

## Data, Competition, and Digital Platforms<sup>†</sup>

By DIRK BERGEMANN AND ALESSANDRO BONATTI\*

*A monopolist platform uses data to match heterogeneous consumers with multiproduct sellers. The consumers can purchase the products on the platform or search off the platform. The platform sells targeted ads to sellers that recommend their products to consumers and reveals information to consumers about their match values. The revenue-optimal mechanism is a managed advertising campaign that matches products and preferences efficiently. In equilibrium, sellers offer higher qualities at lower unit prices on than off platform. The platform exploits its information advantage to increase its bargaining power vis-à-vis the sellers. Finally, privacy-respecting data-governance rules can lead to welfare gains for consumers. (JEL D11, D42, D44, D82, D83, M37)*

**Motivation.**—The role of data in shaping competition in online markets is a critical issue for both economics and policy. Digital platforms such as Amazon, Facebook, and Google in the United States and Alibaba, JD, and Tencent in China operate as matching engines that connect users and content providers. These platforms collect increasingly large and precise datasets that allow better pairing of consumers and sellers. They also monetize their data by selling sponsored content and targeted advertising. Therefore, digital platforms not only serve as gatekeepers of information online but also act as competition managers.<sup>1</sup>

Regulators fear that platforms may leverage their gatekeeper position to increase sellers' market power, thereby raising their willingness to pay for advertising.<sup>2</sup> The optimal regulatory response to the current business practices of digital platforms, if any, depends on the answers to a number of open questions, including the following:

\* Bergemann: Department of Economics, Yale University (email: [dirk.bergemann@yale.edu](mailto:dirk.bergemann@yale.edu)); Bonatti: Sloan School of Management, MIT (email: [bonatti@mit.edu](mailto:bonatti@mit.edu)). Jeff Ely was the coeditor for this article. We acknowledge financial support through NSF Grant SES-1948692, the Omidyar Network, and the Sloan Foundation. We thank the coeditor and four referees for productive suggestions. We thank audiences at Bank of Canada, Berkeley, ESMT, IESE, Indiana, Northwestern, NYU, Paris-Panthéon, Toulouse, as well as participants in the CEPR Virtual IO Seminar, the Luohan Academy Webinar, the FTC Microeconomics Conference, the Asia-Pacific Industrial Organization Conference, the 2022 Market Innovation Workshop, and the 2022 NBER Economics of Privacy Conference. We are grateful to our discussants Andrei Hagiu, Alessandro Pavan, and Vasiliki Skreta for helpful comments and suggestions, and to Lucas Barros, Andrew Koh, Roi Orzach, and Nick Wu for stellar research assistance.

<sup>†</sup>Go to <https://doi.org/10.1257/aer.20230478> to visit the article page for additional materials and author disclosure statement(s).

<sup>1</sup>The European Union's Digital Markets Act establishes a set of criteria for a large online platform to be a "gatekeeper" (Article 3): (i) a strong economic position in the EU market, (ii) a strong intermediation position, and (iii) an entrenched and durable position in the market.

<sup>2</sup>The report by Crémer, Montjoye, and Schweitzer (2019, p. 62) explicitly warns that "one cannot exclude the possibility that a dominant platform could have incentives to sell 'monopoly positions' to sellers by showing buyers alternatives which do not meet their needs."

(i) How does the precision of a digital platform's data affect the creation and distribution of surplus, both on and off the platform? (ii) How do these effects depend on the intensity of competition among sellers? (iii) How do they depend on the mechanisms for collecting and sharing consumer data?

In this paper, we develop a model of an intermediated online marketplace and trace how a data-rich platform creates and distributes surplus among market participants. Our goal is to provide a tractable and flexible framework to study digital markets where different privacy regimes can be compared. Our model captures three ubiquitous features of digital platforms. First, the platform leverages an informational advantage to personalize the *sponsored content* at the individual consumer level through managed advertising campaigns. Second, while price discrimination is rare, targeted advertising and personalized search results and recommendations are common and amount to *product steering*.<sup>3</sup> Third, sellers have parallel sales channels; i.e., consumers can buy their products both on and off digital platforms.

To capture these features, we consider a market with differentiated quality-pricing sellers. Consumers have heterogeneous preferences for the product lines of the sellers and are imperfectly informed about their own values. The platform's data augment the consumers' imperfect information. For each consumer, the platform identifies the most valuable seller and the most valuable product within that seller's product line and shows it to the consumer as a sponsored search result or product recommendation. The platform's superior information therefore creates the potential for product steering, whereby consumers with higher willingness to pay are shown higher-quality and higher-priced goods.

A critical issue is then how the platform monetizes the value of the data. In principle, the platform might use a number of instruments: it could charge commissions for every realized transaction or a fixed fee for each sponsored product listing; it could also run an item-by-item auction for every individual consumer visit to the platform. In our model, the digital platform charges each seller a lump-sum fee for a bundle of realized matches with consumers. This mechanism is often referred to as a *managed campaign* because the digital platform actively manages the sellers' advertising strategy.

Managed campaigns are the predominant mode of selling advertisements in real-world digital markets, including sponsored search, display advertising, and sponsored product listings. In a managed campaign, sellers set a fixed advertising budget, specify high-level objectives (e.g., to maximize clicks or conversions) for their campaigns, and leave the task of targeting individual consumers to "auto-bidding" algorithms offered by the platform, such as Google Performance Max or Meta Advantage+.<sup>4</sup> Some recent estimates—see Deng et al. (2023) and Deng et al. (2022)—suggest that over 80 percent of digital advertising is now generated by managed campaigns. In *pure advertising* digital platforms, such as Google, Yahoo, and social media platforms such as Meta, TikTok, and YouTube, advertising generates the overwhelming majority of revenue, and managed campaigns are

<sup>3</sup>Donnelly, Kanodia, and Morozov (2022) document the effect of personalized recommendations on a retail platform, and Raval (2020) illustrates a recent shift in eBay policy.

<sup>4</sup>For more details about programmatic or "smart" bidding on Google, see <https://support.google.com/google-ads/answer/6268637>. For recent academic work on algorithmic bidding, see Aggarwal, Badanidiyuru, and Mehta (2019); Balseiro et al. (2021); and Deng et al. (2021).

the dominant sales mechanism. Sponsored product listings are also the only source of revenue for Alibaba's Taobao consumer-to-consumer online marketplace and a significant revenue generator for several retail platforms that also charge merchant fees, including eBay, Wayfair, Booking, Orbitz, Amazon, and Alibaba's Tmall business-to-consumer e-commerce platform.<sup>5</sup>

Our framework also introduces a pool of consumers who shop off the platform and face search costs. In the tradition of the Diamond (1971) model, these consumers have zero cost to search the first seller and positive cost to visit additional sellers. In equilibrium, each consumer only visits the website of a single seller—the one whose products they (believe they) value most. The presence of the off-platform sales channel restrains sellers' ability to extract consumer surplus on the platform because the on-platform consumers can seamlessly move from the platform to individual websites off the platform. In particular, the more the seller wants to trade with the loyal consumers off the platform, the less flexibility it has to offer targeted products and prices on the platform. Thus, the consumer's choice of sales channel limits the scope for price discrimination.

*Results.*—The platform's informational advantage over consumers and the search frictions on the platform, no matter how small, give the digital platform significant bargaining power over sellers. Our first main result shows that the platform can completely control consumers' shopping behavior and steer them away from sellers that do not submit an advertising budget (Proposition 2). Consumers understand the managed campaign mechanism and expect that in equilibrium, advertised products generate the highest value. In the presence of search costs, consumers only consider buying from the advertised brand, whether on platform or off platform. As a result, the platform restricts competition among sellers, as each seller only faces consumers who are most interested in their products and competes with their own off-platform offers only (Proposition 3).

This leads to sellers facing an additional opportunity cost of generating surplus off the platform. Not only must they concede information rents to off-platform consumers but they must also lower their prices on the platform. This has two welfare consequences. First, the equilibrium quality levels of off-platform products are distorted downward from the efficient levels even more than in the model proposed by Mussa and Rosen (1978). Second, the platform is able to extract most of the surplus it generates, as it only needs to compensate sellers for the additional distortions in their off-platform menus of products (Proposition 4).

Next, we examine the platform's dual gatekeeper role by considering two sources of the platform's bargaining power: its information advantage over consumers and sellers, and consumers' search costs off the platform. We first show that it would be against the platform's interest to provide consumers with information about nonsponsored products. We assume that consumers on the platform observe their values perfectly, so the platform cannot steer their behavior away

<sup>5</sup> Indeed, Taobao does not assess fees for completing transactions but generates revenue from "paid placement" (i.e., sponsored product listings) alone. Tmall charges very small fees (of around 5 percent) and sells paid placement as well. In this context, Amazon's use of commission fees (15 percent of product prices on average) are the exception, not the rule. For our model, Proposition 4 establishes that lump-sum fees generate greater revenue for the platform than variable sales fees. The online Appendix analyzes an extension of our framework to sales fees.

from their favorite seller even if that seller rejects the platform's offer. Complete information for on-platform consumers does not change the equilibrium prices or products, but it reduces the platform's fees (Proposition 5). We then assume that the platform offers organic links that advertise all off-platform prices and products to all on-platform consumers. We show that the provision of price information introduces menu competition among sellers, reducing all equilibrium prices, both on and off the platform, as well as the platform's fees (Proposition 6).

We also investigate how data governance, which consists of rules governing the collection and usage of consumer data by the digital platform, affects the creation and distribution of social surplus. Specifically, we examine how limiting the use of consumer data may impact the market outcome. A prominent approach in practice is cohort-based advertising, where tracking individual users is replaced by tracking cohorts of anonymized users who share similar characteristics and online behavior. In our model, we define a cohort of consumers as those who have the same ranking of the sellers. Therefore, a cohort shares the same ordinal ranking without revealing any cardinal information about each member's willingness to pay. We show that cohort-based advertising, which protects privacy and allows consumers to retain an information advantage over sellers on the platform, improves consumer surplus (Proposition 7).

So far, the platform has been using all the additional information for product steering and pricing recommendations. We then explore whether the platform can do even better by employing the additional information only partially and stochastically. When the on-platform market is large compared to the off-platform market, using complete information indeed maximizes the platform's advertising revenue (Proposition 8). However, when consumers only know the prior distribution and the off-platform market is sufficiently large, the platform can increase its revenue by offering a more limited disclosure policy, which we fully characterize (Proposition 9).

*Related Literature.*—This paper is most closely related to the literature on information gatekeepers pioneered by Baye and Morgan (2001) and on the conflict of interest between intermediaries and customers they serve. Many recent contributions, including Armstrong and Zhou (2011); Condorelli and Szentes (2024); de Cornière and Taylor (2019); Gomes and Pavan (2016); Gur et al. (2022); Hagiu and Jullien (2011); Inderst and Ottaviani (2012a, b); Ke, Ling, and Lu (2022); Rayo and Segal (2010); and Shi (2022), analyze the steering role of platforms that strategically modify search results, e.g., to match consumers with the sellers that pay the largest commissions. The trade-off between value creation through personalization and consumer surplus extraction is also central in Hidir and Vellodi (2021) and Ichihashi (2020).

The provision of information by a digital platform is central to the model of de Cornière and de Nijs (2016), who examine a platform's incentives to provide match-value information to differentiated sellers in a second-price auction model. More recently, Teh and Wright (2022) study the signaling role of ranking the search results, and Zhu, Moorthy, and Shi (2022) study the impact of a platform's privacy policies on the downstream competition within and across product categories. Relative to these papers, we allow the platform to provide information directly to the consumer, e.g., through product reviews. Moreover, the multiproduct sellers in our

model use the platform's information to tailor their quality level to the consumer's preferences. This captures surplus creation and product steering *within* a match. Relative to the papers above, our model focuses on sponsored links and advertising platforms. Hence, sellers pay fees (or bids) that do not vary with the prices of their products.<sup>6</sup>

A recent body of work, including Choi, Jeon, and Kim (2019); Acemoglu et al. (2022); Ichihashi (2021); Kirpalani and Philippon (2021); and Bergemann, Bonatti, and Gan (2022), documented the *data externalities* that consumers impose on each other when they share their information with a digital platform. In the present paper, the growth of a platform's database (e.g., through the participation of more consumers) influences the ability to match products to tastes but also endogenously reduces each consumer's outside option.

The forces at work in our paper relate to a growing literature on showrooming, product lines, and multiple sales channels, including Bar-Isaac and Shelegia (2020); Idem (2021); Miklós-Thal and Shaffer (2021); and Wang and Wright (2020). In particular, Anderson and Bedre-Defolie (forthcoming) introduce the self-preferencing problem by letting the platform choose whether to be hybrid, i.e., to sell the private label products. Unlike in these papers, the sellers in our model are concerned about showrooming because the opportunity to sell on the platform benefits them through the added value of making personalized offers. The showrooming constraint—whereby on-platform consumers can pretend to be off-platform consumers, but not vice versa—is also similar to the asymmetric incentive compatibility constraint in Krämer and Strausz (2023), where a seller must prevent classical consumers from mimicking privacy-conscious consumers.

Our analysis of parallel sales channels is further related to “partial mechanism design,” or “mechanism design with a competitive fringe,” e.g., Philippon and Skreta (2012); Tirole (2012); Calzolari and Denicolo (2015); and Fuchs and Skrzypacz (2015). In these papers, the platform is limited in the ability to monopolize the market since the sellers have access to an outside option. Our setting shares some of the same features but in an oligopoly environment where sellers compete for heterogeneous consumers. Furthermore, the sellers choose their product menus understanding that customers arrive through two different channels and that they have distinct information in each channel.

At a broad level, this paper relates to information structures in advertising auctions, e.g., Bergemann, Brooks, and Morris (2021) and Gur et al. (2022), and to nonlinear pricing, market segmentation, and competition, e.g., Bergemann, Brooks, and Morris (2015); Bonatti (2011); Elliott et al. (2020); and Yang (2022). Our analysis could also be extended to discuss self-preferencing by a monopoly platform. In this sense, our paper also relates to Hagiu, Teh, and Wright (forthcoming); Kang and Muir (2021); Lam (2022); Lee (2022); Lee and Musolff (2021); and Padilla, Perkins, and Piccolo (2020).

Finally, in parallel work, Bergemann, Bonatti, and Wu (forthcoming) compare managed advertising campaigns and data-augmented auctions for online advertising.

<sup>6</sup>In Gomes and Pavan (forthcoming), a monopolist platform elicits sellers' willingness to pay for consumer attention through a nonlinear tariff for advertising space, which is analogous to our managed campaign. In contrast, we explicitly model consumer search and determine the equilibrium product prices.

In a setting with single-product sellers and unit consumer demand, they show that the optimal managed campaign mechanism increases the platform's revenue but also raises the off-platform equilibrium prices.

## I. Model

*Sellers and Consumers.*—We consider a digital platform and  $J$  differentiated multiproduct sellers. Each seller  $j$  offers a product line (or menu) of quality differentiated products. As in Mussa and Rosen (1978), each seller  $j$  can produce a good of quality  $q_j$  at a cost

$$c(q_j) = \frac{q_j^2}{2}.$$

There is a unit mass of consumers with single-unit demand. Each consumer is described by a vector  $\theta$  of willingness to pay for the quality for each seller's products,

$$\theta = (\theta_1, \dots, \theta_j, \dots, \theta_J) \in [\underline{\theta}, \bar{\theta}]^J,$$

with  $0 \leq \underline{\theta} < \bar{\theta} < \infty$ . We refer to the vector  $\theta$  as the value profile of consumer  $i$ . Given a quality  $q_j$  offered by seller  $j$ , the consumer receives a gross utility:

$$u(\theta, q_j) = \theta_j q_j.$$

*Information.*—The values  $\theta_j$  of each consumer for each seller  $j$  are i.i.d. across consumers and sellers with marginal distribution  $F(\theta_j)$  and density  $f(\theta_j)$ . Each consumer has incomplete information about their true value  $\theta_j$ . In particular, each consumer only has partial, private information with expectation  $m_j$  about  $\theta_j$ . (We could alternatively refer to  $m_j$  and  $\theta_j$  as interim and ex post value, respectively.) The expectations  $m_j$  are assumed to be i.i.d. with marginal distribution  $G(m_j)$  and density  $g(m_j)$ . The distribution of values and expectations,  $F$  and  $G$ , are implicitly related by an information structure. By Blackwell (1951), Theorem 5, there exists a signal  $s$  that induces a distribution  $G$  of expected values if and only if  $F$  is a mean-preserving spread of  $G$ . We recall that  $F$  is defined to be a mean-preserving spread of  $G$  if

$$(1) \quad \int_{\theta}^{\bar{\theta}} F(x) dx \leq \int_{\theta}^{\bar{\theta}} G(x) dx, \forall \theta \in [\underline{\theta}, \bar{\theta}],$$

with equality for  $\theta = \underline{\theta}$ . We say that  $F$  is a strict mean-preserving spread if the above inequality holds as a strict inequality in the interior of the support,

$$\int_{\theta}^{\bar{\theta}} F(x) dx < \int_{\theta}^{\bar{\theta}} G(x) dx, \forall \theta \in (\underline{\theta}, \bar{\theta}).$$

With the strict inequality, we guarantee that a consumer with information represented by  $G$  does not have complete information on any subinterval of  $[\underline{\theta}, \bar{\theta}]$ . If  $F$  is a strict mean-preserving spread of  $G$ , we write  $F \prec G$ , as  $G$  majorizes  $F$ . (Conversely,  $G$  is referred to as a mean-preserving contraction of the distribution  $F$ .)



