Data, Competition, and Digital Platforms*

Dirk Bergemann† Alessandro Bonatti‡

August 3, 2022

Abstract

We propose a model of intermediated digital markets where data and heterogeneity in tastes and products are defining features. A monopolist platform uses superior data to match consumers and multiproduct advertisers. Consumers have heterogenous preferences for the advertisers' product lines and shop on- or off-platform. The platform monetizes its data by selling targeted advertising space that allows advertisers to tailor their products to each consumer's preferences. We derive the equilibrium product lines and advertising prices. We identify search costs and informational advantages as two sources of the platform's bargaining power. We show that privacy-enhancing data-governance rules, such as those corresponding to federated learning, can lead to welfare gains for the consumers.

KEYWORDS: Data, Privacy, Data Governance, Digital Advertising, Competition, Digital Platforms, Digital Intermediaries, Personal Data, Matching, Price Discrimination.

JEL CLASSIFICATION: D18, D44, D82, D83.

*We acknowledge financial support through NSF Grant SES-1948692, the Omidyar Network, and the Sloan Foundation. We thank audiences at Indiana, Northwestern, Paris-Panthéon, as well as participants in the CEPR Virtual IO Seminar, the Luohan Academy Webinar, the 2021 FTC Microeconomics Conference, the 2021 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 and Vasiliki Skreta for helpful comments and suggestions, and to Andrew Koh and Nick Wu for stellar research assistance.

†Department of Economics, Yale University, New Haven, CT 06511, dirk.bergemann@yale.edu

‡Sloan School of Management, MIT, Cambridge, MA 02142, bonatti@mit.edu

1 Introduction

The role of data in shaping competition in online markets is a critical issue for both economics and policy. Over the last decade, digital platforms—such as Amazon, Facebook, and Google in the United States, and Alibaba, JD, and Tencent in China—have collected increasingly large and precise datasets. These platforms operate as matching engines that connect viewers and producers, and they monetize their data by selling sponsored content and targeted advertising space. The quality of a platform's data allows better pairing of viewers and advertisers, but also controls the type of content consumers are exposed to and the products they are informed about. Thus, digital platforms not only serve as gatekeepers of information online but also act as competition managers.

The platforms' dual gatekeeping role has recently come under scrutiny. Regulators fear that platforms may leverage their privileged position to increase merchants' market power and charge higher prices for advertising.[1] 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: how does the precision of a digital platform's data affect the creation and distribution of surplus, both on and off the platform? How do these effects depend on the intensity of competition among advertisers? 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 out how a data-rich platform creates and distributes surplus among market participants. Our model captures three ubiquitous features of personalized sponsored content on digital platforms. First, the platform leverages its informational advantage to personalize the *sponsored content* at the individual consumer level.[2] Second, while price discrimination is conspicuously rare, targeted advertising and personalized recommendations amount to *product steering*.[3] Third, most advertisers have *parallel sales channels*, i.e., consumers can buy their products both on and off digital platforms. Our approach consists of providing a tractable and flexible framework to study digital markets where different privacy regimes can be compared.

We consider an imperfectly competitive market with differentiated quality-pricing firms.

[1] The report by Cremèr, de Montjoye, and Schweitzer (2019) 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."

[2] Sponsored links are the only source of revenue for pure advertising platforms, including display advertising networks such as Google, Meta, Microsoft, Twitter, Tiktok, YouTube, and Criteo. Sponsored content is also a significant revenue generator for several retail platforms that charge merchant fees (eBay, Wayfair, Booking, Orbitz, Amazon) and the only source of revenue for Alibaba's Taobao shopping platform.

[3] Donnelly, Kanodia, and Morozov (2022) document the effect of personalized recommendations on a retail platform, and https://tech.ebayinc.com/product/ebay-makes-search-more-efficient-through-personalization/ illustrates a recent shift in eBay policy.

Consumers have heterogeneous preferences for the firms' product lines but are imperfectly informed about their own types. The platform's data identify the most valuable consumer-advertiser pair and the most valuable product within that advertiser's product line. However, these data also create the potential for price discrimination through product steering, whereby consumers who are perceived to be of high value receive offers to buy higher-quality and higher-priced goods.

A key innovation in our model is that the platform actively manages the firms' advertising campaigns. Managed campaigns are emerging as the predominant mode of selling advertisements in real-world digital markets: advertisers set a fixed budget, specify high-level objectives for their campaigns, and leave the task of bidding to "autobidders."[4] In our model, we assume that the digital platform charges a fixed fee to each advertiser. The platform then shows a single sponsored listing to each consumer that advertises the product generating the highest social surplus for that type. This mechanism is both realistic (in practice, Google auctions use bids, click-through rates, and relevance metrics to assign sponsored links) and optimal in our model (i.e., it maximizes the platform's revenues).

Each firm also has a pool of consumers who shop off the platform and face search costs. As in the Diamond (1971) model, these consumers only visit one firm's website in equilibrium. In our setting, they visit the firm they believe they have the strongest preference for. The presence of the off-platform sales channel restrains firms' ability to extract consumer surplus on the platform because the on-platform consumers can seamlessly move from the platform to individual websites. Thus, we combine elements of nonlinear pricing and market segmentation, where the consumer's choice of sales channel limits the scope for price discrimination. In particular, the more the firm wants to trade with its loyal consumers off the platform, the less flexibility it has to offer targeted promotions on the platform.

The platform's informational advantage vis-à-vis the consumers and the search frictions off the platform—no matter how small—grant the digital platform significant bargaining power over the firms. Indeed, our main result shows that the platform is able to completely control the consumers' shopping behavior and to steer them away from any firm that does not pay the advertising fee. This is because strategic consumers understand the managed-campaign mechanism and expect the advertised products in equilibrium to generate the highest value. Because of search costs, these consumers then contemplate buying (on or off the platform) from the advertised brand only. Therefore, the platform is able to fully restrict competition among firms. In particular, each firm faces only those consumers who like its products the most and competes with its own off-platform offers only.

[4] For a treatment of this problem, see Aggarwal, Badanidiyuru, and Mehta (2019), Balseiro, Deng, Mao, Mirrokni, and Zuo (2021), and Deng, Mao, Mirrokni, and Zuo (2021).

Therefore, firms face an additional opportunity cost of generating surplus off the platform: not only must they concede information rents to the off-platform consumers; they must also lower their prices on the platform. This result has two welfare consequences. First, the off-platform equilibrium quality levels are distorted downward from the efficient levels further than in the Mussa and Rosen (1978) model. Second, the platform is able to extract most of the surplus it generates—it must only compensate the firms for the additional distortions in their off-platform menus of products.

We show that these distortions become more severe as the fraction of on-platform consumers grows and as the platform's information becomes more precise. The implications of the platform's information precision for consumer surplus are more intricate. Holding prices fixed, consumers benefit from shopping on a well-informed platform and obtaining higher-quality matches. As more consumers join the platform, however, prices rise both on and off the platform. In other words, consumers impose a negative externality on each other by shopping on the platform.

We then return to the two sources of the platform's bargaining power: its own information advantage vs. consumers and producers and the consumers' search costs off the platform. We first suppose the consumers on the platform observe their types perfectly. We show that the platform cannot steer their behavior away from their favorite firm, were that firm to reject the platform's offer. Thus, the on-platform consumer's prior information does not change the equilibrium prices or products but rather reduces the platform's fees. We then assume that the platform offers organic links that advertise all off-platform prices to all on-platform consumers. We show that the provision of price information introduces menu competition among firms. This reduces all prices (both on and off the platform) relative to the baseline model, as well as the platform's fees.

Finally, we ask how data governance—the rules governing how the consumer's data can be collected and deployed—influences the creation and distribution of social surplus. In particular, we discuss the implications of cohort-based (vs. personalized) advertising as the outcome of federated learning. We show how this form of privacy protection improves consumer surplus by allowing consumers to retain an information advantage relative to the firms on the platform. Our model is also amenable to study other significant issues in the economics of digital platforms, such as preferential treatment by the platform and other elements of choice architecture.

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 the consumers they serve. Many recent contributions—including Armstrong and Zhou (2011), De Corniere and Taylor (2019), Gomes and Pavan (2016), Gur, Macnamara, Morgenstern, and Saban (2022), Hagiu and Jullien (2011), Inderst and Ottaviani (2012a), Inderst and Ottaviani (2012b), Ke, Ling, and Lu (2022), Rayo and Segal (2010), Shi (2022), and especially Teh and Wright (2022)—analyze the steering role of platforms that strategically modify search results, e.g., to match consumers with the firms that pay the largest commissions.

As in this literature, the platform uses its superior information to choose the most valuable firm for each consumer. Our paper shares with Teh and Wright (2022) the signaling role of ranking the search results but also allows the platform to provide information directly to the consumer, e.g., through product reviews. Moreover, the multiproduct firms in our model can use the platform's information to tailor their quality level to the consumer's preferences. This allows us to capture surplus creation and product steering *within* a match. Finally, relative to the papers above, our model focuses on sponsored links and advertising platforms. Hence, firms pay fees that do not vary with the prices of the goods they sell.

A recent body of work including Choi, Jeon, and Kim (2019), Acemoglu, Makhdoumi, Malekian, and Ozdaglar (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 other consumers) influences its ability to match products to tastes but also affects each consumer's outside option. We trace the implications of these new data externalities for product-line design under alternative privacy regimes.

The forces at work in our paper are also related to a growing literature on showrooming, product lines, and multiple sales channels. Prominent contributions on these topics include Bar-Isaac and Shelegia (2020), Idem (2021), Miklós-Thal and Shaffer (2021), and Wang and Wright (2020). In particular, Anderson and Bedre-Defolie (2021) introduce the self-preferencing problem by letting the platform choose whether to be hybrid, i.e., to sell its private label products. Unlike in these papers, the firms 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.

Our analysis of parallel sales channels is also related to the papers on "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 its ability to monopolize the market since the firms have access to an outside option. Our setting shares some of the same features, but in an oligopoly environment where firms compete for heterogeneous consumers. Furthermore, the firms 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 to nonlinear pricing, market segmentation, and competition, e.g., Bergemann, Brooks, and Morris (2015), Bonatti (2011), Elliott, Galeotti, and Koh (2020), and Yang (2022). Finally, our analysis can be easily extended to discuss self-preferencing by a monopoly platform. In this sense, our paper also relates to Hagiu, Teh, and Wright (2022), Kang and Muir (2021), Lam (2021), Lee (2021), Lee and Musolff (2021), and Padilla, Perkins, and Piccolo (2020).

2 Model

We consider a digital platform and J differentiated multiproduct firms. Each firm j offers a product line à la Mussa and Rosen (1978) and can produce a good of quality q_j at a cost

$$c(q_j) = q_j^2/2.$$

There is a unit mass of consumers with single-unit demand. The consumer's type captures her marginal willingness to pay for each firm j's products,

$$\theta = (\theta_1, ..., \theta_j, ..., \theta_J) \in [0,1]^J,$$

where the restriction of the types to the hypercube is merely a normalization for any given finite support of values. In particular, a consumer of type θ has value

$$u\left(\theta, q_j\right) = \theta_j \cdot q_j$$

for consuming a product of quality q_j produced by firm j.

