* = v_l$, $W^{Hor} = \beta(v_h - v_l)$.
- Vertical Information: $b_j^* = \alpha v_h$, $p^* = v_h$, v_l , $W^{Ver} = \alpha\beta(v_h - v_h) + \alpha(1-\beta)(v_l - v_l)$.
- No Information: $b_j^* = \max\{w_l, \alpha\beta v_h, \alpha v_l, \beta w_h + (1-\beta)\alpha w_h\}$. From previous conditions we know that $\alpha v_h \geq w_h$ implies that $\alpha v_l \geq w_l$, therefore w_l cannot be the maximum. Also, $\beta v_h \leq v_l$ implies that $\alpha\beta v_h$ cannot be the maximum either. Therefore, the max can be either αv_l or $\beta w_h + (1-\beta)\alpha w_h$.

- When $w_h/v_l \leq \frac{\alpha}{\alpha + \beta - \alpha\beta}$, $b_j^* = \alpha v_l$, $p^* = v_l$ and $W^N = \alpha\beta(v_h - v_l)$.
- When $w_h/v_l > \frac{\alpha}{\alpha + \beta - \alpha\beta}$, $b_j^* = \beta w_h + (1-\beta)\alpha w_h$, $p^* = w_h$ and $CSn = \alpha\beta(v_h - w_h) + \alpha(1-\beta)(v_l - w_h) > 0$.

Therefore, when $w_h/v_l \leq \frac{\alpha}{\alpha + \beta - \alpha\beta}$: $W^{Hor} > W^N > W^{Ver} = W^{Com} = 0$ and $S^{Hor} = \{i, t | z_{it}^{Hor} - p_{it}^{Hor} > 0\}$ and $S^N = \{i, t | z_{it}^N - p_{it}^N > 0\}$ with $S_N \subset S^{Hor}$.

References

- Abraham I, Athey S, Babaioff M, Grubb M (2013) Peaches, lemons, and cookies: Designing auction markets with dispersed information. *EC '13 Proceedings of the fourteenth ACM conference on Electronic commerce*.
- Acquisti A, Taylor C, Wagman L (2016) The economics of privacy. *Journal of Economics Literature* 54(2):442–492.
- Bagwell K (2007) The economic analysis of advertising. *Handbook of Industrial Organization* 3:1701–1844.
- Bapna R, Goes P, Gupta A (2003) Analysis and design of business-to-consumer online auctions. *Management Science* 49(1):85–101.
- Boda K, Fldes AM, Gulys GG, Imre S (2011) User tracking on the web via cross-browser fingerprinting. Laud P, ed., *Information Security Technology for Applications*, 31–46 (Springer).
- Calzolari G, Pavan A (2006) On the optimality of privacy in sequential contracting. *Journal of Economic Theory* 130(1):168–204.
- Casadesus-Masanell R, Hervas-Drane A (2015) Competing with privacy. *Management Science* 61(1):229–246.
- Celis LE, Lewis G, Mobius MM, Nazerzadeh H (2011) Buy-it-now or take-a-chance: A simple sequential screening mechanism. *Proceedings of the 20th International Conference on World Wide Web*, 147–156, WWW '11 (New York, NY, USA: ACM).
- Chen H, He C (2011) Paid placement: Advertising and search on the internet. *The Economic Journal* 121(556):309–328.
- Chen J, Stallaert J (2014) An economic analysis of online advertising using behavioral targeting. *MIS Quarterly* 38(2):429–449.
- Eliasz K, Spiegler R (2011) A simple model of search engine pricing. *Economic Journal* 121.
- Esteban L, Hernandez J (2007) Strategic targeted advertising and market fragmentation. *Economics Bulletin* 12(10):1–12.
- FTC (2012) Protecting consumer privacy in an era of rapid change. Technical report.
- Goldfarb A (2014) What is different about online advertising. *Review of Industrial Organization* 44(20):115–129.
- Hagiu A, Jullien B (2011) Why do intermediaries divert search? *RAND Journal of Economics* 42(2):337–362.
- IAB (2015) Internet advertising report. Technical report.
- ITIF (2010) Stricter privacy regulations for online advertising will harm the free internet. Technical report.
- Iyer G, Soberman D, Villas-Boas J (2005) The targeting of advertising. *Marketing Science* 24(3):461–476.

- Kreps D, Wilson R (1982) Reputation and imperfect information. *Journal of Economic Theory* 27(2):253–279.
- Levin J, Milgrom P (2010) Online advertising: Heterogeneity and conflation in market design. *American Economic Review: Papers & Proceedings* 100:603–607.
- Liu D, Chen J, Whinston A (2010) Ex ante information and design of keyword auctions. *Information System Research* 21(1):133–153.
- Pierson J, Heyman R (2011) Social media and cookies: challenges for online privacy. *Digital Policy, Regulation and Governance* 13(6):30–42.
- Pinker E, Seidmann A, Vakrat Y (2003) Managing online auctions: Current business and research issues. *Management Science* 49(11):1457–1484.
- Research, Markets (2016) Real-time bidding market in the us 2016-2020. Technical report.
- Selten R (1965) Spieltheoretische behandlung eines oligopolmodells mit nachfrageträgheit. *Zeitschrift für die gesamte Staatswissenschaft* 121(1):667– 689.
- Soberman D (2004) Additional learning and implications on the role of informative advertising. *Management Science* 50(12):1744–1750.
- Sun Z, Dawande M, Janakiraman G, Mookerjee V (2016) The making of a good impression: Information hiding in ad exchanges. *MIS Quarterly* 40(3):717–739.
- Taylor C (2004) Consumer privacy and the market for customer information. *RAND Journal of Economics* 35(4):631–650.
- Tucker C (2012) The economics of advertising and privacy. *International journal of Industrial organization* 30(3):326–329.
- Varian H (2009) Economic aspects of personal privacy. Lehr W, Pupillo L, eds., *Internet Policy and Economics*, chapter 7, 101–109 (Springer US).
- Vickrey W (1961) Counterspeculation, auctions, and competitive sealed tenders. *The Journal of Finance* 16(1):8–37.
- Zhang K, Katona Z (2012) Contextual advertising. *Marketing Science* 31(6):980–994.