*On Platform.*—A fraction  $\lambda \in [0, 1]$  of all consumers uses the platform to find a product. The platform has access to extensive data and knows each consumer's value profile  $\theta$ , while the sellers only know the corresponding prior distribution  $F$ . The platform offers a single recommendation to each consumer  $\theta$  through a specific price-quality pair  $(q_j(\theta), p_j(\theta))$  by a specific seller  $j$ , referred to as the *sponsored product* or *sponsored link*. In order to generate a recommendation to the consumer, the platform requests a menu from each seller that specifies which product-price pair seller  $j$  intends to advertise to each consumer profile  $\theta$ :

$$(2) \quad q_j : \Theta \rightarrow \mathbb{R}_+, p_j : \Theta \rightarrow \mathbb{R}_+.$$

The managed (advertising) campaign mechanism is defined formally as follows.

**DEFINITION 1** (Managed Advertising Campaign): *In a managed advertising campaign, the platform simultaneously*

- (i) *requests an advertising budget  $t_j \in \mathbb{R}_+$  from each seller  $j$ ;*
- (ii) *requests quality and price schedules  $q_j : \Theta \rightarrow \mathbb{R}_+$  and  $p_j : \Theta \rightarrow \mathbb{R}_+$ ;*
- (iii) *commits to advertising a seller and a product via a selection rule:*

$$\sigma : \Theta \times \mathbb{R}_+^{J\Theta} \times \mathbb{R}_+^{J\Theta} \rightarrow J \times \mathbb{R}_+^2.$$

The selection rule  $\sigma$  maps the consumer profile  $\theta$  and the price-quality menus  $\{q_j(\cdot), p_j(\cdot)\}_{j \in J}$  of all  $J$  sellers (each of “dimension”  $\Theta$ ) to a single seller  $j$  and to the price-quality pair  $(q_j(\theta), p_j(\theta))$  that the selected seller  $j$  wishes to advertise to consumer  $\theta$ .<sup>7</sup> Throughout the paper, we restrict attention to the symmetric mechanisms defined below.

**DEFINITION 2** (Symmetric Managed Campaign): *A symmetric managed campaign requests the same budget from all sellers ( $t_j = t$ ), and the selection rule  $\sigma$  does not condition on the identity of the sellers.*

Until Section V, we shall maintain the assumption that, upon advertising the sponsored product  $q_j$  to consumer  $\theta$ , the platform directly provides additional information to the consumer that fully reveals their value  $\theta_j$  for the advertised product.<sup>8</sup>

*Off Platform.*—The remaining  $1 - \lambda$  consumers buy off the platform, e.g., from the merchants' own websites or physical stores. Off the platform, the consumers have their expectation  $m$  and the sellers know the corresponding distribution  $G$ .

<sup>7</sup> At this stage, we allow the platform to select a seller who does not participate in a managed campaign, though this will not happen in equilibrium. We also assume without loss that a nonparticipating seller submits the single price-quality pair  $(q, p) = (0, 0)$ .

<sup>8</sup> With risk-neutral consumers, this assumption is equivalent to a framework in which the platform provides superior (i.e., Blackwell-ranked but not necessarily perfect) information relative to what is available off platform.

On the platform, there are extensive data, and the platform knows each consumer's value profile  $\theta$ .

The consumers who buy off the platform face positive (and arbitrarily small) search costs beyond the first search, as in Diamond (1971) and Anderson and Renault (1999). The expectation  $m$  is private information of the consumer. Therefore, seller  $j$  elicits the consumer's private information through a menu of (price, quality) pairs<sup>9</sup>

$$(3) \quad \left\{ (\hat{p}_j(m_j), \hat{q}_j(m_j)) \right\}_{m_j \in [\underline{\theta}, \bar{\theta}]}$$

as in Mussa and Rosen (1978) and Maskin and Riley (1984). At this stage, sellers can offer distinct sets of product qualities on and off the platform. Furthermore, the goods being sold are not experience or inspection goods: to learn any of the true values  $\theta$ , consumers and sellers must gain access to the platform's data.

After receiving seller  $j$ 's offer on the platform and learning their value  $\theta_j$  for the advertised product  $q_j$ , each consumer can search off the platform, in which case they also face strictly positive search costs. Unlike the off-platform consumers, however, any on-platform consumer can use the newly gained information to select a product from the off-platform schedule (3) posted by the advertised seller  $j$  or any other seller. Figure 1 summarizes the interaction between the agents and their actions in our model.

*Timing and Equilibrium.*—We consider a sequential game in which the platform first announces a mechanism, and then all sellers choose (on- and off-platform) prices simultaneously. The simultaneous timing in the subgame following the platform's announcement is meant to capture the great flexibility that algorithmic pricing offers sellers both on and off the platform. The exact timing of the game is as follows:

- (i) The platform announces a selection rule  $\sigma$  and requests an advertising budget  $t_j$  from each seller  $j$ .
- (ii) Sellers simultaneously set off-platform products  $\hat{q}_j(m)$  and prices  $\hat{p}_j(m)$ , choose whether to participate (i.e., to submit the budget  $t_j$ ), and if so, what products  $q_j(\theta)$  and prices  $p_j(\theta)$  to advertise.
- (iii) The platform shows a single advertisement—a product  $q_j(\theta)$  and a price  $p_j(\theta)$ —to each on-platform consumer according to the announced selection rule.
- (iv) The on-platform consumers learn their value  $\theta_j$  for the advertised seller and their expectations  $m$  for all sellers; they purchase the advertised product on the platform or search off the platform. The off-platform consumers learn their expectations  $m$  and choose which seller(s) to search for.

<sup>9</sup>Throughout the paper, we use the circumflex to distinguish off-platform from on-platform variables.



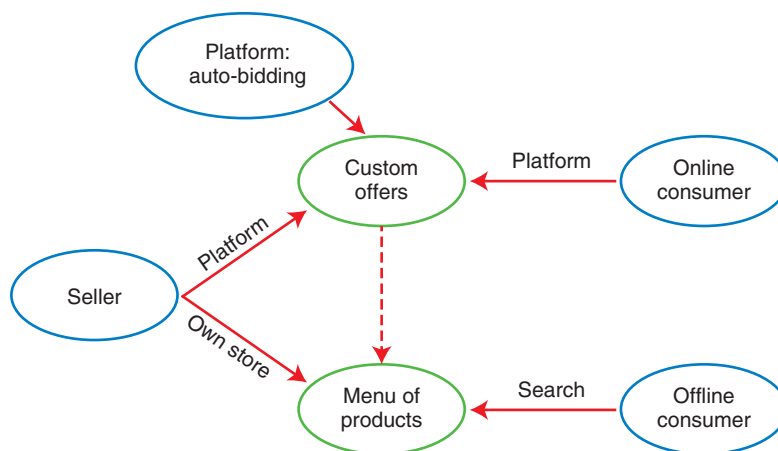


FIGURE 1. ON- AND OFF-PLATFORM MATCHING

**DEFINITION 3** (Symmetric Perfect Bayesian Equilibrium): *We consider symmetric perfect Bayesian equilibria. The consumers have symmetric beliefs over the sellers' off-platform menus both on and off the equilibrium path.*

The symmetry qualification in the definition above means that the ex ante identical sellers receive identical offers from the platform, make identical participation decisions, and offer identical prices and qualities to the consumers in equilibrium. In turn, the consumers expect the sellers to offer identical prices and qualities on the equilibrium path. Finally, consumers continue to hold symmetric (though not necessarily passive) beliefs over every seller's prices and qualities even when they observe a deviation (either on or off the platform).

*Discussion.*—The managed campaign mechanism captures the salient features of automated bidding in real-world digital advertising platforms. First, sellers do not acquire the platform's data, but they condition products and prices on the platform's information about the consumer. In other words, the consumer's value  $\theta$  acts as a *targeting category*. Second, fixed payments for advertising in the form of sponsored links reflect the advertising budget allocated to automated bidding for advertising links. Sellers submit a budget and upload the ads for the products they wish to show to select consumers. Third, as each seller can tailor the product offer to each consumer value, the platform creates an opportunity for surplus extraction through *product steering* without personalized prices.

We represent the informational friction in the off-platform market by means of the Diamond (1971) search model. This leads to the admittedly simplifying implication that the off-platform seller acts like a monopolist (in the absence of the interaction with the platform). This allows us to use the characterization of nonlinear pricing in Mussa and Rosen (1978) to trace the implications of the on-platform to off-platform prices. Yet our analysis does not rely in an essential manner on the Diamond (1971) paradox. A more elaborate representation of the off-platform market in the style of Perloff and Salop (1985) and Wolinsky (1986) would yield similar qualitative

trade-offs. If we were to represent the off-platform market as competition (rather than monopoly) with nonlinear pricing as in Bonatti (2011) or Garrett, Gomes, and Maestri (2019), then we would lose tractability. However, it would still be true that the returns from product steering on the platform would provide incentives to raise the prices off the platform. This key transmission mechanism will remain unaffected by the exact model of the off-platform market.

Several assumptions in the above mechanism can be easily relaxed. In particular, we allow the sellers' schedules  $q_j(\theta)$  and  $p_j(\theta)$  to condition on the entire value  $\theta$  and not just  $\theta_j$ . However, this additional flexibility will be redundant in equilibrium. Our model of a single sponsored link is also simple in that the platform sells both information and recognition to a single seller and a specific product. In Section V, we extend our model to allow for all brands and products to be present on the platform through organic search results that advertise their off-platform offers. Finally, the direct revelation of information to consumers captures the rich contextual detail that some retail platforms provide to their users. Here, we assume that the platform fully utilizes the informational advantage to generate surplus through efficient product-consumer matching. In Section VD, we relax the assumption of perfect revelation of  $\theta_j$ , and we study information design by the platform.

## II. A First Example: Single Seller

We begin by illustrating some of the central implications of our model with a simple example. The example considers a single seller (rather than many sellers) and binary values (rather than a multidimensional continuum of values). In addition, consumers know their true values both on and off the platform; thus,  $F = G$ . Nonetheless, the platform retains an informational advantage over the sellers because it learns the value of the consumer, which remains their private information off the platform.

The central result in this section (Proposition 1) describes the relationship between on-platform and off-platform pricing and quality provision. Thus, even this very basic setup sheds light on the fundamental interaction between on-platform and off-platform allocations.

We consider a single seller that encounters a mass  $\lambda$  of consumers on the platform and a mass  $1 - \lambda$  of consumers off the platform. Consumers can be of two values,  $\theta \in \{\underline{\theta}, \bar{\theta}\}$ , each with probability  $f(\theta)$ . The platform charges an advertising budget  $t$  to the seller. With the provided budget, the seller earns the right to offer a personalized product to each value of the consumer. However, each consumer on the platform can also shop from the seller's own website (i.e., buy products the seller offers off platform). Thus, the consumer's option to "showroom" limits the seller's ability to price discriminate.

If the seller accepts the platform's request of an advertising budget, it offers a menu of products on the platform, which we describe in terms of the product qualities  $q(\theta)$  and information rents  $U(\theta)$ :

$$(4) \quad \{(q(\theta), U(\theta))\}_{\theta \in \{\underline{\theta}, \bar{\theta}\}},$$

where the information rent  $U(\theta)$  is the net utility of the consumer on the platform:

$$U(\theta) \triangleq \theta q(\theta) - p(\theta), \theta \in \{\underline{\theta}, \bar{\theta}\}.$$

The seller also offers a menu off the platform, denoted by

$$(5) \quad \left\{ (\hat{q}(\theta), \hat{U}(\theta)) \right\}_{\theta \in \{\underline{\theta}, \bar{\theta}\}}.$$

The seller's profit is given by

$$\begin{aligned} \max_{q, U} \sum_{\theta \in \{\underline{\theta}, \bar{\theta}\}} f(\theta) & \left( \lambda \left[ \theta q(\theta) - \frac{q(\theta)^2}{2} - U(\theta) \right] \right. \\ & \left. + (1 - \lambda) \left[ \theta \hat{q}(\theta) - \frac{\hat{q}(\theta)^2}{2} - \hat{U}(\theta) \right] \right). \end{aligned}$$

The seller maximizes profit subject to the individual rationality constraints on and off the platform and to the incentive compatibility constraints *off* the platform only. Incentive compatibility is not required on the platform because the consumer receives a single targeted product offer. However, the targeted offer must induce the consumer not to buy off the platform.<sup>10</sup> The seller then faces the following new “showrooming” constraints:

$$U(\theta) \geq \hat{U}(\theta), \theta \in \{\underline{\theta}, \bar{\theta}\}.$$

In other words, each consumer  $\theta$  must prefer to purchase on the platform rather than to use the platform as a showroom and seek an alternative quality-price pair off the platform.

It follows that the seller should offer the socially efficient quality levels on the platform and that the showrooming constraint should bind,

$$(6) \quad q(\theta) = \theta \text{ and } U(\theta) = \hat{U}(\theta), \theta \in \{\underline{\theta}, \bar{\theta}\}.$$

We now characterize the quality levels off the platform. As usual, the equilibrium menu does not distort at the top ( $\hat{q}(\bar{\theta}) = \bar{\theta}$ ) and offers no rents at the bottom ( $\hat{U}(\underline{\theta}) = 0$ ). Furthermore, the incentive compatibility constraint binds for the high value. With these preliminary results, the seller's objective can be written as

$$\begin{aligned} (7) \quad \max_{q, U} & \left\{ \lambda \left( f(\underline{\theta}) \frac{\underline{\theta}^2}{2} + f(\bar{\theta}) \left[ \frac{\bar{\theta}^2}{2} - \hat{U}(\bar{\theta}) \right] \right) \right. \\ & \left. + (1 - \lambda) \left( f(\underline{\theta}) \left[ \underline{\theta} \hat{q}(\underline{\theta}) - \frac{\hat{q}(\underline{\theta})^2}{2} \right] + f(\bar{\theta}) \left[ \frac{\bar{\theta}^2}{2} - \hat{U}(\bar{\theta}) \right] \right) \right\} \end{aligned}$$

<sup>10</sup>This is weakly optimal for the seller because they could always offer the consumer the same (quality, price) pair the consumer would choose from the off-platform menu.

subject to the constraint

$$(8) \quad \hat{U}(\bar{\theta}) = (\bar{\theta} - \underline{\theta})\hat{q}(\underline{\theta}).$$

From this expression, it is immediate that the provision of quality to the low value off the platform is doubly costly for the seller: it forces the seller to lower the price for the high value off the platform, and it forces lower prices on the platform.