Neither the consumers nor the producers initially know the consumer's valuation. Consumers' valuations θ_j are i.i.d. across consumers and firms with marginal distribution $F(\theta_j)$ and log-concave density $f(\theta_j)$. Each consumer observes a signal s with full support about her type θ. The consumer's posterior expectations generated by the prior distribution F and the signal s,

$$m_j \triangleq \mathbb{E}[\theta_j \mid s],$$

are also assumed i.i.d. with marginal distribution $G\left(m_j\right)$ and log-concave density $g(\theta_j)$. 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. Recall that F is defined to be a mean-preserving spread of G if

$$\int_v^\infty F(t)dt \leq \int_v^\infty G(t)dt, \forall v \in \mathbb{R}_+,$$

with equality for $v = 0$. If F is a mean-preserving spread of G, we write $F \succ G$.

Matching on the Platform: Managed Campaigns A fraction $\lambda \in [0, 1]$ of all consumer types uses the digital platform to find a firm.[5] The platform has access to extensive data that allow it to perfectly observe each consumer's type $\theta \in [0, 1]^J$. The platform offers a single sponsored link and posts a fixed fee t for the advertisers (e.g., a required campaign budget). We assume the platform uses the following mechanism for matching firms and consumers, and we establish its revenue optimality in Proposition 4 in Section 5 below:

- each participating firm j submits quality and price schedules $q_j(\theta)$ and $p_j(\theta)$; these represent the products (and prices) firm j would like to advertise to each consumer θ;
- for each $\theta \in [0, 1]^J$, the platform chooses the product q_j that maximizes the *social value* of the match $\theta_j q_j - c(q_j)$ among all the firms that participate in the mechanism;
- upon advertising product q_j to consumer θ, the platform provides additional information to the consumer that reveals her valuation θ_j for the advertised product.

This mechanism has several critical properties. First, fixed payments for advertising slots are akin to managed campaigns with automated bidding: firms submit a budget and upload the ads for the products they wish to show to select consumers. Second, firms 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 type θ acts as a *targeting category.* This corresponds to an indirect sale of information as discussed in Admati and Pfleiderer (1990) and Bergemann and Bonatti (2019). Third, because each firm can tailor its product offer to each consumer type, the platform creates the opportunity for surplus extraction through *product steering.*

Conversely, some of the assumptions in the above mechanism can be easily relaxed. In particular, the firms' schedules $q_j(\theta)$ and $p_j(\theta)$ are allowed to condition on more than the consumer's value for firm j's own products θ_j, but this additional flexibility will be redundant in equilibrium. Instead of selecting the socially efficient product, the platform

[5] We can endogenize the fraction of on-platform consumers λ by introducing heterogeneity in the cost of using the platform (e.g., in the privacy implications of leaving "footprints" online).

could equivalently select the consumer's favorite firm $j^* = \arg\max_j \theta_j$ and the *most profitable* product offered by firm j^*. Finally, the consumer will be able to infer θ_j in equilibrium from the advertised quality level q_j, but our assumption of direct information revelation captures the rich contextual information that some retail platforms provide to their users.

Matching Off the Platform: Consumer Search The remaining $1 - \lambda$ consumers buy off the platform, e.g., from the merchants' own websites or physical stores. Consumers who buy off the platform face positive (and arbitrarily small) search costs $\gamma > 0$ beyond the first search, as in Diamond (1971) and Anderson and Renault (1999). Each firm only knows the prior distribution of the consumer's beliefs. Therefore, firm j elicits the consumer's private information through a menu of (price, quality) pairs

$$\{(\hat{p}_j(m_j), \hat{q}_j(m_j))\}_{m_j \in [0,1]} \tag{1}$$

as in Mussa and Rosen (1978) and Maskin and Riley (1984). Importantly, the goods being sold are not experience or inspection goods. Thus, to learn the vector θ, consumers and producers must gain access to the platform's data.

The on- and off-platform channels are connected. In particular, if a consumer visits the platform, she learns her value θ_j for the advertised product q_j. Upon receiving firm j's offer, the consumer can still search off-platform and use the newly acquired information to select a product. In particular, the consumer will be able to buy from the off-platform schedule (1) posted by the selected firm j. Figure 1 summarizes the key ingredients of our model.

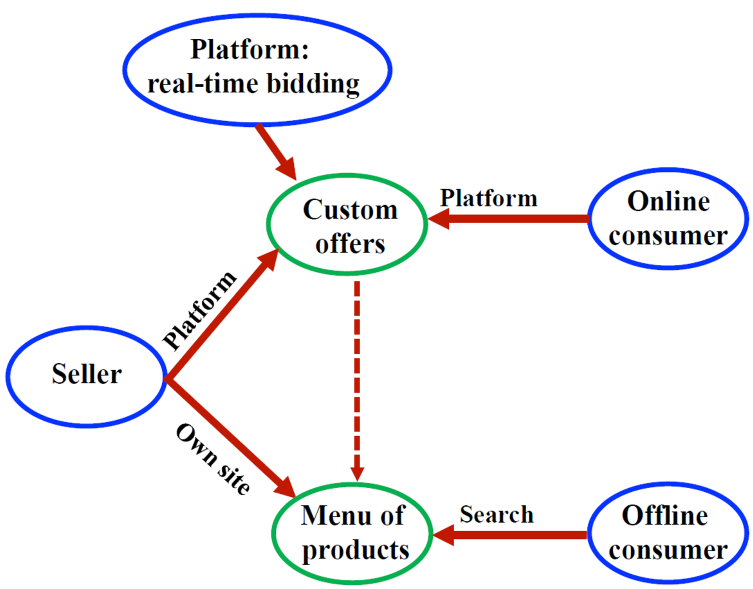

FIGURE 1: On- and Off-Platform Matching

Timing and Equilibrium We consider a game with simultaneous pricing choices that captures the great flexibility that algorithmic pricing offers both on and off the platform. Our baseline model is restrictive in that the platform sells both information and recognition to a single firm. In Section 7.2, we extend our model to allow for all brands to be present on the platform through organic search results that advertise their off-platform offers.

The timing of our baseline setting is as follows:

1. The platform announces the participation fee t.

2. Firms simultaneously set off-platform products $\hat{q}_j(m)$ and prices $\hat{p}_j(m)$, choose whether to pay the fee, and if so, what products $q_j(\theta)$ and prices $p_j(\theta)$ to advertise.

3. The platform shows a single advertisement—a product $q_j(\theta)$ and a price $p_j(\theta)$—to each on-platform consumer.

4. The on-platform consumer learns her valuation θ_j and can buy the advertised product on the platform or search off the platform.

Definition 1 (Symmetric Perfect Bayesian Equilibrium)
We consider symmetric Perfect Bayesian Equilibria in which all consumers hold symmetric beliefs over the firms' off-platform menus both on and off the equilibrium path.

3 A First Example: Single Seller

Before we begin with the analysis of the complete model, we illustrate some of the central implications of our model with a simple example. The example has a single firm (rather than many firms) and binary types (rather than a multidimensional continuum of types). In addition, the distribution of types (and expected willingness to pay) is identical on and off the platform; thus, $F = G$. However, even this simple example illustrates the relation between on-platform and off-platform pricing and quality provision as characterized by Proposition 1. The platform retains an informational advantage as it learns the type of the consumer that remains private information off the platform.

We thus consider a single firm that encounters a mass λ of consumers on the platform and a mass $1 - \lambda$ of consumers off the platform. Consumers can be of two types, $\theta \in \{\theta_L, \theta_H\}$, each with probability $f(\theta)$. The platform charges a fixed fee t to the firm. If the firm pays the fee, it earns the right to offer a personalized product to each type. However, each buyer on the platform can also shop from the firm's own website (i.e., buy the products the firm offers off-platform). Thus, the consumer's option to "showroom" limits the firm's ability to price discriminate.

If the firm pays the platform's fee, it offers a menu of products on the platform, which we describe in terms of the product qualities $q(\theta)$ and information rents $U(\theta)$:

$$\{(q(\theta), U(\theta))\}_{\theta\in\{\theta_L,\theta_H\}},$$

where the information rent is the net utility of the consumer in equilibrium:

$$U(\theta) \triangleq \theta q(\theta) - p(\theta), \quad \theta \in \{\theta_L, \theta_H\}.$$

The firm also offers a menu, distinguished by superscript 0:

$$\{(\hat{q}(\theta), \hat{U}(\theta))\}_{\theta\in\{\theta_L,\theta_H\}}$$

off the platform. The firm's profits are then given by:

$$\max_{q,U} \sum_{\theta\in\{\theta_L,\theta_H\}} f(\theta)[\lambda\left(\theta q(\theta) - q(\theta)^2/2 - U(\theta)\right) + (1-\lambda)(\theta\hat{q}(\theta) - \hat{q}(\theta)^2/2 - \hat{U}(\theta))].$$

The firm maximizes its profits subject to the individual rationality constraints on and off the platform and to the incentive compatibility constraints *off* the platform because the consumers on the platform receive a single targeted product offer. In addition, the firm faces the following new "showrooming" constraints:

$$U(\theta) \geq \hat{U}(\theta), \ \theta \in \{\theta_L, \theta_H\}.$$

Thus, each consumer type θ 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 firm should offer the socially efficient quality levels on the platform and that the showrooming constraint should bind,

$$q(\theta) = \theta \text{ and } U(\theta) = \hat{U}(\theta), \ \theta \in \{\theta_L, \theta_H\}.$$

We now characterize the quality levels off the platform. As usual, the equilibrium menu satisfies no distortion at the top ($\hat{q}(\theta_H) = \theta_H$) and no rents at the bottom ($\hat{U}(\theta_L) = 0$). Furthermore, incentive compatibility binds for the high type. With these preliminary results, the firm's objective can be written as

$$\max_{q,U}\left[\lambda(f(\theta_L)\,\theta_L^2/2+f(\theta_H)\,(\theta_H^2/2-\hat{U}(\theta_H)))\right. \tag{2}$$
$$\left.+(1-\lambda)\left(f(\theta_L)\left(\theta_L\hat{q}(\theta_L)-\hat{q}(\theta_L)^2/2\right)+f(\theta_H)\,(\theta_H^2/2-\hat{U}(\theta_H))\right)\right]$$

subject to the constraint

$$\hat{U}(\theta_H)=(\theta_H-\theta_L)\,\hat{q}(\theta_L)\,. \tag{3}$$

From this expression, it is immediate that the provision of quality to the low type off the platform is doubly costly for the firm: it forces the firm to lower the price for the high type off the platform, and it also forces lower prices on the platform. Therefore, we have the following result.

Proposition 1 (Single Seller, Binary Types)
The optimal off-platform menu of products for a single firm is given by

$$\begin{aligned}\hat{q}(\theta_L) &= \max\left\{0,\theta_L-\frac{f(\theta_H)}{f(\theta_L)}(\theta_H-\theta_L)\left(1+\frac{\lambda}{1-\lambda}\right)\right\}, \\ \hat{q}(\theta_H) &= \theta_H.\end{aligned} \tag{4}$$

Now compare the optimal menu with the classic Mussa and Rosen (1978) solution, which corresponds to the case $\lambda=0$. In that case, we have

$$\begin{aligned}q(\theta_L) &= \max\left\{0,\theta_L-\frac{f(\theta_H)}{f(\theta_L)}(\theta_H-\theta_L)\right\} \\ q(\theta_H) &= \theta_H.\end{aligned} \tag{5}$$

The result in Proposition 1 indicates an additional opportunity cost of serving the low type off the platform. Indeed, it is not difficult to find parameters (e.g., λ large enough) for which the quality level in (5) is strictly positive but the quality in (4) is zero. Thus, without a platform, the firm would offer a low-quality product to the low type. However, for a sufficiently large platform, the low type is only offered a product on the platform, where the firm can make a different personalized offer to the high type. This is not the case off the platform, where the firm prefers to forego sales of the low product, to sell product $q(\theta_H)=\theta_H$ at a higher price on both channels. Indeed, when $\hat{q}(\theta_L)=0$, no consumer type receives any rent on or off the platform.

Finally, to determine the platform's fee t^*, we must specify what the on-platform consumers can do if the firm does not advertise. If these consumers can buy off the platform, then the optimal fee extracts the firm's extra profits relative to offering the Mussa and Rosen

(1978) menu (5) to all consumers. If they do not buy at all, the firm's outside option is scaled by $1-\lambda$, and the optimal fee is correspondingly higher.

Relative to this example, the full model we have described above shows how the platform can use search and information frictions (neither of which were present in this section) to manage information and competition and restore the monopoly environment it can exploit so well in this setting.

4 Search and Showrooming

We now analyze the general environment with many firms and a multidimensional continuum of types as introduced in Section 2. The platform provides a match-quality advantage based on superior information reflected by realized values (from distribution F) on the platform and less precise information (from distribution G) off the platform.

The objective of this section is to establish the equilibrium interaction between search and showrooming generated by the informational advantage of the platform. The informational advantage of the platform is valuable to firms that are willing to pay for the right to make a personalized offer to the consumer. In particular, on the platform, consumers and producers interact under symmetric information. However, if the winning firm wants the consumer to accept its personalized offer, this firm must induce the consumer not to buy from its own off-platform store, i.e., not to use the platform for "showrooming." Thus, the producer's ability to product steer and price discriminate on the platform is limited by the presence of the off-platform channel. Off the platform, each firm j faces "captive" consumers who value its product the most based on their beliefs m. However, trade off the platform occurs under asymmetric information, i.e., the firm must elicit the consumers' willingness to pay.

4.1 Search Patterns

We restrict attention to equilibria where all firms participate and where consumers expect symmetric menus, both on and off the equilibrium path. Search costs for consumers off the platform are significantly larger than those on the platform. In particular, the $1-\lambda$ off-platform consumers with beliefs m face search costs beyond the first firm. As a result, any consumer with beliefs m visits firm $j^{(1)} = \arg\max_j m_j$ only. This result does not depend on the magnitude of the search costs, as in the Diamond (1971) model. It then follows that if the platform has a strict informational advantage ($F \succ G$), each of the λ on-platform consumers infers that the advertised firm j^* maximizes her 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 firm j^* only.

Proposition 2 (Consideration Sets)
Every on-platform consumer θ compares the advertised firm's on-platform offer $(p_{j^*}(\theta), q_{j^*}(\theta))$ and its off-platform offer $(\hat{p}_{j^*}(\theta_{j^*}), \hat{q}_{j^*}(\theta_{j^*}))$ only.

For instance, a consumer who shops off-platform visits firm $j = 1$ when that firm yields the highest expected value m_1. If the same consumer shopped on the platform instead, she may see an ad by a different firm, $j = 2$. In equilibrium, the consumer will learn that $\theta_2 > \theta_1$ and would then either accept firm 2's offer or shop off platform from firm 2's menu. Figure 2 illustrates possible search behavior by a consumer whose initial beliefs m rank the firms differently relative to the beliefs of type θ.

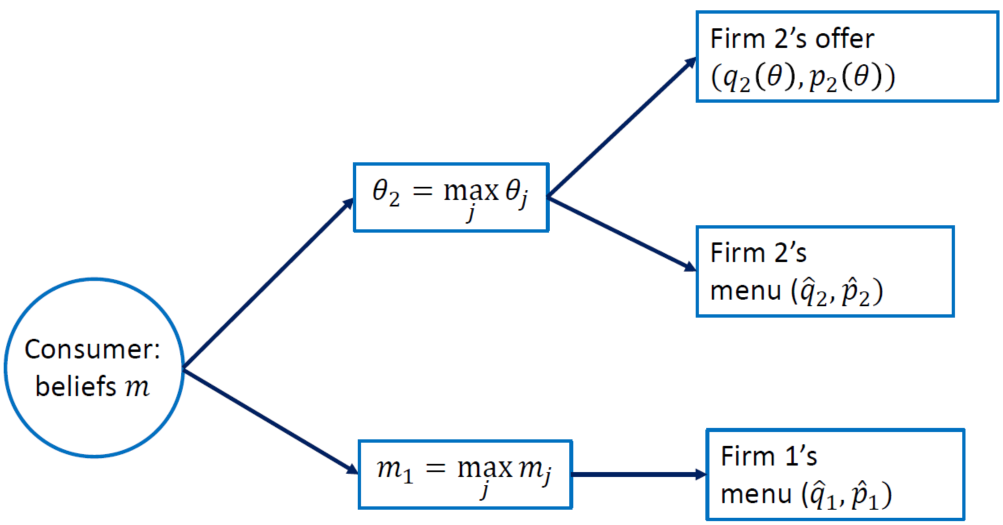

FIGURE 2: Sample Search Patterns

An important implication of Proposition 2 is that every consumer will (possibly incorrectly) buy from a competing firm if she does not see her favorite firm's ad. Thus, every firm realizes that participating in the platform's mechanism is necessary to access *any* of the on-platform consumers.

Indeed, when the platform has better information than the consumers, any symmetric equilibrium of the game is outcome-equivalent to a simpler model where each firm has a group of $(1 - \lambda)/J$ loyal customers with values distributed according to G^J. These customers shop off-platform from their favorite firm j only. The remaining λ consumers are currently not loyal shoppers for any brand (i.e., no firm 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 firm in their consideration set—the firm with the sponsored link or advertisement.

This alternative interpretation in terms of endogenous consideration sets requires the platform to hold an (arbitrarily small) informational advantage relative to the consumers. In Section 7, we show that, without an informational advantage, the platform does not control the consumers' outside options. Instead, the consumers' beliefs fully determine which firm they would visit off the platform.

4.2 Showrooming

The platform generates surplus by matching each θ to the product $q_j(\theta)$ that generates the largest match value. As information is symmetric between the consumer and the selected firm, the firm 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. The only limit on surplus extraction by the advertising firm is given by the "showrooming constraint," which is a necessary condition for firm j to make a sale on the platform:

$$\theta_j \cdot q_j(\theta) - p_j(\theta) \geq \max_{m_j} [\theta_j \cdot \hat{q}_j(m_j) - \hat{p}_j(m_j)] \text{ for all } \theta_j. \tag{6}$$

Because firm j offers an incentive-compatible menu off the platform, each on-platform consumer would also report her type truthfully if shopping off-platform. By Proposition 2, we know the consumer chooses between two products by the same firm, and the showrooming constraint (6) reduces to

$$U_j(\theta) := \theta_j q_j(\theta) - p_j(\theta)) \geq \theta_j \hat{q}_j(\theta_j) - \hat{p}_j(\theta_j)) =: \hat{U}_j(\theta). \tag{7}$$

The showrooming constraint prevents the selected firm from extracting the entire surplus of the on-platform consumers. Because the on-platform transaction takes place under symmetric information, it is immediate to show that each firm offers a single product to each consumer θ at the socially efficient quality level

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

The socially efficient quality provision maximizes both the profits from the ad and the probability of being chosen. At the same time, each firm must also offer the consumer a discount that satisfies (7) with equality. Therefore, if firm j offers the off-platform menu $(\hat{p}_j, \hat{q}_j)$ with the associated rent function $\hat{U}_j$, its on-platform profits from type θ are given by

$$\pi_j(\theta, \hat{U}_j) = \begin{cases} \theta_j^2/2 - \hat{U}_j(\theta_j), & \text{if } \theta_j > \theta_{-j}; \\ 0, & \text{otherwise.} \end{cases}$$

5 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 the platform and off the platform. Finally, we establish that the socially efficient matching rule is the revenue optimal mechanism for the platform to deploy.

By Proposition 2, in any symmetric equilibrium of our model, no consumer (off-platform or on-platform) searches past the first firm on the equilibrium path. Combining the on- and off-platform profits, each firm's maximization problem after accepting the platform's offer can be written as

$$\Pi_j^* = \max_{\hat{q},\hat{U}} (1-\lambda) \int_0^1 [\theta_j \hat{q}(\theta_j) - \hat{q}(\theta_j)^2/2 - \hat{U}(\theta_j)] G^{J-1}(\theta_j) \mathrm{d}G(\theta_j) \tag{8}$$

$$+ \lambda \int_0^1 [\theta_j^2/2 - \hat{U}(\theta_j)] F^{J-1}(\theta_j) \mathrm{d}F(\theta_j),$$

$$\text{s.t.} \quad \hat{U}(\theta_j) \geq 0, \tag{IR}$$

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

Maximizing (8) over rent and quality functions $\hat{U}$ and $\hat{q}$ using standard optimal-control methods, we obtain the following characterization of the optimal menus.

Proposition 3 (Symmetric Equilibrium Menus)

1. *The unique symmetric equilibrium quality levels are given by*

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

$$\hat{q}_j^*(\theta_j) = \max\left\{0, \theta_j - \frac{1 - \lambda F^J(\theta_j) - (1-\lambda) G^J(\theta_j)}{(1-\lambda) J G^{J-1}(\theta_j) g(\theta_j)}\right\}. \tag{9}$$

2. *The consumer's information rents are identical on- and off-platform,*

$$U_j^*(\theta) = \hat{U}_j^*(\theta) = \int_0^{\theta_j} \hat{q}_j^*(m_j)\, dm_j. \tag{10}$$

The equilibrium quality provision off the platform has several intuitive properties. First, the efficient quality is sold to each consumer i on the platform, on the basis of her favorite firm, i.e., $\max_j\{\theta_j\}$. Second, matching is inefficient off the platform because it is based on insufficient information, i.e., on beliefs m instead of fundamentals θ.