**PROPOSITION 1** (Single Seller and Binary Values): *The optimal off-platform menu of products for the seller is*

$$(9) \quad \hat{q}(\underline{\theta}) = \max \left\{ 0, \underline{\theta} - \frac{f(\bar{\theta})}{f(\underline{\theta})}(\bar{\theta} - \underline{\theta}) \left( 1 + \frac{\lambda}{1 - \lambda} \right) \right\},$$

$$\hat{q}(\bar{\theta}) = \bar{\theta}.$$

We relegate the formal proof of all our results to the Appendix. Let us now compare the optimal menu with the classic nonlinear pricing solution as in Mussa and Rosen (1978), which corresponds to the case  $\lambda = 0$ . In that case, we would have

$$(10) \quad \hat{q}(\underline{\theta}) = \max \left\{ 0, \underline{\theta} - \frac{f(\bar{\theta})}{f(\underline{\theta})}(\bar{\theta} - \underline{\theta}) \right\},$$

$$\hat{q}(\bar{\theta}) = \bar{\theta}.$$

Proposition 1 indicates an additional opportunity cost of serving the low value off the platform. Indeed, it is not difficult to find parameters (e.g.,  $\lambda$  large enough) for which the quality level on platform as in (10) is strictly positive but the quality off platform as in (9) is zero. Thus, without a platform, the seller would offer a low-quality product to the low value. However, for a sufficiently large platform, the low-value consumer is only offered a product on the platform, where the seller can make a different personalized offer to the high value. Off the platform, the seller prefers to forgo sales of the low product altogether, in order to sell product  $q(\bar{\theta}) = \bar{\theta}$  at a higher price on both channels. Indeed, when  $\hat{q}(\underline{\theta}) = 0$ , no consumer value receives any rent on or off the platform.

Finally, to determine the optimal advertising budget  $t^*$ , we need to consider what the on-platform consumers would do if the seller did not advertise. If these consumers can buy off the platform, the optimal advertising budget extracts the seller's extra profit relative to offering the menu in (10) to all consumers. If these consumers were not to buy at all, the seller's outside option is scaled by a factor  $1 - \lambda$ , and the optimal advertising budget is correspondingly higher.

The environment with many sellers and many values that we consider next requires a richer analysis. With many sellers, we must consider how the information on the platform impacts the search behavior off the platform (Proposition 2). In turn, this determines the shape of the menu offered by the sellers in the presence of competition (Proposition 3) and the nature of the revenue-maximizing mechanism for the platform (Proposition 4).

### III. Managed Campaign and Showrooming

We now analyze the environment with many sellers and a multidimensional continuum of values as introduced in Section I. Our first objective is to establish the equilibrium patterns of consumer search induced by the informational advantage of the platform. This advantage is captured by the distinction between values  $\theta_j$  with distribution  $F$  on the platform and expected values  $m_j$  with distribution  $G$  off the platform.

#### A. Managed Campaign

As we introduced in Section I, the platform designs a managed advertising campaign to select a sponsored product. The platform enables consumers and sellers to interact under symmetric information. Thus, the sellers are willing to pay for the right to make a personalized offer to each consumer they are matched to under the platform's managed campaign mechanism. However, the winning seller must induce the consumer to accept their personalized offer instead of buying from the off-platform store, i.e., not to use the platform for "showrooming." Thus, the seller's ability to product steer and price discriminate on the platform is limited by the presence of the off-platform channel.

The specific selection rule  $\sigma^*$  that matches consumers and products efficiently is defined below and plays an important role throughout the paper.

**DEFINITION 4 (Efficient Steering):** *A socially efficient steering policy  $\sigma^*$  recommends a price-quality pair  $(q_{j^*}, p_{j^*})$  by the seller  $j^*$  that maximizes social welfare among all participating sellers:*

$$(11) \quad \sigma^*(\theta, q(\cdot), p(\cdot)) = (j^*, q_{j^*}(\theta), p_{j^*}(\theta)), \text{ where} \\ j^* = \operatorname{argmax}_j \{ \theta_j q_j(\theta) - c(q_j(\theta)) \}.$$

We note that the socially efficient steering rule is independent of the price  $p_{j^*}(\theta)$  at which the item  $q_{j^*}(\theta)$  is offered on the platform. Indeed, Proposition 4 establishes the revenue optimality of this particular rule  $\sigma^*$  among all managed campaign mechanisms.

#### B. Choice Patterns with Efficient Steering

We now derive the consumers' choice patterns in a symmetric equilibrium under the efficient steering policy. We begin with the off-platform consumers. These consumers (who have mass  $1 - \lambda$ ) face positive search costs beyond the first seller. As a result, a consumer with expected value  $m$  visits only the seller that offers the highest expected value:

$$j^{(1)} = \operatorname{argmax}_j m_j.$$

This result does not depend on the magnitude of the search costs, as established famously by Diamond (1971).

If the platform has a strict informational advantage ( $F \succ G$ ), each on-platform consumer infers that the advertised seller  $j^*$  maximizes their willingness to pay, i.e.,

$$\theta_{j^*} = \max_j \theta_j.$$

Because these consumers expect symmetric menus off the platform and the information rent function associated with those menus is strictly increasing, these consumers consider products offered by the advertised seller  $j^*$  only.

**PROPOSITION 2 (Consideration Sets):** *Consider the socially efficient steering policy  $\sigma^*$ . In any symmetric equilibrium with full seller participation, every on-platform consumer  $\theta$  compares the advertised seller's on-platform offer  $(p_{j^*}(\theta), q_{j^*}(\theta))$  only with that seller's corresponding off-platform offer  $(\hat{p}_{j^*}(\theta_{j^*}), \hat{q}_{j^*}(\theta_{j^*}))$ .*

The platform augments the expectation of each consumer with additional data that lead to a revision from the expected value  $m$  to the (true) value  $\theta$ . Figure 2 illustrates the choice behavior by a consumer whose expected value  $m$  ranks the sellers differently relative to (true) value  $\theta$ . The consumer has two possible choices, thus,  $J = 2$  and the expected value  $m$  suggests that seller 1 offers a higher value, and thus,  $m_1 > m_2$ . Now suppose that on the platform the consumer is shown an advertisement by seller,  $j = 2$ . Indeed, the platform reveals to the consumer the value  $\theta_2$ , and thus, they will infer that  $\theta_2 > \theta_1$ . Therefore, by Proposition 2, the consumer either accepts seller 2's offer or shops off the platform from seller 2 but now with full knowledge of their value  $\theta_2$ .

An important implication of Proposition 2 is that every consumer will buy from a competing seller if they do not see their favorite seller's ad. Thus, participating in the platform's mechanism is necessary for each seller to access *any* of the on-platform consumers.

Indeed, as the platform has better information than the consumers do, any symmetric equilibrium of the game is outcome-equivalent to a simpler model where each seller has a group of  $(1 - \lambda)/J$  customers who only consider their brand. These consumers buy a product variety that depends on their expected value  $m_j$ , which is distributed according to  $G^J$ . The remaining  $\lambda$  consumers are currently not loyal shoppers for any brand (i.e., no seller is in their consideration set), but they become aware of a buying opportunity upon seeing an ad. In this case, they can buy from the only seller in their consideration set—the seller with the sponsored link or advertisement.<sup>11</sup>

This alternative interpretation in terms of endogenous consideration sets requires the platform to hold an (arbitrarily small) informational advantage relative to the

<sup>11</sup>Ronayne (2021) considers fixed-sample search with homogeneous sellers and a price-comparison website. The price comparison website generates “endogenous captivity” as in our model because, in equilibrium, consumers see the best price on the site and have no need to search elsewhere. Mekonnen, Murra-Anton, and Pakzad-Hurson (2023) consider the impact of exogenous information at each stage of a consumer's sequential search process, while Chen (2022) studies the evolution of a consumers' consideration set under the equilibrium advertising levels.



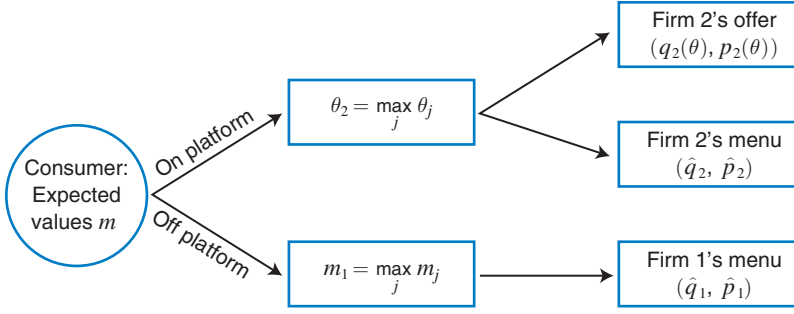


FIGURE 2. POSSIBLE CONSUMER CONSIDERATION SETS

consumers. In Section V, we show that, without an informational advantage, the platform does not control the consumers' outside options. Instead, the consumers' expected value fully determines which seller they would visit off the platform.

### C. Showrooming

The platform generates surplus by matching each  $\theta$  to the product with quality  $q_j(\theta)$  that generates the largest match value. As information is symmetric between the consumer and the selected seller, the seller can extract a substantial share of the created social surplus. To wit, the extraction of the surplus does not occur through personalized price discrimination but through product steering (thus a form of second-degree price discrimination). The only limit on surplus extraction by the advertising seller is given by the “showrooming constraint,” which is a necessary condition for seller  $j$  to make a sale on the platform:

$$(12) \quad \theta_j q_j(\theta) - p_j(\theta) \geq \max_{m_j} \{ \theta_j \hat{q}_j(m_j) - \hat{p}_j(m_j) \} \text{ for all } \theta_j.$$

Because seller  $j$  offers an incentive-compatible menu off the platform, each on-platform consumer would also report their value truthfully if shopping off platform. By Proposition 2, we know the consumer chooses between two products by the same seller, and the showrooming constraint (12) reduces to

$$(13) \quad U_j(\theta) \triangleq \theta_j q_j(\theta) - p_j(\theta) \geq \theta_j \hat{q}_j(\theta_j) - \hat{p}_j(\theta_j) \triangleq \hat{U}_j(\theta).$$

The showrooming constraint prevents the selected seller from extracting the entire surplus of the on-platform consumers. Because the on-platform transaction takes place under symmetric information, it is optimal for each seller to offer a single product to each consumer  $\theta$  at the socially efficient quality level

$$q_j^*(\theta) = \theta_j.$$

The socially efficient quality provision maximizes both the profit from the ad and the probability of being chosen by the platform. Similarly, it is optimal for each

seller to offer the consumer a discount that satisfies the showrooming constraint (13) with equality. Thus, despite the flexibility awarded by the platform, the quality and utility offered on platform by seller  $j$  are a function of  $\theta_j$  only.

Finally, if seller  $j$  offers the off-platform menu  $(\hat{p}_j, \hat{q}_j)$  with the associated rent function  $\hat{U}_j$ , the on-platform profit from a consumer with value profile  $\theta$  is given by

$$(14) \quad \pi_j(\theta, \hat{U}_j) = \begin{cases} \frac{\theta_j^2}{2} - \hat{U}_j(\theta_j), & \text{if } \theta_j > \max_{k \neq j} \theta_k; \\ 0, & \text{otherwise.} \end{cases}$$

#### IV. Equilibrium Product Lines

We now characterize the symmetric equilibrium menus off the platform and trace their implications for on-platform quantities and prices. We can then analyze the expected consumer surplus on and off the platform. Finally, we establish that the socially efficient steering mechanism is the revenue-maximizing managed campaign for the platform.

By Proposition 2, in any symmetric equilibrium of our model, no consumer (off platform or on platform) searches past the first seller on the equilibrium path. Combining the off-platform profit with the on-platform profit (14), each seller solves the following problem:

$$(15) \quad \Pi_j^* = \max_{\hat{q}_j, \hat{U}_j} \left\{ (1 - \lambda) \int_{\underline{\theta}}^{\bar{\theta}} \left[ m_j \hat{q}_j(m_j) - \frac{\hat{q}_j(m_j)^2}{2} - \hat{U}_j(m_j) \right] G^{J-1}(m_j) g(m_j) dm_j \right. \\ \left. + \lambda \int_{\underline{\theta}}^{\bar{\theta}} \left[ \frac{\theta_j^2}{2} - \hat{U}_j(\theta_j) \right] F^{J-1}(\theta_j) f(\theta_j) d\theta_j \right\},$$

subject to

$$(IR) \quad \hat{U}_j(m_j) \geq 0,$$

$$(IC) \quad \hat{U}_j'(m_j) = \hat{q}_j(m_j).$$

We observe that the objective function of each seller takes into account the competition across the seller's own sales channels. Each seller  $j$  generates an expected value  $m_j$  for the consumer off the platform with density  $g(m_j)$  but makes a sale only if  $j$  has a higher expected value than the remaining  $J - 1$  sellers, which happens with probability  $G^{J-1}(m_j)$ . A similar expression involving  $f(\theta_j)$  and  $F^{J-1}(\theta_j)$  holds for the on-platform revenue. Thus, the above pricing and revenue formulas take into account the competition among the  $J$  sellers by taking expectation with respect to the highest-order statistics. The resulting formula for the virtual utility of the off-platform consumer reflects the influence of the competition on as well as off the platform:

$$(16) \quad \theta_j - \frac{1 - [\lambda F^J(\theta_j) + (1 - \lambda) G^J(\theta_j)]}{(1 - \lambda) J G^{J-1}(\theta_j) g(\theta_j)}.$$

The solution of the seller's relaxed maximization problem (15) leads to a monotone allocation  $\hat{q}_j$  with respect to the value  $\theta_j$  if the virtual utility (16) is monotone in  $\theta_j$  whenever the virtual utility is nonnegative. In the following characterization of the equilibrium menus, we shall directly assume the monotonicity of the virtual utility expression (16).

Furthermore, because we have assumed  $m_j \in [\underline{\theta}, \bar{\theta}]$ , we state all our results with  $\theta_j$  as the argument, letting the distributions  $F$  and  $G$  indicate whether we refer to on- or off-platform variables. Maximizing (15) over rent and quality functions  $\hat{U}_j$  and  $\hat{q}_j$ , we obtain the following characterization of the optimal menus.

**PROPOSITION 3** (Equilibrium Menus with Efficient Steering): *Consider the socially efficient steering policy  $\sigma^*$ . Any symmetric equilibrium with full seller participation yields the following menus on and off the platform.*

(i) *The consumer's information rents are identical on and off platform:*

$$(17) \quad U_j^*(\theta_j) = \hat{U}_j^*(\theta_j) = \int_{\underline{\theta}}^{\theta_j} \hat{q}_j^*(x) dx.$$

(ii) *Assuming the monotonicity of the virtual utility (16), the quality levels are given by*

$$(18) \quad q_j^*(\theta_j) = \theta_j,$$

$$(19) \quad \hat{q}_j^*(\theta_j) = \max \left\{ 0, \theta_j - \frac{1 - [\lambda F^J(\theta_j) + (1 - \lambda) G^J(\theta_j)]}{(1 - \lambda) J G^{J-1}(\theta_j) g(\theta_j)} \right\}.$$

As in Myerson (1981), we could provide sufficient conditions in terms of  $F$  and  $G$  for the virtual utility (16) to be monotone. However, as the monotonicity of the virtual utility is not material for our main result—the revenue optimality of the socially efficient steering policy—we do not pursue this direction. If the virtual utility were not monotone, then the equilibrium quality schedule (19) off the platform would be given by the “ironed” version of (19), without affecting the efficient provision (18) on the platform.

The equilibrium quality provision on and off the platform has several intuitive properties. First, the efficient quality is sold to each consumer  $i$  on the platform, on the basis of their favorite seller, i.e.,  $\max_j \{\theta_j\}$ . Second, matching is inefficient off the platform because it is based on imperfect information, i.e., on expected value  $m$  instead of value  $\theta$ .

The consumer's private information off the platform requires the sellers to resolve the efficiency versus rent extraction trade-off. The information rents of each value  $m_j$  are as usual increasing in the quality level provided to all lower values. To resolve this trade-off, each seller  $j$  could offer the Mussa and Rosen (1978) tariff for the distribution of off-platform consumer values  $G^J(m_j)$ , which is the distribution of the highest-order statistic out of  $J$  variables  $m_j$ . However, any information rent  $\hat{U}(m)$  provided to the off-platform consumers has an additional shadow cost: it makes buying off platform more attractive for the on-platform consumers, too. As

we saw, by leaving positive rents for the consumers off the platform, each seller must also provide rents on the platform. Indeed, for all  $\theta_j > \underline{\theta}$ ,

$$(20) \quad U_j^*(\theta_j) = \hat{U}_j^*(\theta_j) > 0 \quad \text{iff} \quad \hat{q}_j^*(\theta_j) > 0.$$

Conversely, by limiting the off-platform rents, the seller is able to capture a greater share of the efficient social surplus that personalized on-platform offers generate.