The consumer's private information off the platform requires the firms to resolve the efficiency vs. rent extraction trade-off. The information rents of each type m_j are as usual increasing in the quality level provided to all lower types. To resolve this trade-off, each firm j could offer the Mussa and Rosen (1978) tariff for the distribution of off-platform consumer types $G^J(m_j)$, 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 off the platform, each firm must also provide rents on the platform

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

Conversely, by limiting the off-platform rents, the firm 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, we can rewrite the equilibrium off-platform qualities (9) in Proposition 3 as

$$\hat{q}_j^*(\theta_j) = \underbrace{\theta_j - \frac{1 - G^J(\theta_j)}{JG^{J-1}(\theta_j)g(\theta_j)}}_{\text{Mussa and Rosen (1978) quality}} - \frac{\lambda}{1-\lambda}\frac{1 - F^J(\theta_j)}{JG^{J-1}(\theta_j)g(\theta_j)}, \tag{11}$$

where the first two terms identify the optimal quality level for the distribution of types $G^J(\theta_j)$. The last term captures the intuition that any rent given off-platform to type θ_j must also be given to all higher types *on* the platform.

Figure 3 illustrates the optimal quality schedule relative to the socially efficient allocation and to the monopoly benchmark as analyzed in Mussa and Rosen (1978).

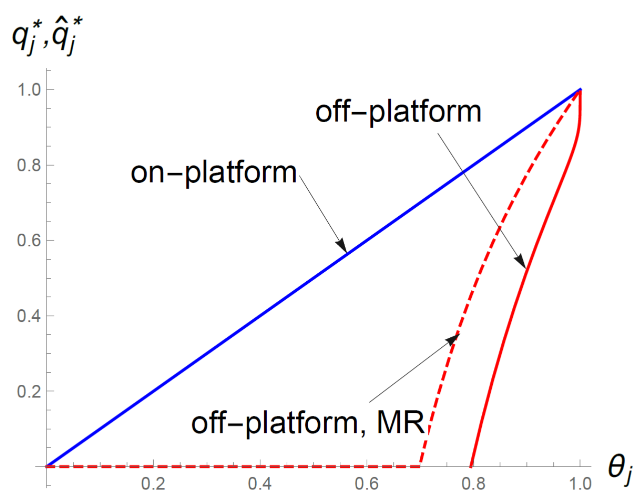

FIGURE 3: Quality Levels, $\lambda = 1/2, J = 5, G(m_j) = m_j, F(\theta_j) = \text{Beta}(\theta_j, 1/4, 1/4)$

The formulation of the optimal off-platform menu (11) allows us to establish several intuitive properties of the equilibrium. First, each type θ receives a higher quality level and pays a higher price on the platform than off the platform. However, while each type receives a better product at a higher price, each quality level q is sold at a lower price on the platform, i.e., $p(q) \leq \hat{p}(q)$ for all q. In other words, each firm is forced to introduce "on-platform only" discounts due to the threat of showrooming.

Figure 4 shows the nonlinear pricing schedules for the parameter values above. Note that for a set of low types, the nonlinear tariff is equal to the gross surplus generated by the efficient quality (i.e., $p(q) = q^2$), and for types that receive a positive rent off-platform, it is correspondingly lower.

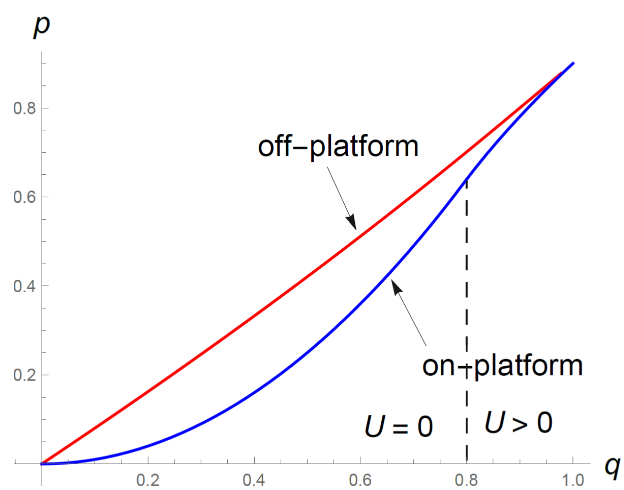

FIGURE 4: Nonlinear tariffs: every variety q is sold at a lower price on-platform.

Our results illustrate how advertising platforms that run managed campaigns face very different incentives than retail platforms that charge merchant fees. In the case of Amazon, for example, firms typically want consumers to showroom to avoid the fees. This is not the case in our model, as fixed-price contracts between advertisers and the platform eliminate the need to add most-favored-nation clauses.

Consumer Surplus An implication of Proposition 3 is that, on aggregate, consumer surplus is higher on the platform than off the platform. Indeed, for each type θ, we have

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

However, we also know that $F \succ G$. This implies that the highest order statistics satisfy $\mathbb{E}_{F^J}[\theta_j] > \mathbb{E}_{G^J}[\theta_j]$, and thus also

$$\mathbb{E}_{F^J}[U_j^*(\theta)] > \mathbb{E}_{G^J}[\hat{U}_j^*(\theta)],$$

because 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 her type, which enables better matching of products to preferences. However, the consumer does not necessarily benefit from the presence of the platform *in equilibrium*. Indeed, Proposition 5 considers the effects of a larger platform (λ) and finds that *all* consumers are worse off as λ increases.

Platform Revenue To examine the implications for the firms and the platform, we characterize the equilibrium fees paid by the firms. We define the firms' outside option $\bar{\Pi}_j$ as the Mussa and Rosen (1978) profits achievable for the off-platform consumers:

$$\bar{\Pi}_j \triangleq \max_{\hat{q},\hat{U}} (1-\lambda) \int_0^1 \left[\theta_j \hat{q}(\theta_j) - \hat{q}(\theta_j)^2/2 - \hat{U}(\theta_j)\right] G^{J-1}(\theta_j) \mathrm{d}G(\theta_j). \tag{12}$$

The platform's fee is then given by

$$t^* = \Pi_j^* - \bar{\Pi}_j.$$

Another way to interpret the platform's fee is the following: the firms are willing to give up all the on-platform profits to participate, but need to be compensated for distorting the off-platform menus away from the monopoly benchmark.

We can now show that this mechanism—whereby the platform charges a fixed payment for a managed advertising campaign that selects the socially optimal product varieties—maximizes the platform's profit.

Proposition 4 (Platform Revenue)
The managed campaign mechanism that advertises the product $q_j(\theta)$, which maximizes the social surplus, also maximizes the platform's profit.

A sketch of the proof of this result is as follows. The managed campaign mechanism we have considered maximizes the firms' and the platform's joint profits across both markets (on- and off-platform). Furthermore, the platform's fee extracts all the firms' surplus over and above their exogenous outside option. Thus, this mechanism maximizes the platform's revenue.

6 Platform Size, Information, and Competition

We established the nature of the equilibrium and the basic welfare properties for a given market structure. Next, we investigate how the size of the platform in terms of its market share and in terms of its informational advantage affects the welfare and distribution of the social surplus. We then analyze how an increase in competition in terms of the number of competing firms affects the welfare outcomes on and off the platform. We begin with the size of the platform in terms of its market share λ.

Platform Size The opportunity cost of serving consumers off the platform increases as the platform becomes (exogenously) larger. Intuitively, the information rents of the off-platform consumers must also be paid to a mass λ of on-platform consumers. This should lead to further distortions in the off-platform quality levels. We formalize this intuition in Proposition 5. Recall that the individual types θ_j are distributed over the unit interval (as a representation of any compact interval). Thus, the highest possible value of θ_j is 1.

Proposition 5 (Platform Size)

1. *The equilibrium quality levels $\hat{q}_j^*(\theta_j)$ are decreasing in λ for all $\theta_j < 1$, and the information rents $\hat{U}_j^*(\theta_j)$ are decreasing in λ for all θ_j.*

2. *For every $\theta_j < 1$, there exists $\bar{\lambda} < 1$ such that $\hat{q}_j^*(\theta_j) = 0$ for all $\lambda \geq \bar{\lambda}$.*

In Figure 5, we illustrate how the off-platform quality provision changes as the market share λ of the platform increases. The allocation on the platform remains unchanged and is given by the socially efficient quality provision. However, as the platform grows larger, each firm attempts to minimize the information rents on the platform and in turn renders the menu off the platform less attractive. Thus, for every type θ_j, the equilibrium quality-match off the platform $\hat{q}_j^*(\theta_j)$ decreases, the price per unit of quality increases, and the consumer surplus $\hat{U}_j^*$ off the platform decreases as the size of the platform increases.

Information Precision An important question for welfare and policy is whether consumers or the platform benefit from more precise information, and how these results change when the platform's size and information precision are related. To capture the impact of better information on the platform, consider the true type distribution F and the distribution

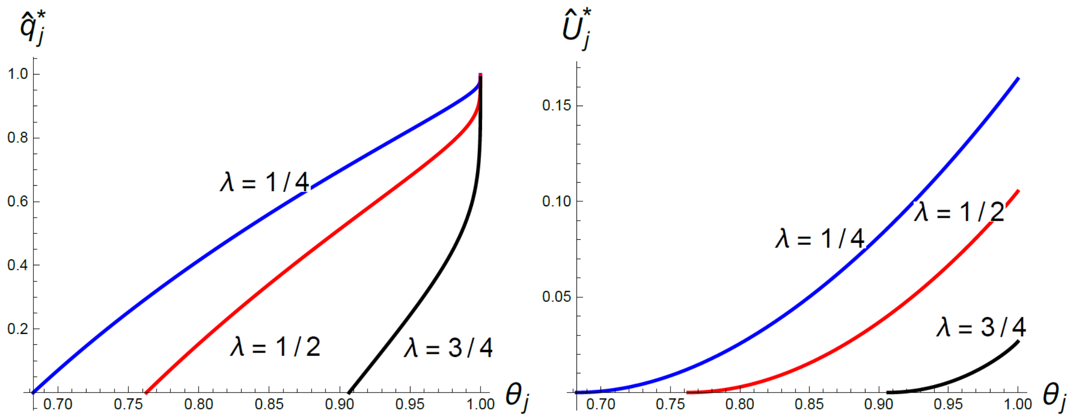

FIGURE 5: Off-Platform Menus, $J = 3, G(m_j) = m_j, F(\theta_j) = \text{Beta}(\theta_j, 1/4, 1/4)$.

of prior consumer beliefs G. Now suppose the platform observes an imperfectly informative signal about each on-platform consumer. These signals are strictly more informative than the consumer's beliefs and they supersede the consumer's initial information.

We assume that the platform's posterior beliefs about each θ_j are distributed according to $\widehat{F}$. We then say the platform is endowed with information $\widehat{F}$, which satisfies

$$F \succ \widehat{F} \succ G.$$

Through the managed campaign mechanism, the platform shares information $\widehat{F}$ with all the firms, i.e., it enables prices and ads to condition on this information. It also shares its posterior beliefs about θ_{j^*} with any consumer who is shown an ad by firm j^*.

We now compare the equilibrium outcomes for a fixed λ and G under two on-platform distributions $\widehat{F} \in \{F_1, F_2\}$ with $F_1 \succ F_2 \succ G$, such as those in Figure 6.

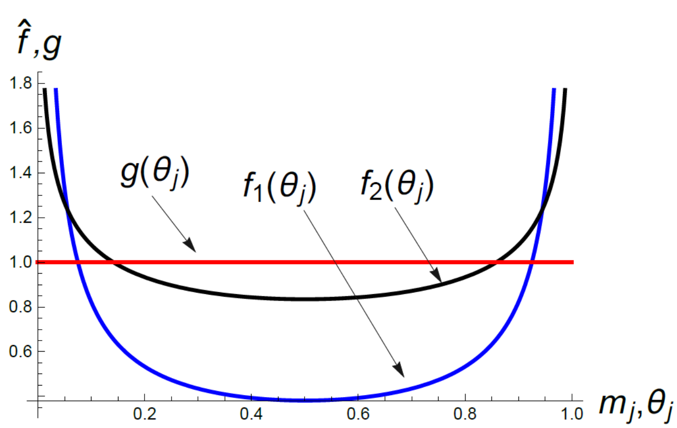

FIGURE 6: $G(m_j) = m_j, F_1(\theta_j) = \text{Beta}(\theta_j, 1/4, 1/4), F_2(\theta_j) = \text{Beta}(\theta_j, 3/4, 3/4)$.

We denote the equilibrium off-platform menus by $\hat{q}_j^*(\theta_j; F_1)$ and $\hat{q}_j^*(\theta_j; F_2)$ and the platform's fees by $t^*(F_1)$ and $t^*(F_2)$, respectively. For our next result, we strengthen the ranking of the two distributions by assuming that the more informative distribution first-order stochastically dominates the less informative distribution *over the range of types who receive a positive quality off the platform.*

Proposition 6 (Information Precision)
Let $F_1 \succ F_2$ and assume $F_1(\theta_j) \leq F_2(\theta_j)$ for all θ_j such that $\hat{q}_j^*(\theta_j; F_2) > 0$.

1. *The equilibrium menus satisfy $\hat{q}_j^*(\theta_j; F_1) \leq \hat{q}_j^*(\theta_j; F_2)$ for all θ_j, with strict inequality whenever $\hat{q}_j^*(\theta_j; F_2) \in (0, 1)$.*
2. *The platform's fees satisfy $t^*(F_1) > t^*(F_2)$.*

In Propositions 5 and 6, we separately traced out the implications of a larger market and a more informed platform. We might in fact conjecture that a larger market share generates more user data for the platform and hence more precise estimates regarding the willingness to pay of the consumer. Thus, the *combined* effect of a larger platform is stronger than each of its components—as more consumers join the platform, the opportunity cost of off-platform quality provision increases, and as information becomes more precise, the relevant range of the type distribution shifts in a first-order stochastic dominance sense. The resulting larger (absolute and relative) mass of high types induces firms to charge higher prices on and off the platform at the expense of market coverage off the platform.

Despite the more severe distortions off the platform, the firms stand to gain from a more informative distribution of on-platform types under the assumptions of Proposition 6. This is intuitive because the profit level obtained on the platform is strictly increasing in θ_j. Therefore, a shift in the distribution of on-platform types from $F_2(\theta_j)$ to $F_1(\theta_j)$ would improve the firms' profits even if they did not modify their off-platform menus $q_j^*(\theta_j, F_2)$. If seller j does not participate, however, its outside option $\bar{\Pi}_j$ in (12) does not depend on the precision of the platform's information—it is solely determined by the monopoly profits under type distribution G^J. The platform can therefore capture the entire additional producer surplus it generates thanks to more precise information.

The effects of the platform's information precision on consumer surplus amount to asking whether a less informative distribution $\widehat{F}$ is preferable to the true distribution of types F. The main tension is that less precise information yields inefficiently noisy matching and quality provision but also reduces quality distortions off the platform.

In Figure 7, we examine the consequences of information precision on consumer surplus, separately on and off the platform. We focus on a parametrized example with Beta distributions. In this example, similar to Figure 6, consumer valuations θ_j are distributed according to a symmetric Beta distribution with parameter $1 - \bar{a}$. Consumer beliefs m_j are uniformly and independently distributed over $[0, 1]$ for each j (i.e., a Beta with $\bar{a} = 0$). The platform reveals to all the firms and to the consumer a signal that induces beliefs over valuations that follow a symmetric Beta distribution with parameter $1 - a$, where $a \in [0, \bar{a})$. In the right panel, we note that with fiercer competition ($J = 12$), on-platform consumers benefit from greater information precision, while the diminishing information rents (for each type) dominate both off platform and on platform if the number of firms is low.

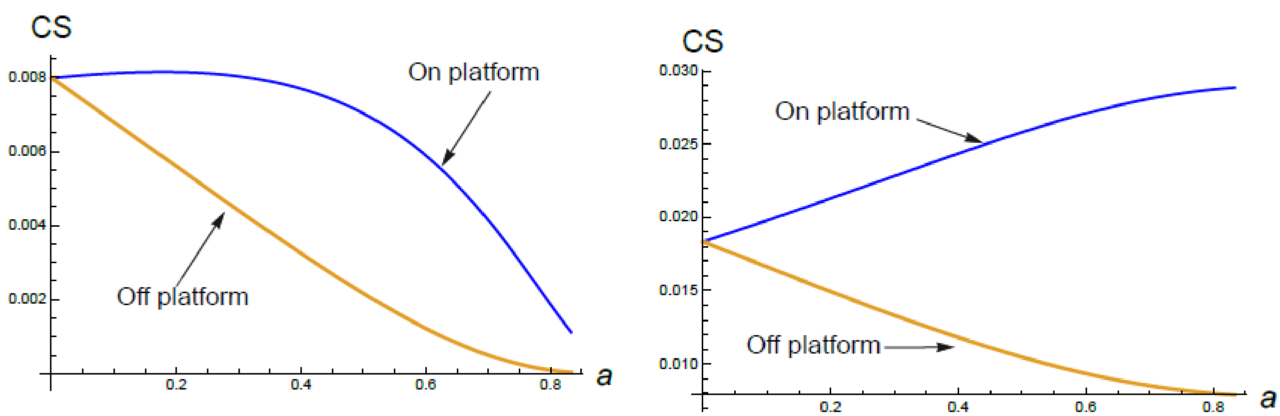

FIGURE 7: Consumer Surplus, $\lambda = 4/5, G(m_j) = m_j, \bar{a} = 5/6, J = 2$ (L) and $J = 12$ (R).

Number of Firms As the number of firms increases, a larger number of draws J for each type θ_j improves the distributions F^J and G^J in the likelihood-ratio order. This leads to lower information rents: in the limit, a firm will know that every consumer who shops on its site, or receives its ads, has a valuation arbitrarily close to 1, and therefore information rents vanish. This result is a direct implication of the Diamond (1971) model adapted to our setting. We illustrate this result for the benchmark case of an off-platform market only (i.e., $\lambda = 0$) in Figure 8.

Relative to the Diamond model, quality distortions decrease faster in our setting for lower types and slower for higher types. This effect is due to the interaction of showrooming and the different distributions of types. In particular, one can show that the additional distortion term in the equilibrium quality (11), i.e.,

$$\frac{\lambda}{1-\lambda} \frac{1 - F^J(\theta_j)}{JG^{J-1}(\theta_j)g(\theta_j)}$$

is decreasing in J when θ_j is close to 1. For a small number of firms, this implies that high-θ_j types receive a higher quality as J increases while low types receive a lower quality.

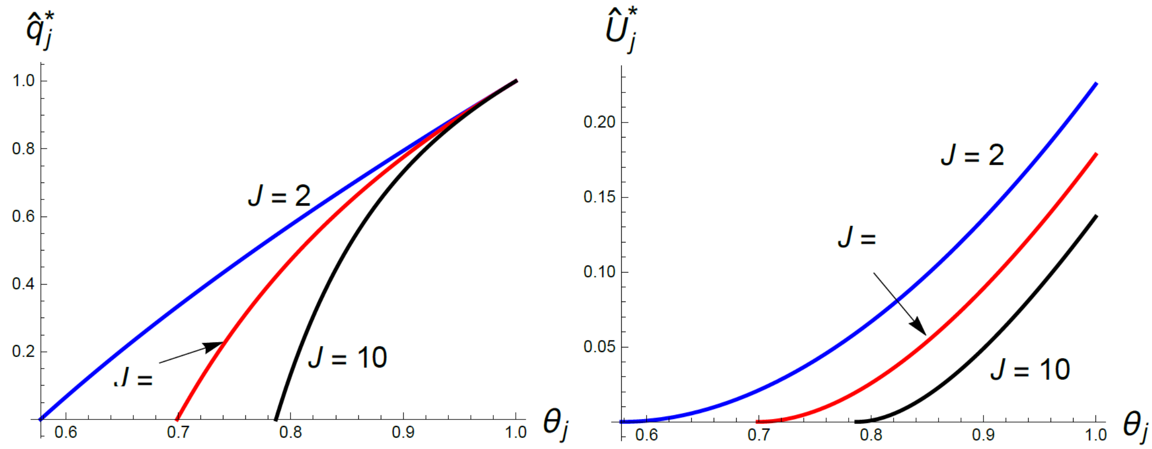

FIGURE 8: Off-Platform Menus, $\lambda = 0, G(m_j) = m_j, F(\theta_j) = \text{Beta}(\theta_j, 1/4, 1/4)$.

As Figure 9 illustrates, this effect may not be sufficient to generate a larger rent for any type.

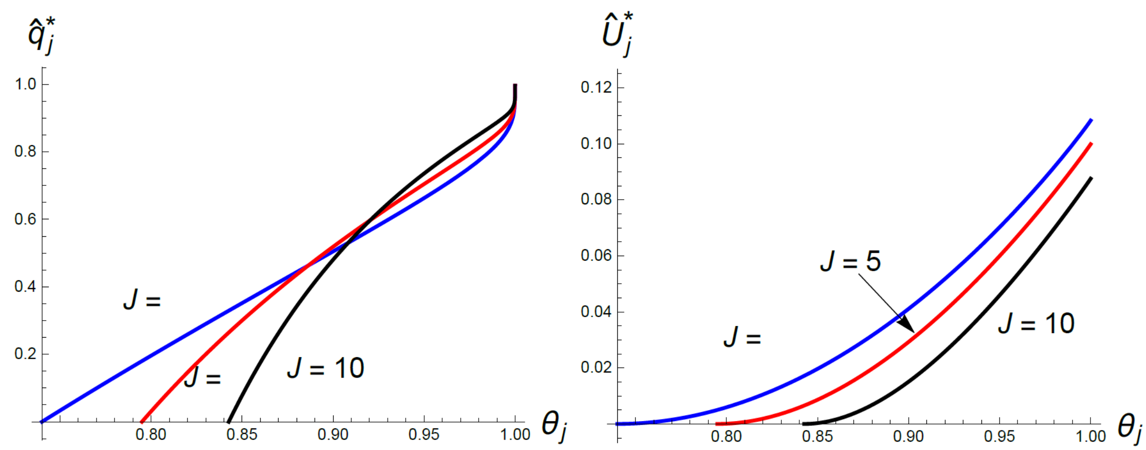

FIGURE 9: Off-Platform Menus, $\lambda = 1/2, G(m_j) = m_j, F(\theta_j) = \text{Beta}(\theta, 1/4, 1/4)$.

Furthermore, Proposition 7 shows that as J grows large, every type's quality allocation eventually decreases in the number of firms.