Because of the shadow cost of showrooming, the off-platform quality schedule  $\hat{q}$  is further distorted downward. In particular, whenever  $\hat{q}_j^*(\theta_j) > 0$ , the equilibrium off-platform qualities (19) in Proposition 3 can be written as

$$(21) \quad \hat{q}_j^*(\theta_j) = \underbrace{\theta_j - \frac{1 - G^J(\theta_j)}{JG^{J-1}(\theta_j)g(\theta_j)}}_{\text{Mussa and Rosen (1978) quality}} - \frac{\lambda}{1 - \lambda} \frac{1 - F^J(\theta_j)}{JG^{J-1}(\theta_j)g(\theta_j)}.$$

The first two terms identify the optimal quality level for the distribution of values  $G^J(\theta_j)$ . The last term captures the intuition that any rent given off-platform to value  $\theta_j$  must also be given to all higher values *on* the platform. Note that, when  $\lambda$  is high, the shadow cost of these rents is so large that the seller forgoes off-platform sales to consumer  $\theta_j$  altogether.

Figure 3 displays the equilibrium off-platform quality schedule, the socially efficient allocation, and the monopoly benchmark of Mussa and Rosen (1978). The consumer's expected values  $m_j$  are uniformly distributed, and their values  $\theta_j$  follow a Beta distribution.<sup>12</sup>

The formulation of the optimal off-platform menu (21) allows us to establish several intuitive properties of the equilibrium. First, each value  $\theta$  receives a higher quality level, namely the socially efficient allocation, and pays a higher price on the platform than off the platform. However, while each value receives a better product at a higher price, each quality level  $q$  is sold at a lower price on the platform. Thus, let us define the equilibrium price for a given quality  $q$  on and off platform as  $p(q)$  and  $\hat{p}(q)$ , respectively. We then find that  $p(q) \leq \hat{p}(q)$  for all  $q$ . In other words, each seller is forced to introduce “on-platform only” discounts to prevent consumers from showrooming.

Figure 4 displays the nonlinear pricing schedules, namely the price  $p_j(q_j)$  for every offered quality  $q_j$  under the same parameters as in Figure 3. Note that for a set of low values, namely those below 0.8, the nonlinear tariff is equal to the gross surplus generated by the efficient quality (i.e.,  $p_j(q_j) = q_j^2$ ). By contrast, values above 0.8 receive a positive rent off platform, and hence, on the platform the price is below the gross surplus.

<sup>12</sup>Because quality provision is efficient for all  $\theta_j$  on platform but only efficient “at the top” off platform, the quality range is weakly larger off than on platform; i.e., downward distortions extend the quality range to the bottom. The quality ranges in Figure 3 are identical only because types  $\theta_j$  are supported on  $[0, 1]$ .

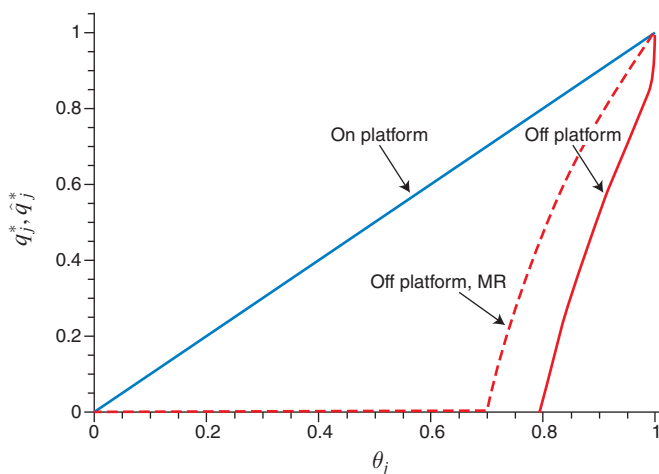


FIGURE 3. MATCH VALUES AND QUALITIES ON PLATFORM VERSUS OFF PLATFORM,  
 $\lambda = 1/2$ ,  $J = 5$ ,  $G(m_j) = m_j$ ,  $F(\theta_j) = \text{Beta}(\theta_j, 1/4, 1/4)$

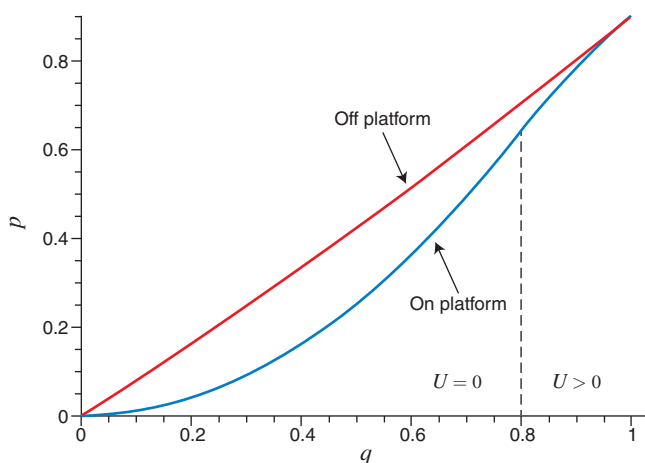


FIGURE 4. NONLINEAR TARIFFS: EVERY VARIETY  $q$  IS SOLD AT A LOWER PRICE ON PLATFORM

*Consumer Surplus.*—Proposition 3 implies that, on aggregate, consumer surplus is strictly larger on the platform than off the platform. Indeed, we have

$$\hat{U}_j^*(\theta_j) = U_j^*(\theta_j)$$

for each  $\theta_j$ , but we also know that  $F \succ G$ . This implies

$$(22) \quad E_{F'}[U_j^*(\theta_j)] > E_{G'}[\hat{U}_j^*(\theta_j)]$$

because the highest-order statistics satisfy  $E_{F'}[\theta_j] > E_{G'}[\theta_j]$  and incentive compatibility requires the function  $\hat{U}_j^*$  to be increasing and convex. Thus, at the equilibrium prices, every consumer would rather be on the platform (ex ante) than off the

platform. A stronger result is that, holding prices fixed, the consumer would like the platform to have as precise information as possible about their value, which enables better matching of products to preferences. However, the consumer does not necessarily benefit from the presence of the platform *in equilibrium*.

*Comparative Statics.*—Section A in the online Appendix examines the comparative statics of the optimal menus with respect to the platform size and the number of sellers. In particular, Proposition 10 shows that the equilibrium utility of every realized consumer value  $\theta$  weakly decreases as  $\lambda$  increases. As the fraction of on-platform consumers grows, the distortions in off-platform quality become more severe because deterring showrooming becomes more important for sellers. This drives off-platform information rents down and drives on-platform prices up.

An *ex ante* comparison, which is stated in expectation over all value profiles, must also account for the positive welfare effect described of shifting consumers from the off-platform to the on-platform market, as captured in (22). However, as  $\lambda \rightarrow 1$ , the platform captures the entire (first best) social surplus it creates, and rents vanish for all consumers' value profiles.

*Platform Revenue.*—To examine the implications for the sellers' profit and the platform's revenue, we characterize the advertising budgets the platform can demand in equilibrium under a given steering policy.

Consider the subgame following the platform's announcement of  $\sigma$  and  $t$ . Define each seller  $j$ 's outside option  $\bar{\Pi}_j(\sigma)$  as the profit that seller  $j$  can obtain if they do not participate in a managed campaign with selection rule  $\sigma$ . Similarly, define each seller's equilibrium profits (gross of the advertising budget) as  $\Pi_j^*(\sigma)$ . Because sellers are homogeneous *ex ante*, the platform can then request an advertising budget equal to

$$t^*(\sigma) \triangleq \Pi_j^*(\sigma) - \bar{\Pi}_j(\sigma).$$

Now consider the efficient selection rule  $\sigma^*$  we have examined so far. In the efficient steering managed campaign, consumers follow the platform's recommendation on and off the equilibrium path (Proposition 2). Each seller's outside option then consists of the profit level achievable from the off-platform consumers only, i.e.,

$$(23) \quad \bar{\Pi}_j(\sigma^*) = \max_{\hat{q}, \hat{U}} (1 - \lambda) \int_{\underline{\theta}}^{\bar{\theta}} \left[ \theta_j \hat{q}(\theta_j) - \frac{\hat{q}(\theta_j)^2}{2} - \hat{U}(\theta_j) \right] G^{J-1}(\theta_j) dG(\theta_j).$$

Given the equilibrium profit levels  $\Pi_j^*(\sigma^*)$  defined in (15), the platform then requests the following advertising budget:

$$(24) \quad t^*(\sigma^*) = \Pi_j^*(\sigma^*) - \bar{\Pi}_j(\sigma^*).$$

Another way to interpret the advertising budget is the following: the sellers are willing to give up all their on-platform profits to participate, but they must be compensated for distorting their off-platform menus to discourage showrooming.



We now consider the platform's problem when announcing a selection rule, i.e.,

$$\max_{\sigma} \{t^*(\sigma)\},$$

and we show that the efficient-steering policy solves this problem.<sup>13</sup>

**PROPOSITION 4 (Optimality of Efficient Steering):** *The managed campaign with efficient steering policy  $\sigma^*$  and associated budget  $t^*(\sigma^*)$  yields a unique symmetric subgame equilibrium. In this equilibrium, all sellers participate, and the platform obtains the profit level*

$$Jt^*(\sigma^*) = J[\Pi_j^*(\sigma^*) - \bar{\Pi}_j(\sigma^*)].$$

*The managed campaign with efficient steering policy  $\sigma^*$  maximizes the platform's profit among all managed campaign mechanisms.*

The proof of this result establishes that the efficient-steering managed campaign attains an exogenous upper bound on the platform's profit, taken over all selection rules. In particular, the equilibrium profits of the sellers,  $\Pi_j^*(\sigma^*)$ , coincide with the vertically integrated benchmark where the platform controls all sellers' menus on both sales channels. Thus, this mechanism maximizes the sellers' and the platform's joint profits across the on- and off-platform markets. Moreover, the seller's outside option,  $\bar{\Pi}_j(\sigma)$ , is bounded from below, for all managed campaigns  $\sigma$ , by the profits  $\bar{\Pi}_j(\sigma^*)$  a seller can obtain by posting the optimal menu for the off-platform consumers only. Because the advertising budget  $t^*(\sigma^*)$  extracts each seller's surplus in excess of this exogenous outside option, this mechanism maximizes the platform's revenue.<sup>14</sup>

The efficient-steering managed campaign relies only on advertising and steering through sponsored products. In particular, the optimal revenue can be attained without levying commission or transaction fees on sellers or consumers. Thus, the optimal matching mechanism fits directly the digital platforms that generate the majority of their revenue from digital advertising, such as Google or Meta.

*Fixed versus Variable Payments.*—Our results illustrate how advertising platforms that run managed campaigns face very different incentives than retail platforms that charge variable sales fees. Section B in the online Appendix contrasts the optimal mechanism with advertising budgets with a mechanism that more directly reflects the revenue model of Amazon, which is predominantly based on sales commissions. Proposition 13 shows that, under positive sales commissions, sellers want to encourage some high-value consumers to showroom and buy off platform, so to

<sup>13</sup>Consistent with this result, Bergemann, Bonatti, and Gan (2022) show that the optimal managed campaign improves the platform's revenue relative to running auctions for advertising to each individual consumer. Furthermore, the experimental analysis of Decarolis et al. (2023) shows that, when a platform provides less detailed information to the bidders' algorithms, its revenues are "substantially and persistently higher."

<sup>14</sup>In the presence of complete information by the platform, it is optimal to offer a single sponsored product. We suspect that in richer environments, where consumers have some independent private information, many sponsored links would optimally screen for this additional information.

avoid paying the platform's fees. The platform is then forced to impose price-parity clauses (also known as most-favored-nation clauses) that require each seller to always offer the lowest price on platform (i.e., to post menus as in Figure 4). In contrast, the lump-sum payments (i.e., advertising budgets) in our optimal mechanism incentivize sellers to prevent showrooming without the need for additional constraints.

## V. Value of Information and Privacy

In this section, we explore the platform's bargaining power by examining the role of its informational advantage and the implications of privacy policies. In Section VA, we remove the information advantage of the platform. Instead, we assume that every on-platform consumer learns their entire value profile  $\theta$ , not just their value for the sponsored seller. One possible reason for this could be that reviews and recommendations are available on the platform and online more generally.

Next, in Section VB, we examine the role of price information. We consider the provision of organic search links by the platform that enable consumers to learn about all off-platform prices and products.

In Section VC, we introduce privacy policies that safeguard the consumers' information from the sellers. We consider cohort-based privacy protection where the platform informs the sellers only about the consumer's ranking of the sellers, while disclosing the exact value for the sponsored seller to the consumer only. Consequently, the platform targets ads at the level of a *cohort* of consumers, and each consumer within a cohort has the same preference ranking over the  $J$  sellers.<sup>15</sup>

The results in these extensions paint a consistent picture of the platform's value of information: removing the platform's information advantage about consumer values reduces its revenue but does not change the equilibrium products and prices (Proposition 5); providing the consumers with off-platform price information fosters seller competition, increases quality and rents, and reduces platform revenues; and implicitly reducing the platform's information advantage relative to the sellers (i.e., limiting its ability to target sponsored content) also benefits consumers at the expense of platform revenue.

We conclude this section by examining whether revealing the full value for the sponsored seller to the consumer maximizes the platform's revenue. Thus, in Section VD, we introduce information design in our managed campaign mechanism and provide conditions under which full or partial information revelation is optimal.

### A. Symmetric Information

To assess the value of the platform's information advantage, we now assume all consumers who visit the platform learn their *entire value profile*  $\theta$  (i.e., not just their value for the sponsored seller). The consumers off the platform remain imperfectly informed with expected value profile  $m$ . In Proposition 5, we establish that the ensuing symmetric information limits the platform's ability to steer the consumers'

<sup>15</sup>This is in line with the recent Google Privacy Sandbox proposals to replace third-party cookies.

search behavior and reduces the advertising budget the platform can request from the sellers.

**PROPOSITION 5 (Symmetric Information):** *With complete information about  $\theta$  for all on-platform consumers, the equilibrium quality levels on and off the platform remain as in Proposition 3. But the equilibrium advertising budget  $t^*$  is strictly lower relative to when the platform has exclusive information about  $\theta$ .*

In equilibrium, both on- and off-platform consumers have the same information as in Section IV, where the platform has initially exclusive information about  $\theta$  (henceforth, the “baseline” model). Thus, any seller who participates in the managed campaign mechanism offers the optimal menu in Proposition 3. Facing informed consumers, however, changes the seller’s value of turning down the platform’s offer because consumers who know their values visit their favorite seller off platform regardless of the identity of the sponsored seller on the platform. Suppose a consumer sees a product by a seller that they did not expect on the platform. Under our symmetry refinement, this consumer continues to believe that all sellers offer symmetric menus off the platform.

Therefore, each seller  $j$  can choose not to participate in the managed campaign and poach any consumer to whom they offer the highest value:  $\theta_j = \max_k \theta_k$ . Seller  $j$  can achieve this by offering the consumers an off-platform information rent  $\hat{U}_j(\theta_j)$  above the level  $\hat{U}_j^*(\theta_j)$  in (17), which is offered by the competitors in equilibrium. When contemplating such a menu, the deviating seller  $j$  solves the following problem:

$$(25) \quad \hat{\Pi} \triangleq \max_{\hat{q}, \hat{U}} \int_{\underline{\theta}}^{\bar{\theta}} \left[ \theta_j \hat{q}(\theta_j) - \frac{\hat{q}(\theta_j)^2}{2} - \hat{U}(\theta_j) \right] \left[ (1 - \lambda) G^{J-1}(\theta_j) g(\theta_j) + \lambda F^{J-1}(\theta_j) f(\theta_j) \right] d\theta_j$$

subject to

$$(26) \quad \hat{U}(\theta_j) \geq \hat{U}_j^*(\theta_j).$$

The equilibrium rent function (17) in the baseline model satisfies the constraint (26) and yields a strictly larger profit. Therefore, the sellers’ outside option with known values exceeds the outside option  $\bar{\Pi}$  of the baseline model characterized in (23).