Proposition 7 (Number of Sellers)

1. *For every $\theta_j < 1$, the equilibrium quality $\hat{q}_j^*(\theta_j)$ and information rent $\hat{U}_j^*(\theta_j)$ are decreasing in J if J is large enough.*

2. *For every $\theta_j < 1$, there exists $\hat{J}$ such that $\hat{q}_j^*(\theta_j) = 0$ for all $J \geq \hat{J}$.*

To conclude this section, we examine the impact of the size of the platform λ and of the number of firms J on all parties' surplus levels. An immediate consequence of Propositions

5 and 7 is that expected consumer surplus on-platform and off-platform is always decreasing in λ, and is eventually decreasing in J too. At the same time, the platform's revenue is increasing in both λ and J. Furthermore, as J grows without bound, the platform captures the entire (first best) social surplus it creates. Intuitively, the consumers have no information rents (as the highest type component is converging in probability to 1), and therefore firms need not distort the off-platform menus when they participate in the platform's mechanism. Figure 10 illustrates the platform revenue, consumer surplus, producer surplus (i.e., the outside option $\bar{\Pi}_j$), and the profits generated on the platform for various numbers of firms.

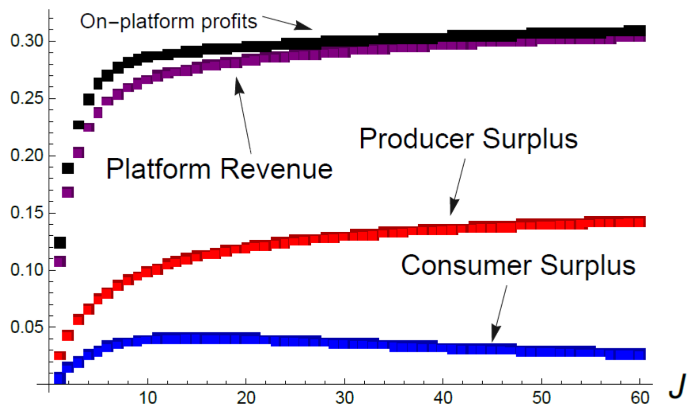

FIGURE 10: Surplus Levels, $\lambda = 2/3, G(m_j) = m_j, F(\theta_j) = \text{Beta}(\theta_j, 1/3, 1/3)$.

7 Information and Search Frictions

In this section, we deepen our understanding of the platform's bargaining power by exploring the role of informational advantages and search frictions. In particular, we assume that *at the onset of the game*, every on-platform consumer has as much information about her willingness to pay for the firms' product lines as the platform. Next, we consider what happens when the platform offers information about the off-platform prices by means of organic search links.

7.1 Symmetric Information

Suppose all consumers who visit the platform know their type θ, while the consumers off the platform remain imperfectly informed with beliefs m. One reason this may be the case is that the platform provides organic information in the form of product rankings and recommendations. We deal with organic links separately in Section 7.2, as they require more careful modeling. For now, we simply introduce the assumption of symmetric complete information on the platform.

As we show in Proposition 8, informed consumers limit the platform's ability to steer their search behavior, which in turn reduces the amount the platform can charge the firms.

Proposition 8 (Symmetric Information)
Suppose the consumers on the platform have complete information about θ at the onset of the game. The equilibrium quality levels on the platform remain socially efficient, and the platform's fees are strictly lower relative to when the platform has superior information about θ (as in Proposition 3).

Because consumers know their types and maintain those beliefs off the equilibrium path, they visit their favorite firm off-platform regardless of the displayed firm. This means that each firm j can refuse to pay the platform's fee and poach any consumer with $\theta_j = \max_k \theta_k$ by offering them an off-platform information rent $\hat{U}_j(\theta_j)$ above the level $\hat{U}_j^*(\theta_j)$ in (10) that is offered by its competitors in equilibrium.[6] The deviating firm solves the following problem:

$$\hat{\Pi} \triangleq \max_{\hat{q},\hat{U}} \int_0^1 [\theta_j \hat{q}(\theta_j) - \hat{q}(\theta_j)^2/2 - \hat{U}(\theta_j)] \begin{bmatrix} (1-\lambda)G^{J-1}(\theta_j)g(\theta_j) \\ +\lambda F^{J-1}(\theta_j)f(\theta_j) \end{bmatrix} \mathrm{d}\theta_j \tag{13}$$

$$\text{s.t. } \hat{U}(\theta_j) \geq \hat{U}_j^*(\theta_j). \tag{14}$$

The baseline model's equilibrium rent function (10) satisfies the constraint (14) and yields a strictly larger profit. Therefore, we immediately conclude that the firms' outside option with known types exceeds the outside option $\bar{\Pi}$ of the baseline model characterized in (12).

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

$$\hat{q}(\theta_j) = \max\left\{0, \theta_j - \frac{1-\lambda F^J(\theta_j) - (1-\lambda)G^J(\theta_j)}{\lambda J F^{J-1}(\theta_j)f(\theta_j) + (1-\lambda)JG^{J-1}(\theta_j)g(\theta_j)}\right\}. \tag{15}$$

The quality in (15) is larger than the equilibrium $\hat{q}_j^*(\theta_j)$ in (9) and yields higher utility to the consumers. Thus, constraint (14) does not bind in the optimal deviation—the best off-platform menu for firm j exceeds the utility level offered to its favorite consumers by all other firms. Finally, note that by posting a more attractive menu off the platform, the deviating firm cannot attract any consumers who value its competitors' products more than its own. This is because those consumers still face search costs off the platform and will not know that the deviating firm has in fact lowered its prices.

[6] There is no profitable deviation for a seller who accepts the platform's offer. Therefore, the menus on the equilibrium path have the same functional form as in the baseline case.

Proposition 8 has important implications for a platform's choice of information policy. For this purpose, let us consider for the moment a model where there is symmetric information across the consumers and the platform with one specific feature. Namely, the platform has no additional data or insights beyond the information of the individual consumer. The platform therefore allows firms to target ads based solely on input from consumers. Now, the consumers are imperfectly informed as represented by the belief distribution G. Thus, one may think that a comparison of the equilibrium outcome under F and under G, would reduce to setting $\widehat{F} = G$ in the language of Proposition 6 in the previous Section 6. After all, there we considered how a decrease in information from F to $\widehat{F}$ would affect the allocation and the revenue of the platform.

However, Proposition 8 above shows that the equilibrium menus are qualitatively different when the on-platform consumers know their types. The reason for the qualitative difference is that the platform loses its ability to steer the consumer. In the absence of additional information, the platform cannot grant monopoly power to any firm by displaying its advertisement and recommending its products. Without additional information, each consumer continues to evaluate the different products independent of the recommendation implicit in the ad. In turn, the value to a firm of having its advertisement shown decreases, as does its willingness to pay for the platform's services. We thus have the following important implication.

Corollary 1 (Value of Additional Information)
The revenue of the platform from arbitrarily small additional information is strictly positive:

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

In particular, the platform's fees in the case $\widehat{F} = G$ are strictly lower than in the limit as $\widehat{F} \to G$ of Proposition 6 in. In the latter case, the platform still holds considerable bargaining power vis-à-vis the firms because it sells the exclusive right to trade with the consumer. In the former case, the platform cannot prevent a seller from reaching its favorite consumers.

To summarize, not only are the platform's fees increasing in the precision of its information (Proposition 6), but the equilibrium fees jump up as soon as the platform has any informational advantage relative to the consumers. This means that the platform benefits discontinuously from investing in acquiring superior information. We have not modeled information acquisition, but the results in Propositions 6 and 8 suggest that some degree of information asymmetry will be optimal for the platform for a large class of cost functions.

7.2 Organic Links

In the equilibrium of our baseline model, the consumer correctly infers that the advertised brand provides her with the highest value and only considers that brand's on- and off-platform offers. In practice, platforms manage both the provision of information and the degree of competition on their web pages.

We now consider an extension to our model, whereby the platform displays "organic links" that make all off-platform prices observable to consumers. This implies that any on-platform buyer can buy from *any* off-platform firm. Therefore, each firm can undercut its competitors by offering a larger utility level to off-platform consumers.

The timing of the game is as follows and is summarized in Figure 11 below.

1. The platform announces the participation fee t.
2. Firms simultaneously set off-platform products $\hat{q}_j(m_j)$ and prices $\hat{p}_j(m_j)$, choose whether to pay the fee, and if so, choose what products $q_j(\theta)$ and prices $p_j(\theta)$ to advertise.
3. For each consumer θ, the platform displays a single advertised product $q_j(\theta)$ at price $p_j(\theta)$, *in addition to all off-platform products and prices.*
4. The consumer learns her full type θ and can buy the advertised product or any off-platform product.

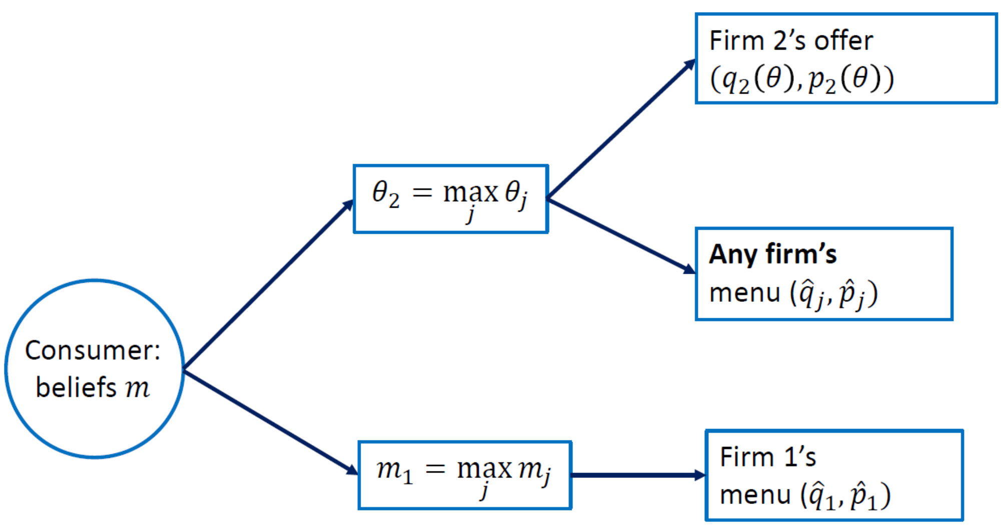

FIGURE 11: On-platform consumers can access every off-platform menu.

In this version of the model, it is still optimal for firms to advertise the socially efficient product varieties and to set prices that make the showrooming constraints bind. Thus, the platform assigns the sponsored link to each consumer's favorite firm. Nonetheless, off-platform menu competition can affect market shares on the platform. In particular, starting from a symmetric strategy profile, any firm can now post lower prices off the platform and "poach" some consumers on the platform. These are consumers who would not learn of the lower off-platform prices without the platform's organic links.

To compute the firms' market shares of on-platform types, consider the off-platform information rents $\{\hat{U}_j(\theta_j)\}_{j=1}^N$. The outside option of the on-platform buyer θ is given by $\max_j\{\hat{U}_j(\theta_j)\}$. For a symmetric strategy profile by firms $-j$ and for each type θ_j, define the indifferent type $\theta^*_{-j}(\theta_j)$ as

$$\hat{U}_{-j}(\theta^*_{-j}) = \hat{U}_j(\theta_j) \tag{16}$$

Firm j's best-response problem can be written as

$$\begin{aligned}\max_{\hat{q},\hat{U}}\,(1-\lambda)\int_0^1 &\underbrace{[\theta_j\hat{q}(\theta_j)-\hat{q}(\theta_j)^2/2-\hat{U}(\theta_j)]}_{\text{off-platform sales}}G^{J-1}(\theta_j)g(\theta_j)\mathrm{d}\theta_j \\ +\lambda\int_0^1 &\underbrace{(\theta_j^2/2-\hat{U}(\theta_j))}_{\text{on-platform sales}}\min\{F^{J-1}(\theta^*_{-j}(\theta_j)),F^{J-1}(\theta_j)\}f(\theta_j)\mathrm{d}\theta_j \\ +\lambda\int_0^1 &\underbrace{[\theta_j\hat{q}(\theta_j)-\hat{q}(\theta_j)^2/2-\hat{U}(\theta_j)]}_{\text{off-platform sales}}\max\{0,F^{J-1}(\theta^*_{-j}(\theta_j))-F^{J-1}(\theta_j)\}f(\theta_j)\mathrm{d}\theta_j,\end{aligned} \tag{17}$$

where θ^*_{-j} is given in (16).

The first term in (17) captures the off-platform consumers. The second term captures the sales to on-platform consumers for which firm j offers both the highest utility level $\hat{U}_j$ and the highest marginal valuation θ_j. The third term, whenever positive, captures on-platform consumers with $\theta_j \neq \max_k \theta_k$ to whom firm j offers the highest utility. Firm j is not advertised to these consumers, but they can nonetheless showroom and buy from firm j off the platform. Firm j's problem can therefore be described as follows: it can poach any consumer and make an off-platform sale by offering low enough prices; if, in addition, the consumer values its product the most, firm j wins the sponsored link and makes a more profitable sale on the platform (i.e., a higher quality product at a higher price).

The difference in profit margins on- vs. off-platform has implications for firms' incentives to expand their market shares. Intuitively, suppose firm j is charging higher prices than its competitors. Then, it will not be able to poach any consumers and will only sell to (a subset of) consumers who value its products the most. Raising prices further then yields a large marginal loss because the firm would be giving up on-platform sales, which occur at higher prices. Conversely, if firm j is already the low-price firm, the marginal benefit of raising prices is high, because market shares are large and the marginal consumer is buying off-platform, where markups are lower.

The symmetric equilibrium quality levels of this game can be characterized through a system of differential equations, as in Bonatti (2011). In Proposition 9, we compare the game with organic links to our baseline setting.

Proposition 9 (Equilibrium with Organic Links)

1. *The equilibrium quality and utility levels $\hat{q}_j^*(\theta_j)$ and $\hat{U}_j^*(\theta_j)$ are weakly higher for all θ_j with organic links, relative to the baseline model.*

2. *Firms' profits are lower and their outside options are higher with organic links, relative to the baseline model.*

To establish the next result, we consider the symmetric equilibria of the subgame following the platform's announcement of a fee t. We show that competition among the firms is fiercer in any continuation equilibrium. In particular, quality and utility levels are higher and firms' profits are lower than without organic links.

In the symmetric Perfect Bayesian equilibrium of the full game, the platform sets the highest fee t^* such that there exists a continuation equilibrium with full firm participation. This is not a straightforward computation because the firms' outside options are determined as the best response to the equilibrium strategies of a deviating (nonparticipating) firm. Nonetheless, the firms' outside options in any continuation equilibrium are higher than in the baseline model. Therefore, the platform must charge a lower fee.

Intuitively, the presence of organic information benefits consumers but reduces the platform's ability to restrain competition and extract surplus from firms. In a symmetric equilibrium, each firm's market segment consists of all consumers who like its 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 firms are lower, and consumer surplus is higher.

However, the firms' profits net of the platform fee are equal to the value of their outside option. With organic links, any firm can respond to competitors' prices without participating in the mechanism. Let θ_{-j}^* be given by (16), with $\hat{U}_{-j} = \hat{U}_j^*$. Because all of its sales necessarily happen off the platform, a deviating firm j can then obtain profits of

$$\tilde{\Pi}_j \triangleq \max_{\hat{q},\hat{U}} \int_0^1 \left[\theta_j \hat{q}(\theta_j) - \hat{q}(\theta_j)^2/2 - \hat{U}(\theta_j)\right] \left[\begin{array}{c} (1-\lambda)\, G^{J-1}(\theta_j) g(\theta_j) \\ +\lambda F^{J-1}(\theta_{-j}^*(\theta_j)) f(\theta_j) \end{array}\right] \mathrm{d}\theta_j. \qquad (18)$$

Unlike in the baseline model, each deviating firm has the opportunity to win over *some* on-platform consumers for which $\theta_j \geq \theta_{-j}$. The outside option $\tilde{\Pi}_j$ in (18) then exceeds the value $\bar{\Pi}_j$ defined in (12). In other words, the firms' outside options are higher with organic links than without, and the platform's fees are correspondingly lower.

Finally, with known on-platform types and no organic links, the deviating firm wins *all* on-platform consumers for which $\theta_j \geq \theta_{-j}$. Thus, the outside option $\hat{\Pi}_j$ defined in (13) is even higher than $\tilde{\Pi}_j$ in (18). Because the equilibrium with organic links does not change if consumers know their types, the platform may be able to raise its fees by posting organic links in addition to the sponsored link when facing informed consumers.

8 Privacy Protection

Thus far, we have imposed no restrictions on the platform's ability to share data θ with the firms. In practice, the amount of data sharing can be limited both by regulation and by the platform's design choices. In this section, we assume that the valuation for the advertised product θ_j is shared with the on-platform consumer but not with the firms. In contrast, the platform allows firms to condition their offers on the *ranking* of consumer valuations θ_j only. Thus, the platform targets ads at the level of a *cohort* of consumers, and each consumer within a cohort ranks the J firms in the same way.[7]

Under this regime, the efficient matching of firms to consumers is still feasible. However, on-platform consumers have residual private information about their tastes. Unlike in the baseline case, with cohort-based ads, firm j only knows the distribution of the consumer's type based on the order statistics implied by her cohort. As a result, the advertising firm offers a menu of products both on and off the platform.

The symmetric equilibrium menus under cohort-based ads can be obtained as the solution to two linked screening problems, one per sales channel. In this problem, the showrooming constraints act as type-dependent participation constraints. To solve this problem, we strengthen the ranking of the on- and off-platform distributions by assuming that the on-platform distribution likelihood-ratio dominates the off-platform distribution over the range of types that receive a positive quality. In particular, define the (Myerson) virtual values for distribution $F^J(\theta_j)$ as

$$\phi_F(\theta_j) := \theta_j - \frac{1 - F^J(\theta_j)}{JF^{J-1}(\theta_j)f(\theta_j)},$$

and similarly $\phi_G(\theta_j)$ for distribution $G^J(\theta_j)$.

[7] This governance regime is reminiscent of the recent Google Privacy Sandbox proposals to replace third-party cookies.

Proposition 10 (Coarse Targeting)
In the unique symmetric equilibrium, each firm offers quality levels

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

if and only if $F^J \succ_{lr} G^J$ *for all* θ_j *such that* $\min\{\phi_{F^J}(\theta_j), \phi_{G^J}(\theta_j)\} \geq 0$.

Proposition 10 shows that if the distribution of the highest θ_j dominates that of the highest m_j in likelihood ratio (over the relevant range), then each firm offers the same menu to consumers both on and off the platform. In this menu, the equilibrium quality schedule is the same as in (15), 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 θ 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 type θ.

A critical implication of Proposition 10 is that all consumers receive higher information rents relative to the baseline setting because of the greater quality provision off the platform. The total surplus can also be higher as a consequence of greater off-platform quality, although on-platform quality is lower. Finally, as the firms' outside options are unchanged relative to the baseline model (i.e., they are given by $\bar{\Pi}_j$), the platform's equilibrium fees are unambiguously lower.

9 Conclusion

We have developed a flexible framework of product-line pricing where consumers can buy on and off a digital platform. The platform generates value by showing consumers their favorite brand and monetizes its superior information by selling selective access to their attention. The producers' ability to price discriminate on the platform—and hence their willingness to pay for advertising—is limited by the presence of the off-platform channel through the showrooming mechanism.

We have shown that consumer surplus on and off the platform is driven by the information rents consumers obtain off-platform and by the availability of organic information. In particular, the digital platform can leverage arbitrarily small information and search frictions to capture a substantial fraction of the surplus it generates. The growth of a platform's database (through more consumers or better data) reduces each consumer's outside option and leads to higher prices.

Our model is stylized and simplified along many dimensions, some of which are readily amenable to extensions. For example, differentiated products are also heterogeneous in their relative on- vs. off-platform presence. Furthermore, our baseline model is one of "perfect steering"—the advertised firm competes only with its own off-platform price-quality schedule. In practice, however, it is standard practice for multiple advertisements—often for related products—to be shown to a single consumer. The competition that ensues could then be modeled as a version of Bergemann, Brooks, and Morris (2021) with quality-differentiated products. Overall, our paper emphasizes the need for a more complete picture of the role of data in managing competition in the Internet economy.

Appendix

This Appendix contains the proofs of all our results.