The deviating seller can do even better by offering the optimal menu of products when consumer values are distributed according to the mixture  $(1 - \lambda) G^J + \lambda F^J$ . These quality levels are given by

$$(27) \quad \hat{q}(\theta_j) = \max \left\{ 0, \theta_j - \frac{1 - \lambda F^J(\theta_j) - (1 - \lambda) G^J(\theta_j)}{\lambda J F^{J-1}(\theta_j) f(\theta_j) + (1 - \lambda) J G^{J-1}(\theta_j) g(\theta_j)} \right\}.$$

The equilibrium quality in (27) is larger for every value than the equilibrium  $\hat{q}_j^*(\theta_j)$  in (19) and yields higher utility to the consumers. Thus, constraint (26) does not bind in the optimal deviation—the best off-platform menu for seller  $j$  offers a higher utility level to  $j$ 's favorite consumers than all other sellers' menus.<sup>16</sup>

To summarize, Proposition 5 shows that the equilibrium advertising budgets are qualitatively different when the on-platform consumers know their values from when they learn their values through the platform's information. After all, the platform loses the ability to steer the consumer. In the absence of additional information, the platform cannot grant monopoly power to any seller by displaying their advertisement and recommending their products. Without additional information, each consumer evaluates the different products independently of the recommendation implicit in the ad. In turn, the value to a seller of showing an advertisement decreases, as does the willingness to pay for the platform's services.

To quantify the value of the platform's steering power, fix the value distribution  $F$  and consider the distribution  $G$  of expectations generated by an imperfectly informative signal that each (on- and off-platform) consumer observes about their value. Denote by  $t^*(G)$  the equilibrium advertising budgets in (24) under distribution  $G$ . Now let the consumer's signal become arbitrarily precise, so that the distribution  $G$  of expectations converges to the distribution  $F$  of true values. The equilibrium menu for the limit case is obtained by setting  $G = F$  in (19). Proposition 5 and condition (27) then imply the following observation.

**COROLLARY 1** (Value of Additional Information): *For all  $J > 1$ , the platform gains strictly positive profits from any information advantage:*

$$\lim_{G \rightarrow F} t^*(G) > t^*(F).$$

Corollary 1 has important implications for a platform's choice of information design, to which we turn in Section VD. In particular, the equilibrium advertising budgets jump up as soon as the platform has any informational advantage relative to the consumers. This suggests that some degree of information asymmetry—revealing some additional information to consumers—is always part of the optimal design.

### B. Organic Links

In the equilibrium of our baseline model, the consumer chooses the advertised seller that offers the highest value, and only considers that seller's on- and off-platform offers. However, in practice, platforms may also display “organic links” that provide additional, free information to consumers. We extend our model to consider a scenario where the platform shows all off-platform prices to consumers. In this setting, on-platform consumers do not incur a search cost and can buy from any seller without incurring search costs.

<sup>16</sup>Note that the deviating seller cannot attract any consumer who values a competitor's products more than their own. This is because those consumers still face search costs off platform and would not learn that the deviating seller has lowered their prices.

Sellers still advertise the socially efficient product varieties and set prices to make the showrooming constraints bind. The platform assigns the sponsored link to each consumer's favorite seller, but the off-platform menus can now affect market shares on the platform. Sellers can attract some of their competitors' on-platform consumers by offering lower prices off the platform. These consumers would not learn about the lower prices without the presence of organic links.

To calculate the sellers' market shares of on-platform consumers, we consider the off-platform information rents  $\hat{U}_j(\theta_j)$ . The outside option of the on-platform consumer  $\theta$  is given by  $\max_j \{ \hat{U}_j(\theta_j) \}$ . For a symmetric strategy profile by all sellers  $k \neq j$ , and for each value  $\theta_j$ , we define the indifferent value  $\theta_k^*(\theta_j)$  as

$$(28) \quad \hat{U}_k(\theta_k^*(\theta_j)) = \hat{U}_j(\theta_j).$$

With  $\theta_k^* = \theta_k^*(\theta_j)$  defined as in (28), seller  $j$ 's best-response problem is given by

$$(29) \quad \max_{\hat{q}, \hat{U}} \left\{ (1 - \lambda) \int_{\underline{\theta}}^{\bar{\theta}} \underbrace{\left[ \theta_j \hat{q}(\theta_j) - \frac{\hat{q}(\theta_j)^2}{2} - \hat{U}(\theta_j) \right]}_{\text{off-platform sales}} G^{J-1}(\theta_j) g(\theta_j) d\theta_j \right. \\ + \lambda \int_{\underline{\theta}}^{\bar{\theta}} \underbrace{\left[ \frac{\theta_j^2}{2} - \hat{U}(\theta_j) \right]}_{\text{on-platform sales}} \min \{ F^{J-1}(\theta_k^*(\theta_j)), F^{J-1}(\theta_j) \} f(\theta_j) d\theta_j \\ + \lambda \int_{\underline{\theta}}^{\bar{\theta}} \underbrace{\left[ \theta_j \hat{q}(\theta_j) - \frac{\hat{q}(\theta_j)^2}{2} - \hat{U}(\theta_j) \right]}_{\text{off-platform sales}} \max \{ 0, F^{J-1}(\theta_k^*(\theta_j)) \\ \left. - F^{J-1}(\theta_j) \} f(\theta_j) d\theta_j \right\}.$$

The first term in (29) captures the off-platform consumers. The second term captures the sales to on-platform consumers for which seller  $j$  offers both the highest utility level  $\hat{U}_j$  and the highest marginal value  $\theta_j$ . The third term, whenever positive, captures on-platform consumers with  $\theta_j < \max_{k \neq j} \theta_k$  to whom seller  $j$  nonetheless offers the highest utility level  $\hat{U}_j$ . Seller  $j$  is not advertised to these consumers, who instead showroom and buy from seller  $j$  off the platform.

Seller  $j$ 's problem can therefore be restated as follows. Undercutting the other sellers (so that  $\theta_k^*(\theta_j) > \theta_j$ ) yields some additional on-platform consumers who buy off platform. Conversely, raising prices above the other sellers' (so that  $\theta_k^*(\theta_j) < \theta_j$ ) causes seller  $j$  to lose some consumers who would otherwise buy on platform (i.e., a higher-quality product at a higher price, relative to off-platform sales). Therefore, relative to the baseline model with a sponsored link only, each seller  $j$  has an incentive to raise  $\theta_k^*$  through a higher  $\hat{U}_j$ . This incentive, which is entirely due to organic links, explains the result in Proposition 6, where we compare the game with organic links to the baseline setting.

**PROPOSITION 6 (Equilibrium with Organic Links):**

- (i) *The equilibrium quality and utility levels  $\hat{q}_j^*(\theta_j)$  and  $\hat{U}_j^*(\theta_j)$  are weakly higher for all  $\theta_j$  with organic links than without.*
- (ii) *The sellers' profits are lower and their outside options are higher with organic links than without.*

To establish this result, we consider the symmetric equilibria of the subgame following the platform's announcement of an advertising budget  $t$  and a selection rule  $\sigma$ . The symmetric equilibrium quality levels of this game can be characterized through a system of differential equations, as in Bonatti (2011). We show that competition among the sellers is fiercer in any such equilibrium. In particular, quality and utility levels are higher and sellers' profits are lower than without organic links. Conversely, the sellers' outside options in any continuation equilibrium are higher than in the baseline model, and the platform demands a lower advertising budget.

Intuitively, the presence of organic information benefits consumers but reduces the platform's ability to restrain competition and extract surplus from sellers. In a symmetric equilibrium, each seller's market segment consists of all consumers who like the products the most. These market shares, unlike the baseline case, are endogenous to the choice of  $\hat{U}_j$ . Because the off-platform menus can affect the on-platform market shares, offering higher rents to consumers has an additional benefit. The equilibrium utility and quality levels are then higher than without organic links, the on-path gross profits of the sellers are lower, and consumer surplus is higher.

However, the sellers' profits net of the advertising budget are equal to the value of their outside option. With organic links, any seller can respond to competitors' prices without participating in the mechanism. Let  $\theta_k^*$  be given by (28), with  $\hat{U}_k = \hat{U}_j^*$ . Because all of the sales necessarily happen off the platform, a deviating seller  $j$  can then obtain profit of

$$(30) \quad \tilde{\Pi}_j \triangleq \max_{\hat{q}, \hat{U}} \int_{\underline{\theta}}^{\bar{\theta}} \left[ \theta_j \hat{q}(\theta_j) - \frac{\hat{q}(\theta_j)^2}{2} - \hat{U}(\theta_j) \right] \left[ (1 - \lambda) G^{J-1}(\theta_j) g(\theta_j) + \lambda F^{J-1}(\theta_k^*(\theta_j)) f(\theta_j) \right] d\theta_j.$$

Unlike in the baseline model, each deviating seller has the opportunity to win over some (but not necessarily all) on-platform consumers for which  $\theta_j \geq \max_{k \neq j} \theta_k$ . The outside option  $\tilde{\Pi}_j$  in (30) then exceeds the value  $\bar{\Pi}_j$  defined in (23). In other words, the sellers' outside options are higher with organic links than without, and the equilibrium advertising budgets are correspondingly lower.

Finally, recall that with symmetric information and no organic links, the deviating seller wins *all* on-platform consumers for which  $\theta_j \geq \max_{k \neq j} \theta_k$ . Thus, the outside option  $\hat{\Pi}_j$  defined in (26) is even higher than  $\tilde{\Pi}_j$  in (30). Because the equilibrium menus with organic links do not change if consumers know their values, it is possible that the platform will profitably raise the requested advertising budget by showing organic links if consumers are already fully informed about their values.

### C. Privacy Protection and Cohort-Based Advertising

We now investigate how data governance, the rules governing how consumer data are deployed by the digital platform, influences the distribution of social surplus. In particular, we analyze how a more limited use of the consumer data may influence the market outcome. Up to this point, we have not limited the platform's ability to share information about the values of the consumers,  $\theta$ , with the sellers. In reality, the extent of data sharing may be restricted by regulation or design choices made by the platform.

We specifically consider cohort-based advertising, where anonymized groups of people with shared characteristics or online behavior replace individual tracking. In the context of our model, we define a cohort of consumers as those consumers who share the same ordinal ranking of the sellers. Thus, a cohort shares the same experience ("every user who bought Nike shoes") without revealing any cardinal information about its members in terms of their willingness to pay.<sup>17</sup>

Thus, we now assume that the value of the advertised product  $\theta_j$  is shared with consumers on the platform but not with sellers. Sellers are only allowed to base their offers on the *ranking* of consumer valuations  $\theta_j$  within a *cohort* of consumers, where each consumer ranks the  $J$  sellers in the same way.

Despite this change, efficient matching of sellers and consumers remains feasible. However, consumers on the platform still have some private information about their preferences. Unlike the baseline case where each seller could make personalized offers to consumers, cohort-based ads mean that seller  $j$  only knows the distribution of consumer values based on the order statistics implied by their cohort. Consequently, each seller must screen consumers both on and off the platform.

The symmetric equilibrium menus under cohort-based ads are the solution to two linked screening problems. In the on-platform problem, the showrooming constraints act as value-dependent participation constraints. To solve this problem, we strengthen the ranking of the distributions  $F$  and  $G$  by assuming that the on-platform distribution  $F$  dominates the off-platform distribution  $G$  in the *likelihood-ratio order*, denoted  $F \succ_{lr} G$ . The distribution  $F$  dominates  $G$  in the likelihood-ratio order if  $g(\theta_j)/f(\theta_j)$  is decreasing in  $\theta_j$  (see Definition 1.C.1 in Shaked and Shanthikumar 1994). We define the (Myerson) virtual values for the two distributions as

$$\phi_F(\theta_j) \triangleq \theta_j - \frac{1 - F^J(\theta_j)}{JF^{J-1}(\theta_j)f(\theta_j)}, \quad \text{and} \quad \phi_G(\theta_j) \triangleq \theta_j - \frac{1 - G^J(\theta_j)}{JG^{J-1}(\theta_j)g(\theta_j)}.$$

We only require that the *likelihood-ratio order* is maintained over the range of values with a positive virtual utility under  $F$  and thus denote the corresponding ranking by  $F \succ_{lr+} G$ .

<sup>17</sup>The notion of cohort-based advertising gained prominence after Google announced that it would no longer support third-party cookies on its browser that facilitated the tracking of individual users. Google further suggested to create broad categories, cohorts, into which individual users would be classified following their online history (see Xiao and Karlin 2024). The cohort-based model of data sharing is often referred to as federated learning, as each local device maps or summarizes the local data into an assignment and only forwards the assignment of an individual to a cohort to the central server. This creates a federation of local devices that are coordinated by a central server (see MacMahan et al. 2017).



**PROPOSITION 7 (Cohort Targeting):** *If  $F^J \succ_{lr+} G^J$ , then in the unique symmetric equilibrium, each seller offers quality levels*

$$\hat{q}_j^*(\theta_j) = q_j^*(\theta_j) = \max \left\{ 0, \theta_j - \frac{1 - \lambda F^J(\theta_j) - (1 - \lambda) G^J(\theta_j)}{\lambda J F^{J-1}(\theta_j) f(\theta_j) + (1 - \lambda) J G^{J-1}(\theta_j) g(\theta_j)} \right\}.$$

Proposition 7 shows that if the distribution of the highest  $\theta_j$  dominates that of the highest  $m_j$  in likelihood ratio (over the relevant range), then each seller offers the same menu to consumers both on and off the platform. In this menu, the equilibrium quality schedule is the same as in (27), i.e., the Mussa and Rosen (1978) quality level for a mixture with weights  $(\lambda, 1 - \lambda)$  of the distributions of the highest-order statistics of  $\theta$  and  $m$ , respectively. Cohort-based ads thus yield higher quality provision off platform but lower quality on platform relative to the baseline model with full disclosure of the value  $\theta$ .

A critical implication of Proposition 7 is that all consumers receive higher information rents relative to the baseline setting because of the greater quality provision off the platform. Total surplus can also be higher because of greater off-platform quality, although on-platform quality is lower. Finally, as the sellers' outside options are unchanged relative to the baseline model, the equilibrium advertising budgets are unambiguously lower.

#### D. Information Design

In the analysis so far, we have assumed that the platform reveals the true value  $\theta_j$  of the sponsored brand  $j$  to the consumer. Proposition 4 has shown that the optimal mechanism is then to match each consumer with their favorite seller  $j^* = \operatorname{argmax} \theta_j$  according to their true preferences, which is the efficient managed campaign mechanism.

In this section, we investigate the optimal information design by the platform. We derive conditions under which full information revelation is approximately or exactly optimal and conditions for no information revelation to be optimal. We then focus on the case of uninformed off-platform consumers and analyze how the optimal information policy changes with the platform's size  $\lambda$  and the distributions of values  $F(\theta)$ . Finally, we discuss the information revealed by the matching mechanism when the platform does not provide full information about consumers' value for the sponsored seller.

We assume that the platform knows each consumer's value  $\theta$  and their expected value  $m$ . This assumption simplifies the analysis, but it is also a reasonable approximation since if the platform knows every consumer's true preferences, it may also have information about their past *experiences*.<sup>18</sup> For instance, the platform may have access to the consumer's cookies and browsing history, which would enable it to estimate the consumer's expected value.

In a managed campaign, each seller  $j$  maximizes total profit by choosing prices and product qualities given the platform's designed information. The platform, in

<sup>18</sup> See Liang and Madsen (2023) for a formal distinction.

turn, maximizes the sellers' profits by choosing the distribution of expected values. Advertising budgets then extract the sellers' willingness to pay for this information. Hence, we can think of the platform as designing both the information and the prices to maximize the sellers' profits.

However, revealing information to consumers presents a trade-off for the platform. On the one hand, better information improves the efficiency of matching between consumers and sellers and product varieties. On the other hand, better information increases the consumer's expected rent off the platform, tightens the showrooming constraint, and reduces the sellers' willingness to pay. To solve the platform's problem, we first show that it is without loss to focus on symmetric information structures.

**LEMMA 1 (Symmetric Information):** *The optimal information structure enables trade under symmetric (possibly partial) information on the platform.*

This result first appeared as Lemma 1 in Bergemann, Bonatti, and Gan (2022). The intuition in our model is that (i) holding off-platform menus fixed, the platform increases the sellers' profits by eliminating the consumers' private information and (ii) any private information signaled by the sellers to the consumers through their prices can be profitably revealed up front to the consumers.

*Large Platform.*—We begin by considering the limit case where  $\lambda = 1$ . In other words, a measure one of consumers shop on the platform, and hence, sellers have no reason to post off-platform menus. In this case, we show that the platform maximizes the advertising budgets by committing to the efficient managed campaign and by fully revealing  $\theta_j$  to both consumer  $\theta$  and to the sponsored seller  $j$ .

**PROPOSITION 8 (Large Platform):** *When  $\lambda = 1$ , for any number of sellers  $J$  and distributions  $F$  and  $G$ , it is optimal for the platform to match consumer  $\theta$  to the efficient seller  $j^*$  and to fully reveal  $\theta_j$ .*

When the platform becomes arbitrarily large ( $\lambda \rightarrow 1$ ), the rents off platform vanish, and the sponsored seller can appropriate the entire surplus it generates. The sponsored seller's profit under complete and symmetric information on each value  $\theta_j$  is then given by the first-best surplus  $\pi_j^*(\theta) = \theta_j^2/2$ . Because  $\pi_j^*(\cdot)$  is strictly convex, the platform-optimal information design reveals to each consumer their true value for the sponsored seller. Furthermore, by Proposition 4, it is optimal to match consumers and sellers efficiently when the platform reveals all the information.

*Zero Private Information.*—We now characterize the platform's optimal information policy in the special case where the off-platform consumers have no private information about their expected values. Thus, the distribution  $G$  places a unit mass on the expected value  $\mu \triangleq E_F[\theta_j]$  for all  $j$ ; thus,  $G(m_j) = \mathbf{1}_{\{m_j \geq \mu\}}$ .

Because the off-platform consumers have no private information, each seller  $j$  offers just one product of quality  $\hat{q}_j \in \mathbb{R}_+$  off platform at a price that extracts the consumer's expected willingness to pay, i.e.,  $\hat{p}_j = \mu \hat{q}_j$ . As we know from our baseline model, the seller's choice of off-platform quality  $\hat{q}_j$  directly controls the

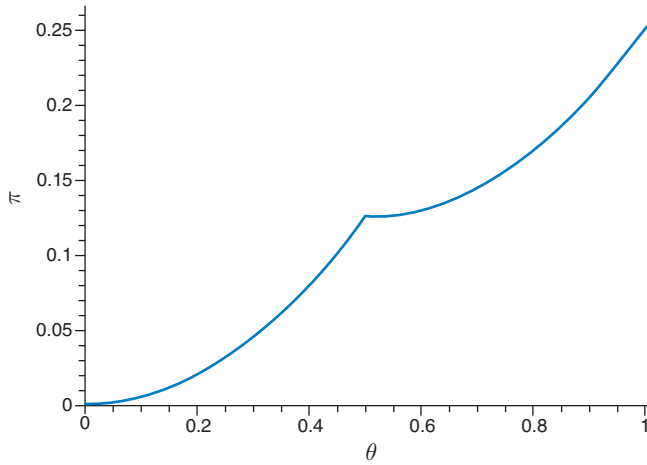


FIGURE 5. ON-PLATFORM PROFIT LEVELS

information rent of all on-platform consumers. In particular, a consumer with an expected value  $\theta_j$  (given the information revealed to them by the platform) obtains a rent

$$U(\theta_j, \hat{q}_j) = \max\{0, (\theta_j - \mu) \hat{q}_j\}$$

when buying from seller  $j$ . Consequently, seller  $j$ 's on-platform profits as a function of the realized value  $\theta_j$  are given by

$$(31) \quad \pi(\theta_j, \hat{q}_j) = \frac{\theta_j^2}{2} - U(\theta_j, \hat{q}_j).$$

Figure 5 illustrates the profit function  $\pi(\cdot, \hat{q}_j)$  for an example where  $\mu = 1/2$  and  $\hat{q}_j = 1/2$ . The seller extracts the entire willingness to pay of all on-platform values  $\theta_j \leq 1/2$  but leaves a rent to values  $\theta_j \geq 1/2$ , hence the downward kink.

As a first step toward characterizing the optimal information design, we establish that the platform shows each consumer an ad by their favorite seller and reveals an informative signal about their value  $\theta_j^*$ .

**LEMMA 2 (Efficient Steering):** *If consumers are uninformed, the efficient managed campaign matching mechanism is optimal for the platform.*

To gain intuition, observe that the seller's profit from value  $\theta_j$  is  $\pi(\theta_j, \hat{q}_j)$ , which is strictly increasing in  $\theta_j$ . Thus, the platform's payoff increases when the distribution of the underlying "state" (i.e.,  $\theta_j$ ) improves in the first-order stochastic sense. Because the distribution of the highest-order statistic  $F^J(\theta_j)$  first-order dominates the distribution of values  $\hat{F}$  that is induced by any other matching mechanism, the sender (i.e., the platform) chooses to design information about the consumer's highest value component  $\theta_j$ .

Therefore, we can write the platform's problem as

$$(32) \quad \max_{\hat{q}_j, \hat{F}} \left\{ (1 - \lambda) \left( \mu \hat{q}_j - \frac{\hat{q}_j^2}{2} \right) + \lambda \int_{\underline{\theta}}^{\bar{\theta}} \left( \frac{\theta_j^2}{2} - \max\{0, (\theta_j - \mu) \hat{q}_j\} \right) d\hat{F}(\theta_j) \right\}$$

subject to

$$F^J \succ \hat{F}.$$

In order to solve this problem, we adapt the toolkit of Dworzak and Martini (2019) for persuasion problems where the receiver's posterior mean is a sufficient statistic for their beliefs. We first fix an off-platform quality level  $\hat{q}$  and optimize over information structures. We then characterize the optimal quality level off platform.

**PROPOSITION 9** (Optimal Information Design): *Fix  $\hat{q}$  and suppose the off-platform consumers have zero private information.*

- (i) *There exist two thresholds  $x_1 \leq \mu \leq x_2$  such that the optimal distribution of posteriors  $\hat{F}^*(\theta_j)$  coincides with  $F^J(\theta_j)$  on  $[0, x_1]$  and  $[x_2, 1]$  and has an atom at  $\mu$ .*
- (ii) *The pair of optimal thresholds  $(x_1, x_2)$  are the unique solution to*

$$x_1 + 2\hat{q} = x_2 \quad E_{F^J}[\theta | x_1 \leq \theta \leq x_2] = \mu.$$

Thus, the platform matches consumers and sellers efficiently but does not enable efficient trade for all values. In particular, values closest to the mean of the marginal distribution  $\mu$  all receive the efficient quality *for the average value*. The pooling region allows the seller to optimally trade off higher on-platform profit with lower off-platform rents. Figure 6 illustrates the solution for the case of  $\lambda = 3/8$ , with  $F(\theta_j) = \theta_j$  and  $J = 2$ .

Having characterized the optimal information design for any choice of quality off the platform, we can compute the optimal  $\hat{q}^*$  from the sponsored seller's profit function.

$$\begin{aligned} \Pi(\hat{q}) = & \lambda \int_0^{x_1(\hat{q})} \frac{\theta_j^2}{2} dF^J(\theta_j) + \frac{\mu^2}{2} [F^J(x_2(\hat{q})) - F^J(x_1(\hat{q}))] \\ & + \lambda \int_{x_2(\hat{q})}^1 \left[ \frac{\theta_j^2}{2} - \hat{q}(\theta_j - \mu) \right] dF^J(\theta_j) + (1 - \lambda) \left( \mu \hat{q} - \frac{\hat{q}^2}{2} \right). \end{aligned}$$

It is then immediate to show that as  $\lambda \rightarrow 1$  (as in Proposition 8), the optimal  $\hat{q}^* \rightarrow 0$  and  $x_1, x_2 \rightarrow \mu$  so that the platform reveals the consumer's value with probability 1. In some special cases, the solution is in closed form and yields the conclusion of Proposition 8 even if  $\lambda$  is bounded away from 1.

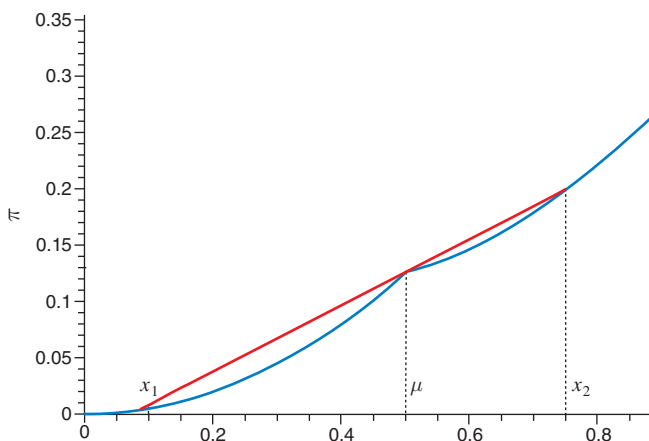


FIGURE 6. OPTIMAL SEPARATING AND POOLING INTERVALS

*Discussion.*—When  $\lambda \in (0, 1)$  and the distribution  $G$  of expected values  $m_j$  is not degenerate, the problem becomes significantly more complex. If the platform does not know the consumer's expected value, then it faces a persuasion problem where the receiver's private value is correlated with the state, unlike in Kolotilin et al. (2017). A potentially fruitful approach to this problem could be to focus attention to the case of *public persuasion*, i.e., to signal structures that do not condition on the consumer's expected value  $m$ .

## VI. Conclusion

We have developed a model of competition in the digital economy with heterogeneous consumer preferences and products. A digital platform serves as an intermediary between consumers and sellers, utilizing its superior information to form matches between them. The platform monetizes its information advantage by presenting consumers with their preferred products and generates revenue by selling access to their attention through managed advertising campaigns. However, the ability for sellers to showcase their products off the platform limits their ability to price discriminate on the platform and their willingness to pay for advertising. This force leads to higher prices on both sales channels as the platform's user base grows.

Our model is simplified, but it can be extended to consider differentiated products with varying on- and off-platform presences. This may introduce distortions in the managed campaign allocation of advertising space, as smaller sellers exploit higher margins and consumer search costs.<sup>19</sup> Our model also assumes “perfect steering,” but it can be expanded to incorporate multiple related advertisements to each consumer. Overall, our paper highlights the role of data in shaping competition and allocating surplus in the digital economy.

<sup>19</sup>The recent evidence in Mustri, Adjerid, and Acquisti (2022) is consistent with a mechanism like the one we outlined.

## APPENDIX

## PROOF OF PROPOSITION 1:

To derive the low-value consumer's optimal quality, we substitute the expression for the binding incentive compatibility constraint for the high-value consumer (8) into both the on-platform and off-platform profit in objective (7). Differentiating with respect to  $\hat{q}(\underline{\theta})$  yields the result in (9). ■

## PROOF OF PROPOSITION 2:

In any symmetric equilibrium, each consumer  $\theta$  with expected value  $m$  learns their true value  $\theta_j^*$  for the advertised seller  $j^*$  and believes that  $j^* = \operatorname{argmax}_j \theta_j$  with probability 1. Moreover, each consumer expects symmetric menus off platform and knows that the rent function  $\hat{U}_j(\theta_j)$  is strictly increasing in  $\theta_j$  for all  $j$ . Therefore, the consumer searches for the advertised seller's off-platform prices. She does not search any further and does not learn any other seller's prices. Because the menu off the platform is incentive compatible, it is sufficient for consumer  $\theta$  to compare the two items  $q_{j^*}(\theta_{j^*})$  and  $\hat{q}_{j^*}(\theta_{j^*})$ .

This holds both on and off the equilibrium path. Indeed, suppose a seller  $\hat{j}$  deviates and does not participate in the platform's mechanism. In this case, all consumers with  $\hat{j} = \operatorname{argmax}_j \theta_j$  are shown an advertisement by a different seller. These consumers are unable to detect seller  $\hat{j}$ 's deviation and hence search for the sponsored seller's menu only. ■

## PROOF OF PROPOSITION 3:

Seller  $j$ 's gross profit (i.e., after paying the platform's required advertising budget) can be written as

$$(33) \quad \max_{\hat{q}, \hat{U}} \left\{ (1 - \lambda) \int_{\underline{\theta}}^{\bar{\theta}} \left[ \theta_j \hat{q}(\theta_j) - \frac{\hat{q}(\theta_j)^2}{2} - \hat{U}(\theta_j) \right] G^{J-1}(\theta_j) dG(\theta_j) \right. \\ \left. + \lambda \int_{\underline{\theta}}^{\bar{\theta}} \left[ \frac{\theta_j^2}{2} - \hat{U}(\theta_j) \right] F^{J-1}(\theta) dF(\theta) \right\},$$

subject to

$$\hat{U}'(\theta_j) = \hat{q}(\theta_j), \quad \hat{U}(\theta_j) \geq 0 \quad \text{for all } \theta_j, \text{ and } \hat{q} \text{ nondecreasing.}$$

The necessary pointwise conditions for  $\hat{q}$  and  $\hat{U}$  can be obtained from the control problem with the associated Hamiltonian and costate variable  $\hat{\gamma}(\theta_j)$ :

$$H(\theta_j, \hat{q}, \hat{U}, \hat{\gamma}) = (1 - \lambda) \left[ \theta_j \hat{q}(\theta_j) - \frac{\hat{q}(\theta_j)^2}{2} - \hat{U}(\theta_j) \right] G^{J-1}(\theta_j) g(\theta_j) \\ + \lambda \left[ \frac{\theta_j^2}{2} - \hat{U}(\theta_j) \right] F^{J-1}(\theta_j) f(\theta_j) + \hat{\gamma}(\theta_j) \hat{q}(\theta_j).$$

At a symmetric equilibrium, the optimality conditions are given by

$$(34) \quad (1 - \lambda)[\theta_j - \hat{q}(\theta_j)] G^{J-1}(\theta_j)g(\theta_j) + \hat{\gamma}(\theta_j) = 0,$$

$$(35) \quad -(1 - \lambda) G^{J-1}(\theta_j)g(\theta_j) - \lambda F^{J-1}(\theta_j)f(\theta_j) + \hat{\gamma}'(\theta_j) = 0,$$

$$\hat{\gamma}(\bar{\theta}) = 0.$$

Integrating, we obtain

$$(36) \quad \hat{\gamma}(\theta_j) = \frac{1}{J}[(1 - \lambda) G^J(\theta_j) + \lambda F^J(\theta_j) - 1].$$

Therefore, the equilibrium quality level is given by

$$(37) \quad \hat{q}_j^*(\theta_j) = \theta_j - \frac{1 - (1 - \lambda) G^J(\theta_j) - \lambda F^J(\theta_j)}{(1 - \lambda) J G^{J-1}(\theta_j)g(\theta_j)}$$

if the right-hand side is nonnegative, and nil otherwise, as in (19). This ends the proof. ■

#### PROOF OF PROPOSITION 4:

- (i) The uniqueness of the symmetric equilibrium with seller participation follows from the strategic substitutability of the participation decisions. Indeed, suppose  $J - 1$  sellers turn down the platform's request of  $t^*(\sigma^*)$ . By joining the platform, seller  $j$  is at least as well-off as when all other sellers participate. Under the efficient steering policy, this seller would obtain a profit level strictly higher than  $\Pi_j^*(\sigma^*)$  defined in (15).
- (ii) We argue the optimality of the efficient steering policy in two steps. First, we show that the efficient steering policy attains an exogenous upper bound on the profits of the coalition of all sellers and the platform. Second, we show that it attains an exogenous lower bound on the sellers' net profits.

Consider first the platform-sellers coalition: the coalition's profits are maximized by matching each on-platform consumer  $\theta$  to seller  $j^* = \arg\max_j \theta_j$ ; by matching each off-platform consumer  $m$  to  $\hat{j} = \arg\max_j m_j$ ; and by maximizing each seller's profit with respect to the on-platform offers  $(q, U)$  and the off-platform menus  $(\hat{q}, \hat{U})$ , holding the seller's market segment fixed. The coalition's optimal on-platform offers and off-platform menus then solve problem (33). Therefore, coalition-optimal offers and menus coincide with the symmetric equilibrium outcome of the managed campaign mechanism with efficient steering.

Now consider the seller's profits net of the advertising budget. Suppose there exists a symmetric equilibrium of a mechanism under which a seller  $j$  earns strictly less than the profit level  $\bar{\Pi}_j$  defined in (23). This seller can forgo participation in the mechanism and post the optimal Mussa and Rosen (1978)



menu for type distribution  $G^J$  off platform instead. Any off-platform consumer with the highest expected value  $m_j$  for seller  $j$  searches for seller  $j$  first and observes this menu. Under symmetric beliefs, the consumer believes that all other sellers are also offering the same menu and will therefore not search further. This means seller  $j$  can secure the Mussa-Rosen profits on the segment of off-platform consumers who assign the highest expected value to product  $j$ . Because this profit level is given by  $\bar{\Pi}_j$  in (23), the managed campaign mechanism with efficient steering attains a lower bound on the sellers' net profits across all mechanisms. ■

#### PROOF OF PROPOSITION 5:

Suppose on-platform consumers know their value  $\theta$ . If seller  $j$  participates in the mechanism but offers an out-of-equilibrium menu off the platform, only consumers who search for seller  $j$  in equilibrium observe this deviation. Therefore, every seller that participates in the mechanism can do no better than to advertise the efficient quality levels and post the off-platform menus that solve (33). However, if seller  $j$  does not participate, it can match the competitors' information rents  $\hat{U}_{k \neq j}$  and attract all the consumers on the platform who value their products the most. (Under the symmetric beliefs refinement, these consumers search for seller  $j$ 's off-platform offer regardless of the ads shown to them by the platform.) Furthermore, if a nonparticipating seller  $j$  maximizes profit with respect to  $(\hat{q}, \hat{U})$  over the combined off- and on-platform market segments, it solves the problem in (26). This is a standard second-degree price discrimination problem where consumer values are distributed according to  $\lambda F^J + (1 - \lambda) G^J$ . The optimal quality provision in such a deviation is given by

$$(38) \quad \hat{q}(\theta_j) = \max \left\{ 0, \theta_j - \frac{1 - (1 - \lambda) G^J(\theta_j) - \lambda F^J(\theta_j)}{\lambda F^{J-1}(\theta_j) f(\theta_j) + (1 - \lambda) G^{J-1}(\theta_j) g(\theta_j)} \right\}.$$

Because the quality level  $\hat{q}$  in (38) is pointwise larger than  $\hat{q}_j^*$  in (37), the resulting information rent is correspondingly higher for each  $\theta_j$ . Thus, the deviating seller's optimal choice of menu yields an outside option  $\hat{\Pi}_j$  larger than  $\bar{\Pi}_j$  in (23). In addition, because the on-path profits are unchanged relative to the case of asymmetrically informed consumers, the advertising budget requested by the platform must decrease. ■

#### PROOF OF PROPOSITION 6:

We characterize all symmetric equilibria with full participation of the subgame following the platform's announcement of the required advertising budget. We first rewrite problem (29) as follows:

$$\begin{aligned} \max_{\hat{q}, \hat{U}} & \left\{ (1 - \lambda) \int_{\underline{\theta}}^{\bar{\theta}} \left[ \theta_j \hat{q}(\theta_j) - \frac{\hat{q}(\theta_j)^2}{2} - \hat{U}(\theta_j) \right] G^{J-1}(\theta_j) g(\theta_j) d\theta_j \right. \\ & + \lambda \int_{\underline{\theta}}^{\bar{\theta}} \left[ \theta_j \hat{q}(\theta_j) - \frac{\hat{q}(\theta_j)^2}{2} - \hat{U}(\theta_j) \right] F^{J-1}(\theta_k^*(\theta_j)) f(\theta_j) d\theta_j \\ & \left. + \lambda \int_{\underline{\theta}}^{\bar{\theta}} \left( \frac{\theta_j^2}{2} - \theta_j \hat{q}(\theta_j) + \frac{\hat{q}(\theta_j)^2}{2} \right) \min \{ F^{J-1}(\theta_k^*(\theta_j)), F^{J-1}(\theta_j) \} f(\theta_j) d\theta_j \right\}, \end{aligned}$$

where, as in (28),  $\theta_k^*$  satisfies

$$\hat{U}_k(\theta_k^*(\theta_j)) = \hat{U}_j(\theta_j).$$

The associated Hamiltonian can be written as

$$\begin{aligned} H(\theta_j, \hat{q}, \hat{U}, \gamma) &= (1 - \lambda) [\theta_j \hat{q}(\theta_j) - \hat{q}(\theta_j)^2 - \hat{U}(\theta_j)] G^{J-1}(\theta_j) g(\theta_j) \\ &\quad + \lambda \left( \theta_j \hat{q}(\theta_j) - \frac{\hat{q}(\theta_j)^2}{2} - \hat{U}(\theta_j) \right) F^{J-1}(\theta_k^*(\theta_j)) f(\theta_j) \\ &\quad + \lambda \left( \frac{\theta_j^2}{2} - \theta_j \hat{q}(\theta_j) + \frac{\hat{q}(\theta_j)^2}{2} \right) \min \{ F^{J-1}(\theta_k^*(\theta_j)), F^{J-1}(\theta_j) \} f(\theta_j) \\ &\quad + \gamma(\theta_j) \hat{q}(\theta_j). \end{aligned}$$

Totally differentiating (28), we obtain

$$\frac{dF^{J-1}(\theta_k^*(\theta_j))}{d\hat{U}_j(\theta_j)} = \frac{(J-1) F^{J-2}(\theta_j) f(\theta_j)}{\hat{q}(\theta_j)} \geq 0.$$

Because seller  $j$ 's market share is increasing in  $\hat{U}_j$  and  $\theta_j^2/2 \geq \theta_j \hat{q}(\theta_j) - \hat{q}(\theta_j)^2/2$ , the Hamiltonian  $H$  has a downward kink in  $\hat{U}_j$  at  $\theta_k^* = \theta_j$ . Therefore, every symmetric equilibrium satisfies the following necessary conditions,

$$(39) \quad (1 - \lambda) [\theta_j - \hat{q}(\theta_j)] G^{J-1}(\theta_j) g(\theta_j) + \gamma(\theta_j) = 0,$$

$$\begin{aligned} (40) \quad & -\lambda \frac{(J-1) F^{J-2}(\theta_j) f^2(\theta_j)}{\hat{q}(\theta_j)} \left[ \alpha \frac{\theta_j^2}{2} + (1 - \alpha) \left( \theta_j \hat{q}(\theta_j) - \frac{\hat{q}(\theta_j)^2}{2} \right) - \hat{U}(\theta_j) \right] \\ & + (1 - \lambda) G^{J-1}(\theta_j) g(\theta_j) + \lambda F^{J-1}(\theta_j) f(\theta_j) = \gamma'(\theta_j), \end{aligned}$$

for some  $\alpha \in [0, 1]$ .

Now compare the costate equation (40) with (35) in the baseline model. Because the term in parentheses in the first line of (40) is nonnegative, we have

$$\gamma'(\theta_j) \leq \hat{\gamma}'(\theta_j) \text{ for all } \theta_j.$$

Furthermore, the transversality conditions in the two problems require

$$\gamma(\bar{\theta}) = \hat{\gamma}(\bar{\theta}) = 0.$$

We can then conclude that

$$\hat{\gamma}(\theta_j) \leq \gamma(\theta_j) \text{ for all } \theta_j.$$

Together with (34) and (39), this implies that the quality and utility levels  $\hat{q}_j(\theta_j)$  and  $\hat{U}_j(\theta_j)$  are weakly higher for all  $\theta_j$  with organic links than without.

We now show that the sellers' profits are lower and their outside options are higher with organic links than without. In a symmetric equilibrium with off-platform quality and rent functions  $(\hat{q}, \hat{U})$ , the seller's profits are given by

$$(41) \quad \Pi_j(\hat{q}, \hat{U}) = (1 - \lambda) \int_{\underline{\theta}}^{\bar{\theta}} \left[ \theta_j \hat{q}(\theta_j) - \frac{\hat{q}(\theta_j)^2}{2} - \hat{U}(\theta_j) \right] G^{J-1}(\theta_j) dG(\theta_j) \\ + \lambda \int_{\underline{\theta}}^{\bar{\theta}} \left[ \frac{\theta_j^2}{2} - \hat{U}(\theta_j) \right] F^{J-1}(\theta_j) dF(\theta_j).$$

The equilibrium menu in the baseline model  $(\hat{q}_j^*, \hat{U}_j^*)$  maximizes (41), while the equilibrium menu maximizes (29) and hence achieves a weakly lower profit level. Now consider the deviating seller's profit. For any choice of  $(\hat{q}, \hat{U})$  off path, the deviation profits are weakly larger with organic links than without. Without organic links, the deviating seller posts the Mussa and Rosen (1978) menu and makes no sales on the platform. With organic links and posting the same menu, the seller wins a fraction  $F^{J-1}(\theta_k^*(\theta_j)) \in [0, 1]$  of on-platform values  $\theta_j$ . Consequently, we have  $\tilde{\Pi}_j \geq \bar{\Pi}_j$ , which implies a fortiori that the advertising budgets are lower. ■

#### PROOF OF PROPOSITION 7:

We construct an equilibrium where each seller  $j$  sets on- and off-platform menus to maximize profit, given that seller  $j$  expects to face all consumers that rank  $j$  the highest. Therefore, consider the joint optimization problem over menus  $(q, U)$  and  $(\hat{q}, \hat{U})$  when facing distributions  $F^J$  and  $G^J$ , respectively, under the showrooming constraint. Seller  $j$  solves

$$(42) \quad \max_{q, \hat{q}, U, \hat{U}} \left\{ \lambda \int_{\underline{\theta}}^{\bar{\theta}} \left[ \theta_j q(\theta_j) - \frac{q(\theta_j)^2}{2} - U(\theta_j) \right] F^{J-1}(\theta_j) dF(\theta_j) \right. \\ \left. + (1 - \lambda) \int_{\underline{\theta}}^{\bar{\theta}} \left( \theta_j \hat{q}(\theta_j) - \frac{\hat{q}(\theta_j)^2}{2} - \hat{U}(\theta_j) \right) G^{J-1}(\theta_j) dG(\theta_j) \right\}$$

subject to

$$U'(\theta_j) = q(\theta_j)$$

$$\hat{U}'(\theta_j) = \hat{q}(\theta_j)$$

$$U(\theta_j) \geq \hat{U}(\theta_j) \geq 0.$$

We now show that the solution to (42) is given by

$$(43) \quad q_j^*(\theta_j) = \hat{q}_j^*(\theta_j) = \theta_j - \frac{1 - (1 - \lambda) G^J(\theta_j) - \lambda F^J(\theta_j)}{J \lambda F^{J-1}(\theta_j) f(\theta_j) + J(1 - \lambda) G^{J-1}(\theta_j) g(\theta_j)}$$

if and only if the  $F^J$  likelihood-ratio dominates  $G^J$ . To this end, consider the necessary conditions for optimality. These conditions are sufficient because the problem is linear in  $q$ , concave in  $U$ , and additively separable in these two variables. In particular, the Hamiltonian is given by

$$\begin{aligned} H = & \lambda \left[ \theta_j q(\theta_j) - \frac{q(\theta_j)^2}{2} - U(\theta_j) \right] F^{J-1}(\theta_j) f(\theta_j) \\ & + (1 - \lambda) \left[ \theta_j \hat{q}(\theta_j) - \frac{\hat{q}(\theta_j)^2}{2} - \hat{U}(\theta_j) \right] G^{J-1}(\theta_j) g(\theta_j) \\ & + \gamma(\theta_j) q(\theta_j) + \hat{\gamma}(\theta_j) \hat{q}(\theta_j) + \bar{\gamma} [U(\theta_j) - \hat{U}(\theta_j)]. \end{aligned}$$

The pointwise necessary conditions for this problem are the following:<sup>20</sup>

$$\begin{aligned} [\theta_j - q(\theta_j)] \lambda F^{J-1}(\theta_j) f(\theta_j) + \gamma(\theta_j) &= 0 \\ [\theta_j - \hat{q}(\theta_j)] (1 - \lambda) G^{J-1}(\theta_j) g(\theta_j) + \hat{\gamma}(\theta_j) &= 0 \\ - \lambda F^{J-1}(\theta_j) f(\theta_j) + \bar{\gamma}(\theta_j) + \gamma'(\theta_j) &= 0 \\ - (1 - \lambda) G^{J-1}(\theta_j) g(\theta_j) - \bar{\gamma}(\theta_j) + \hat{\gamma}'(\theta_j) &= 0 \\ \bar{\gamma}(\theta_j) [U(\theta_j) - \hat{U}(\theta_j)] &= 0 \\ \bar{\gamma}(\theta_j) &\geq 0. \end{aligned}$$

If  $q(\theta_j) = \hat{q}(\theta_j)$  as in (43), we obtain the following expressions for the costate variables:

$$\begin{aligned} \gamma(\theta_j) &= - \frac{\lambda J F^{J-1}(\theta_j) f(\theta_j) [1 - (1 - \lambda) G^J(\theta_j) - \lambda F^J(\theta_j)]}{(1 - \lambda) J G^{J-1}(\theta_j) g(\theta_j) + \lambda J F^J(\theta_j) f(\theta_j)} \\ \hat{\gamma}(\theta_j) &= - \frac{(1 - \lambda) J G^{J-1}(\theta_j) g(\theta_j) [1 - (1 - \lambda) G^J(\theta_j) - \lambda F^J(\theta_j)]}{(1 - \lambda) J G^{J-1}(\theta_j) g(\theta_j) + \lambda J F^J(\theta_j) f(\theta_j)}. \end{aligned}$$

<sup>20</sup>The last condition ( $\bar{\gamma} \geq 0$ ) is analogous the one in Jullien (2000), Theorem 2. There, the shadow cost of the value-dependent participation constraint is a cumulative distribution function; i.e., it is nondecreasing. The multiplier  $\bar{\gamma}(\theta_j)$  in our formulation can be interpreted as the corresponding density function.

Differentiating both expressions with respect to  $\theta_j$  and using the necessary conditions above, we can solve for the multiplier on the showrooming constraint  $\bar{\gamma}$ . We obtain

$$\begin{aligned} \bar{\gamma}(\theta_j) = & \frac{J\lambda(1-\lambda)[1 - (1-\lambda)G^J(\theta_j) - \lambda F^J(\theta_j)]}{[(1-\lambda)JG^{J-1}(\theta_j)g(\theta_j) + \lambda JF^{J-1}(\theta_j)f(\theta_j)]^2} \\ & \times \left( \frac{dF^{J-1}(\theta_j)f(\theta_j)}{d\theta_j} G^J(\theta_j) - \frac{dG^{J-1}(\theta_j)g(\theta_j)}{d\theta_j} F^J(\theta_j) \right), \end{aligned}$$

which is positive if and only if  $dF^J/dG^J$  is increasing in  $\theta_j$ , i.e., if and only if the  $F^J$  likelihood-ratio dominates  $G^J$ . ■

#### PROOF OF PROPOSITION 8:

When the platform becomes arbitrarily large ( $\lambda \rightarrow 1$ ), rents off platform vanish, and the sponsored seller appropriates the entire surplus it generates. The sponsored seller's profit under complete and symmetric information on each value  $\theta_j$  is then given by the first-best surplus  $\pi_j^*(\theta) = \theta_j^2/2$ . Fix any matching mechanism, and let  $F^*$  denote the distribution of  $\theta_j$  that is matched to seller  $j$ . Therefore, the information design problem of the platform is given by

$$\max_{\hat{F} \prec F^*} \int_{\underline{\theta}}^{\bar{\theta}} \pi_j^*(\theta_j) d\hat{F}(\theta_j).$$

Because  $\pi_j^*(\cdot)$  is strictly convex, the platform-optimal information design sets  $\hat{F} = F^*$ , i.e., it reveals to each consumer their true value for the sponsored seller. Furthermore, by Proposition 4, it is optimal to match consumers and sellers efficiently (i.e., to further let  $F^* = F^J$ ) when the platform reveals all the available information. ■

#### PROOF OF PROPOSITION 9:

Fix  $\hat{q}$  and let  $\pi(\theta_j)$  denote the on-platform profit function. By Dworczak and Martini (2019, Theorem 1), if there exists a distribution  $\hat{F} \prec F^J$  and a convex supporting function  $y(\theta_j)$  such that

$$y(\theta_j) \geq \pi(\theta_j),$$

$$\int_{\underline{\theta}}^{\bar{\theta}} y(\theta_j) dF^J(\theta_j) = \int_{\underline{\theta}}^{\bar{\theta}} y(\theta_j) d\hat{F}(\theta_j), \text{ supp}(\hat{F}) \subseteq \{\theta_j \in [\underline{\theta}, \bar{\theta}] : y(\theta_j) = \pi(\theta_j)\},$$

then  $\hat{F}$  solves the problem for the given  $\hat{q}$ .

As the function  $\pi(\theta_j)$  satisfies the regularity conditions of Dworczak and Martini (2019), we can compute the supporting function and the associated distribution: by Dworczak and Martini (2019, Proposition 1), the support of the optimal  $\hat{F}$  is found by solving

$$\min_y \int_{\underline{\theta}}^{\bar{\theta}} y(\theta_j) dF^J(\theta_j)$$

subject to

$$y(\theta_j) \geq \pi(\theta_j), \forall \theta_j; \quad y \text{ convex.}$$

Moreover, the optimal  $\hat{F}$  is then supported only on points where  $y^* = \pi$ . Thus, we can compute the supporting function independent of the distribution.

To solve the problem, we first reduce it to the choice of one variable, namely the slope  $s$  of the affine function  $y$  (when  $y \neq \pi$ ). Call  $x_1, x_2$  the intersection points of  $y$  and  $\pi$ . Then it holds that

$$\begin{aligned} \frac{x_2^2}{2} - (x_2 - \mu)\hat{q} &= \frac{\mu^2}{2} + s(x_2 - \mu) \\ \frac{\mu^2}{2} - s(\mu - x_1) &= \frac{x_1^2}{2}. \end{aligned}$$

Therefore, solving, we have

$$x_1 = 2s - \mu \quad \text{and} \quad x_2 = 2s - \mu + 2\hat{q},$$

and we can write the objective as

$$\begin{aligned} \min_s \left\{ \int_0^{x_1(s)} \frac{\theta_j^2}{2} dF^J(\theta_j) + \int_{x_1(s)}^{x_2(s)} \left( \frac{x_1^2(s)}{2} + s[\theta_j - x_1(s)] \right) dF^J(\theta_j) \right. \\ \left. + \int_{x_2(s)}^1 \left[ \frac{\theta_j^2}{2} - \hat{q}(\theta_j - \mu) \right] dF^J(\theta_j) \right\}. \end{aligned}$$

Solving the first-order condition for  $s$  in the problem above yields

$$\int_{x_1(s)}^{x_2(s)} (\theta_j - \mu) dF^J(\theta_j) = 0.$$

Finally, from Dworczak and Martini (2019, Theorem 1), we know the support of  $\hat{F}^*$  is  $[0, x_1] \cup \{\mu\} \cup [x_2, 1]$ . Moreover, duality ensures that the *optimal* supporting function  $y^*(\theta_j)$  yields a mean-preserving contraction of  $F^J$ .

Finally, note that

$$\begin{aligned} \Pi'(\hat{q}) &= -\lambda \left[ \frac{(x_2^2 - \mu^2)}{2} - \hat{q}(x_2 - \mu) \right] dF^J(x_2(s)) x_2'(q) \\ &\quad - \lambda \int_{x_2(s)}^1 (\theta_j - \mu) dF^J(\theta_j) + (1 - \lambda)(\mu - \hat{q}); \end{aligned}$$

hence,

$$\hat{q} = \frac{\mu - \frac{\lambda}{1-\lambda} \left[ \int_{x_2(s)}^1 (\theta_j - \mu) dF^J(\theta_j) + \frac{x_2^2 - \mu^2}{2} JF^{J-1}(x_2) f(x_2) x_2'(q) \right]}{1 - \frac{\lambda}{1-\lambda} (x_2 - \mu) JF^{J-1}(x_2) f(x_2) x_2'(q)}.$$

Therefore,  $\hat{q}$  is decreasing in  $\lambda$ , and consequently,  $x_2$  is decreasing and  $x_1$  increasing in  $\lambda$ . ■

## REFERENCES

- Acemoglu, Daron, Ali Makhdoui, Azarakhsh Malekian, and Asu Ozdaglar. 2022. "Too Much Data: Prices and Inefficiencies in Data Markets." *American Economic Journal: Microeconomics* 14 (4): 218–56.
- Aggarwal, Gagan, Ashwinkumar Badanidiyuru, and Aranyak Mehta. 2019. "Autobidding with Constraints." In *International Conference on Web and Internet Economics*, edited by Ioannnis Caragiannis, Evdokia Nikolova, and Vahab Mirrokni, 17–30. New York, NY: Springer.
- Anderson, Simon P., and Özlem Bedre-Defolie. Forthcoming. "Hybrid Platform Model." *RAND Journal of Economics*.
- Anderson, Simon P., and Regis Renault. 1999. "Pricing, Product Diversity, and Search Costs: A Bertrand-Chamberlin-Diamond Model." *RAND Journal of Economics* 30 (4): 719–35.
- Armstrong, Mark, and Jidong Zhou. 2011. "Paying for Prominence." *Economic Journal* 121 (556): F368–95.
- Balseiro, Santiago R., Yuan Deng, Jieming Mao, Vahab S. Mirrokni, and Song Zuo. 2021. "The Landscape of Auto-Bidding Auctions: Value Versus Utility Maximization." In *Proceedings of the 22nd ACM Conference on Economics and Computation*, edited by Association for Computing Machinery, 132–33. New York, NY: Association for Computing Machinery.
- Bar-Isaac, Heski, and Sandro Shelegia. 2022. "Search, Showrooming, and Retailer Variety." CEPR Discussion Paper DP15448.
- Baye, Michael R., and John Morgan. 2001. "Information Gatekeepers on the Internet and the Competitiveness of Homogeneous Product Markets." *American Economic Review* 91 (3): 454–74.
- Bergemann, Dirk, Alessandro Bonatti, and Tan Gan. 2022. "The Economics of Social Data." *RAND Journal of Economics* 53 (2): 263–96.
- Bergemann, Dirk, Alessandro Bonatti, and Nicholas Wu. Forthcoming. "How Do Digital Advertising Auctions Impact Product Prices?" *Review of Economic Studies*.
- Bergemann, Dirk, Benjamin Brooks, and Stephen Morris. 2015. "The Limits of Price Discrimination." *American Economic Review* 105 (3): 921–57.
- Bergemann, Dirk, Benjamin Brooks, and Stephen Morris. 2021. "Search, Information, and Prices." *Journal of Political Economy* 129 (8): 2275–2319.
- Blackwell, David. 1951. "Comparison of Experiments." In *Proceedings of the Second Berkeley Symposium in Mathematical Statistics and Probability*, edited by Jerzy Neyman, 93–102. Berkeley, CA: University of California Press.
- Bonatti, Alessandro. 2011. "Brand-Specific Tastes for Quality." *International Journal of Industrial Organization* 29 (5): 562–75.
- Calzolari, Giacomo, and Vincenzo Denicolo. 2015. "Exclusive Contracts and Market Dominance." *American Economic Review* 105 (11): 3321–51.
- Chen, Daniel. 2022. "The Market for Attention." Unpublished.
- Choi, Jay Pil, Doh-Shin Jeon, and Byung-Cheol Kim. 2019. "Privacy and Personal Data Collection with Information Externalities." *Journal of Public Economics* 173: 113–24.
- Condorelli, Daniele, and Balazs Szentes. 2023. "Buyer-Optimal Platform Design." Unpublished.
- de Cornière, Alexandre, and Romain de Nijs. 2016. "Online Advertising and Privacy." *RAND Journal of Economics* 47 (1): 48–72.
- de Cornière, Alexandre, and Greg Taylor. 2019. "A Model of Biased Intermediation." *RAND Journal of Economics* 50 (4): 854–82.
- de Montjoye, Yves-Alexandre, Heike Schweitzer, and Jacques Crémer. 2019. *Competition policy for the digital era*. European Commission.
- Decarolis, Francesco, Gabriele Rovigatti, Michele Rovigatti, and Ksenia Shakhgildyan. 2022. "Artificial intelligence, algorithmic bidding and collusion in online advertising." Unpublished.
- Deng, Yuan, Negin Golrezaei, Patrick Jaillet, Jason Cheuk Nam Liang, and Vahab Mirrokni. 2023. "Individual Welfare Guarantees in the Autobidding World with Machine-Learned Advice." <https://arxiv.org/abs/2209.04748>.
- Deng, Yuan, Jieming Mao, Vahab Mirrokni, Hanrui Zhang, and Song Zuo. 2022. "Efficiency of the First-Price Auction in the Autobidding World." <https://arxiv.org/abs/2208.10650>.
- Deng, Yuan, Jieming Mao, Vahab Mirrokni, and Song Zuo. 2021. "Towards Efficient Auctions in an Auto-Bidding World." In *Proceedings of the Web Conference 2021*, edited by Jure Leskovec, 3965–73. New York, NY: Association for Computing Machinery.
- Diamond, Peter A. 1971. "A Model of Price Adjustment." *Journal of Economic Theory* 3 (2): 156–68.
- Donnelly, Robert, Ayush Kanodia, and Ilya Morozov. 2022. "Welfare Effects of Personalized Rankings." Unpublished.



- Dworczak, Piotr, and Giorgio Martini. 2019. "The Simple Economics of Optimal Persuasion." *Journal of Political Economy* 127 (5): 1993–2048.
- Elliott, Matthew, Andrea Galeotti, Andrew Koh, and Wenhao Li. 2021. "Market segmentation through information." Cambridge Working Papers in Economics 2105.
- Fuchs, William, and Andrzej Skrzypacz. 2015. "Government Interventions in a Dynamic Market with Adverse Selection." *Journal of Economic Theory* 158: 371–406.
- Garrett, Daniel F., Renato Gomes, and Lucas Maestri. 2019. "Competitive Screening under Heterogeneous Information." *Review of Economic Studies* 86 (4): 1590–1630.
- Gomes, Renato, and Alessandro Pavan. 2016. "Many-to-Many Matching and Price Discrimination." *Theoretical Economics* 11 (3): 1005–52.
- Gomes, Renato, and Alessandro Pavan. Forthcoming. "Price Customization and Targeting in Matching Markets." *RAND Journal of Economics*.
- Gur, Yonatan, Gregory Macnamara, Ilan Morgenstern, and Daniela Saban. 2022. "Information Disclosure and Promotion Policy Design for Platforms." Unpublished.
- Hagiu, Andrei, and Bruno Jullien. 2011. "Why Do Intermediaries Divert Search?" *RAND Journal of Economics* 42 (2): 337–62.
- Hagiu, Andrei, Tat-How Teh, and Julian Wright. Forthcoming. "Should Platforms be Allowed to Sell on Their Own Marketplaces?" *RAND Journal of Economics*.
- Hidir, Sinem, and Nikhil Vellodi. 2021. "Privacy, Personalization, and Price Discrimination." *Journal of the European Economic Association* 19 (2): 1342–63.
- Ichihashi, Shota. 2020. "Online Privacy and Information Disclosure by Consumers." *American Economic Review* 110 (2): 569–95.
- Ichihashi, Shota. 2021. "The Economics of Data Externalities." *Journal of Economic Theory* 196.
- Idem, Berk. 2021. "Coexistence of Centralized and Decentralized Markets." <https://arxiv.org/abs/2111.12767>.
- Inderst, Roman, and Marco Ottaviani. 2012a. "Competition Through Commissions and Kickbacks." *American Economic Review* 102 (2): 780–809.
- Inderst, Roman, and Marco Ottaviani. 2012b. "How (Not) to Pay for Advice: A Framework for Consumer Financial Protection." *Journal of Financial Economics* 105 (2): 393–411.
- Jullien, Bruno. 2000. "Participation Constraints in Adverse Selection Models." *Journal of Economic Theory* 93 (1): 1–47.
- Kang, Zi Yang, and Ellen V. Muir. 2021. "Contracting and vertical control by a dominant platform." Unpublished.
- Ke, T. Tony, Song Lin, and Michelle Y. Lu. 2022. "Information Design of Online Platforms." Unpublished.
- Kirpalani, Rishabh, and Thomas Philippon. 2020. "Data Sharing and Market Power with Two-Sided Platforms." NBER Working Paper 28023.
- Kolotilin, Anton, Ming Li, Tymofiy Mylovanov, and Andriy Zapechelnuk. 2017. "Persuasion of a Privately Informed Receiver." *Econometrica* 85 (6): 1949–64.
- Krähmer, Daniel, and Roland Strausz. 2023. "Optimal Nonlinear Pricing with Data-Sensitive Consumers." *American Economic Journal: Microeconomics* 15 (2): 80–108.
- Lam, H. Tai. 2022. "Platform Search Design and Market Power." Unpublished.
- Lee, Changhwa. 2023. "Optimal Recommender System Design." Unpublished.
- Lee, Kwok Hao, and Leon Musolff. 2021. "Entry into two-sided markets shaped by platform-guided search." Unpublished.
- Liang, Annie, and Erik Madsen. Forthcoming. "Data and Incentives." *Theoretical Economics*.
- MacMahan, Brendan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. 2017. "Communication-Efficient Learning of Deep Networks from Decentralized Data." In *20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, edited by Aarti Singh and Jerry Zhu, 1273–82.
- Maskin, Eric, and John Riley. 1984. "Monopoly with Incomplete Information." *RAND Journal of Economics* 15 (2): 171–96.
- Mekonnen, Teddy, Zeky Murra-Anton, and Bobak Pakzad-Hurson. 2023. "Persuaded Search." <https://arxiv.org/abs/2303.13409>.
- Miklós-Thal, Jeanine, and Greg Shaffer. 2021. "Input Price Discrimination by Resale Market." *RAND Journal of Economics* 52 (4): 727–57.
- Mussa, Michael, and Sherwin Rosen. 1978. "Monopoly and Product Quality." *Journal of Economic Theory* 18 (2): 301–17.
- Mustri, Eduardo Schnadower, Idris Adjerid, and Alessandro Acquisti. 2022. "Behavioral Advertising and Consumer Welfare: An Empirical Investigation." Unpublished.
- Myerson, Roger. 1981. "Optimal Auction Design." *Mathematics of Operations Research* 6 (1): 58–73.

- Padilla, Jorge, Joe Perkins, and Salvatore Piccolo.** 2020. "Self-Preferencing in Markets with Vertically-Integrated Gatekeeper Platforms." Unpublished.
- Perloff, Jeffrey M., and Steven C. Salop.** 1985. "Equilibrium with Product Differentiation." *Review of Economic Studies* 52 (1): 107–20.
- Philippon, Thomas, and Vasiliki Skreta.** 2012. "Optimal Interventions in Markets with Adverse Selection." *American Economic Review* 102 (1): 1–28.
- Raval, Shreya.** 2020. "eBay Makes Search More Efficient Through Personalization." *eBay*, June 23. <https://innovation.ebayinc.com/tech/product/ebay-makes-search-more-efficient-through-personalization/>.
- Rayo, Luis, and Ilya Segal.** 2010. "Optimal Information Disclosure." *Journal of Political Economy* 118 (5): 949–87.
- Ronayne, David.** 2021. "Price Comparison Websites." *International Economic Review* 62 (3): 1081–1110.
- Shaked, Moshe, and J. George Shanthikumar.** 1994. *Stochastic Orders and their Applications*. San Diego, CA: Academic Press.
- Shi, Peng.** 2022. "Optimal Match Recommendations in Two-Sided Marketplaces with Endogenous Prices." Unpublished.
- Teh, Tat-How, and Julian Wright.** 2022. "Intermediation and Steering: Competition in Prices and Commissions." *American Economic Journal: Microeconomics* 14 (2): 281–321.
- Tirole, Jean.** 2012. "Overcoming Adverse Selection: How Public Intervention Can Restore Market Functioning." *American Economic Review* 102 (1): 29–59.
- Wang, Chengsi, and Julian Wright.** 2020. "Search Platforms: Showrooming and Price Parity Clauses." *RAND Journal of Economics* 51 (1): 32–58.
- Wolinsky, Asher.** 1986. "True Monopolistic Competition as a Result of Imperfect Information." *Quarterly Journal of Economics* 101 (3): 493–511.
- Xiao, Yao, and Josh Karlin.** 2024. "Topics API." Unpublished.
- Yang, Kai Hao.** 2022. "Selling Consumer Data for Profit: Optimal Market-Segmentation Design and its Consequences." *American Economic Review* 112 (4): 1364–93.
- Zhu, Ruizhi, Sridhar Moorthy, and Xianwen Shi.** 2022. "Advertising Platforms and Privacy." Unpublished.