Proof of Proposition 1. The derivation of the low type's optimal quality is standard. We substitute the expression for the binding incentive compatibility constraint for the high type (3) into both the on-platform and off-platform profits in objective (2). Differentiating with respect to $\hat{q}(\theta_L)$ yields the result. ■

Proof of Proposition 2. Each consumer θ with initial beliefs m learns her true valuation θ_{j^*} for the advertised firm j^*. Because the signal s has full support, the consumer believes that $j^* = \arg\max_j \theta_j$ in any symmetric equilibrium, regardless of her initial beliefs m. This holds both on and off the equilibrium path, i.e., if a firm deviates and does not participate in the platform's mechanism, in which case the consumer is shown an ad by a different firm. Furthermore, on and off the equilibrium path, the consumer expects all firms to post identical menus, and she knows the rent function $\hat{U}_j(\theta_j)$ is strictly increasing in θ_j. Therefore, the consumer searches for the advertised firm's off-platform prices. She does not search any further and does not learn any other firm's prices. Finally, because the menu off the platform is incentive compatible, it is sufficient for consumer θ to compare the two items $q_{j^*}(\theta_j)$ and $\hat{q}_{j^*}(\theta_j)$. ■

Proof of Proposition 3. Firm j's gross profits (i.e., after paying the platform's fee) can be written as

$$\max_{\hat{q},\hat{U}} (1-\lambda)\int_0^1 [\theta_j\hat{q}(\theta_j) - \hat{q}(\theta_j)^2/2 - \hat{U}(\theta_j)]G^{J-1}(\theta_j)\mathrm{d}G(\theta_j) \tag{19}$$
$$+\lambda\int_0^1 [\theta_j^2/2 - \hat{U}(\theta_j)]F^{J-1}(\theta)\mathrm{d}F(\theta),$$
$$\text{s.t. } \hat{U}'(\theta_j) = \hat{q}(\theta_j), \text{ and } \hat{U}(\theta_j) \geq 0 \text{ for all } \theta_j \in [0,1].$$

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

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

At a symmetric equilibrium, the necessary conditions are given by

$$\begin{aligned} (1-\lambda)\left(\theta_j - \hat{q}(\theta_j)\right) G^{J-1}(\theta_j) g(\theta_j) + \hat{\gamma}(\theta_j) &= 0, \\ -(1-\lambda) G^{J-1}(\theta_j) g(\theta_j) - \lambda F^{J-1}(\theta_j) f(\theta_j) + \hat{\gamma}'(\theta_j) &= 0, \\ \hat{\gamma}(1) &= 0. \end{aligned}$$

Integrating, we obtain

$$\hat{\gamma}(\theta_j) = \frac{1}{J}\left((1-\lambda) G^J(\theta_j) + \lambda F^J(\theta_j) - 1\right). \tag{20}$$

Therefore, the equilibrium quality level is given by

$$\hat{q}_j^*(\theta_j) = \theta_j - \frac{1-(1-\lambda) G^J(\theta_j) - \lambda F^J(\theta_j)}{(1-\lambda) J G^{J-1}(\theta_j) g(\theta_j)} \tag{21}$$

whenever this expression is nonnegative and nil otherwise. This is the expression in (9), which ends the proof. ■

Proof of Proposition 4. We argue the optimality of the efficient-selection advertising mechanism in two steps. First, by Proposition 2, any firm j that does not participate does not make any sales to the consumers on the platform. This is because every consumer will see an ad by a different firm $k \neq j$ and will only consider on- and off-platform offers by firm k. Therefore, every firm's outside option consists of offering the optimal Mussa and Rosen (1978) menu to its off-platform consumers. This yields the profit level $\bar{\Pi}_j$ in (12), which is a fixed outside option independent of the menus posted by the other firms.

Second, consider the coalition of all firms and the platform. The coalition's profits are maximized by matching each on-platform consumer θ to firm $j^* = \arg\max_j \theta_j$, each off-platform consumer m to $\hat{\jmath} = \arg\max_j m_j$, and by maximizing the firm's profits with respect to the on-platform offers (q, U) and the off-platform menus $(\hat{q}, \hat{U})$.

The solution to the coalition profit-maximization problem coincides with the equilibrium outcome of the efficient-selection mechanism. In that mechanism, each firm maximizes its profits by offering the socially efficient quality $q_j(\theta) = \theta_j$ to each on-platform consumer θ. As a result, the platform assigns each firm to the consumers that value its products the most. Thus, the firms-platform coalition's profits are maximized by the equilibrium menus that solve problem (19). Because the platform's fee extracts the entire producer surplus in excess of the fixed outside option (12), no other mechanism can generate greater revenue for the platform. ■

Proofs of Propositions 5–7. These results can all be obtained by differentiating expression (9) for the equilibrium quality sold to each off-platform type θ_j. In particular, whenever it is strictly positive, the equilibrium $\hat{q}_j^*(\theta_j)$ is strictly decreasing in λ, in $F^J(\theta_j)$, and in J when J is large enough. Because the equilibrium quality provision is equal to the marginal information rent, the comparative statics of quality $\hat{q}_j^*$ immediately extend to the information rent $\hat{U}_j^*$. Finally, the argument for why the platform's fee increases with the platform's information precision (Proposition 6.2) is given in the text. ■

Proof of Proposition 8. Consider a symmetric equilibrium where on-platform consumers know their type. If firm j participates in the mechanism but offers an off-equilibrium menu, only consumers who search for firm j in equilibrium will observe this deviation. Therefore, every firm that participates can do no better than to advertise the efficient quality levels and post the off-platform menus that solve (19). However, if firm j does not participate, it can match its competitors' information rents $\hat{U}_{-j}$ and attract all the consumers on the platform who value its products the most. (These consumers will search for firm j's off-platform offer regardless of the ad shown to them on the platform). Furthermore, if a nonparticipating firm j maximizes its profits with respect to $(\hat{q}, \hat{U})$ over the combined off- and on-platform market segments, it solves the problem in (13). This is a standard second-degree price discrimination problem, where consumer types are distributed according to $\lambda F^J + (1-\lambda) G^J$. Whenever positive, the optimal quality provision in such a deviation is given by

$$\hat{q}(\theta_j) = \theta_j - \frac{1 - (1-\lambda)\, G^J(\theta_j) - \lambda F^J(\theta_j)}{\lambda F^{J-1}(\theta_j) f(\theta_j) + (1-\lambda)\, G^{J-1}(\theta_j) g(\theta_j)}. \tag{22}$$

Finally, because the quality level $\hat{q}$ in (22) is pointwise larger than $\hat{q}_j^*$ in (21), the resulting information rent is correspondingly higher for each θ_j.

Thus, the deviating firm's optimal choice of menu yields an outside option $\hat{\Pi}_j$ larger than $\bar{\Pi}_j$ in (12). In addition, because the on-path profits are unchanged relative to the case of asymmetrically informed consumers, the platform's fee must decrease. ■

Proof of Proposition 9. Consider each firm's best-response problem in the subgame following the platform's offer. For all upward deviations (i.e., for strategies $\hat{U}_j$ such that $\theta^*_{-j}(\theta_j) \geq \theta_j$), the Hamiltonian associated with problem (17) can be written as

$$\begin{aligned} H &= (1-\lambda)\left(\theta_j \hat{q}(\theta_j) - \hat{q}(\theta_j)^2 - \hat{U}(\theta_j)\right) G^{J-1}(\theta_j) g(\theta_j) \\ &\quad + \lambda \left(\theta_j \hat{q}(\theta_j) - \hat{q}(\theta_j)^2/2 - \hat{U}(\theta_j)\right) F^{J-1}(\theta^*_{-j}(\theta_j)) f(\theta_j) + \gamma(\theta_j) \hat{q}(\theta_j). \end{aligned}$$

The necessary conditions for such an upward deviation not to be profitable are given by

$$\begin{aligned}(1-\lambda)\left((\theta_j-\hat{q}(\theta_j))\,G^{J-1}(\theta_j)g(\theta_j)+\gamma(\theta_j)\right. &= 0 \\ -(1-\lambda)\,G^{J-1}(\theta_j)g(\theta_j)-\lambda F^{J-1}(\theta_j)f(\theta_j)+\gamma'(\theta_j) & \\ +\lambda\frac{(J-1)\,F^{J-2}(\theta_j)f^2(\theta_j)}{\hat{q}(\theta_j)}\left(\theta_j\hat{q}(\theta_j)-\hat{q}(\theta_j)^2/2-\hat{U}(\theta_j)\right) &\leq 0.\end{aligned} \tag{23}$$

Now recall the costate variable $\hat{\gamma}(\theta_j)$ in (20) associated with the best response problem (19) in the baseline model. Because the last term on the left-hand side of (23) 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

$$\gamma\left(1\right)=\hat{\gamma}\left(1\right)=0.$$

We can then conclude that

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

which implies that the quality and utility levels $\hat{q}_j(\theta_j)$ and $\hat{U}_j(\theta_j)$ are weakly higher for all θ_j *with* organic links than without.

We now show that the firms' profits are lower and their outside options are higher with organic links than without. In a symmetric equilibrium, the firm's profits are given by

$$\begin{aligned}\Pi_j(\hat{q},\hat{U}) &= (1-\lambda)\int_0^1\left[\theta_j\hat{q}(\theta_j)-\hat{q}(\theta_j)^2/2-\hat{U}(\theta_j)\right]G^{J-1}(\theta_j)\mathrm{d}G(\theta_j) \\ &\quad+\lambda\int_0^1(\theta_j^2/2-\hat{U}(\theta_j))F^{J-1}(\theta_j)\mathrm{d}F(\theta_j).\end{aligned} \tag{24}$$

The equilibrium menu in the baseline model $(\hat{q}_j^*,\hat{U}_j^*)$ maximizes (24), while the equilibrium menu maximizes (17) and hence achieves a weakly lower profit level. Now consider the deviating firm's profit. For any choice of $(\hat{q},\hat{U})$ off path, the deviation profits are weakly larger with organic links than without. With organic links, the firm wins a fraction $F^{J-1}(\theta^*_{-j}(\theta_j))\in[0,1]$ of on-platform types θ_j. Without organic links, the deviating firm makes no sales on the platform. Consequently, we have $\tilde{\Pi}_j\geq\bar{\Pi}_j$, which implies *a fortiori* that the platform's fees are lower. ■

Proof of Proposition 10. We construct an equilibrium where each firm j sets its on- and off-platform menus to maximize profits, given that it 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. Firm j solves:

$$\max_{q,\hat{q},U,\hat{U}} \left[\begin{array}{c} \lambda \int \left(\theta_j q(\theta_j) - q(\theta_j)^2/2 - U(\theta_j)\right) F^{J-1}(\theta_j) \mathrm{d}F(\theta_j) \\ + (1-\lambda) \int \left(\theta_j \hat{q}_j(\theta_j) - \hat{q}(\theta_j)^2/2 - \hat{U}(\theta_j)\right) G^{J-1}(\theta_j) \mathrm{d}G(\theta_j) \end{array} \right] \tag{25}$$

$$\begin{aligned} \text{s.t.} \quad & U'(\theta_j) = q(\theta_j) \\ & \hat{U}'(\theta_j) = \hat{q}(\theta_j) \\ & U(\theta_j) \geq \hat{U}(\theta_j) \geq 0. \end{aligned}$$

We now show that the solution to (25) is given by

$$q_j^*(\theta_j) = \hat{q}_j^*(\theta_j) = \theta_j - \frac{1 - (1-\lambda) G^J(\theta_j) - \lambda F^J(\theta_j)}{J\lambda F^{J-1}(\theta_j) f(\theta_j) + J(1-\lambda) G^{J-1}(\theta_j) g(\theta_j)} \tag{26}$$

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

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

The pointwise necessary conditions for this problem are the following:

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

(The last condition is analogous the one in Jullien (2000), Theorem 2. In that paper, the shadow cost of the type-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).

If $q(\theta_j) = \hat{q}(\theta_j)$ as in (26), we obtain the following expressions for the costate variables:

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

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

$$\begin{aligned}\bar{\gamma}(\theta_j) &= \frac{J\lambda (1-\lambda) \left(1 - (1-\lambda) G^J(\theta_j) - \lambda F^J(\theta_j)\right)}{\left((1-\lambda) J G^{J-1}(\theta_j) g(\theta_j) + \lambda J F^{J-1}(\theta_j) f(\theta_j)\right)^2} \cdot \\ &\quad \cdot \left(\frac{\mathrm{d}F^{J-1}(\theta_j) f(\theta_j)}{\mathrm{d}\theta_j} G^J(\theta_j) - \frac{\mathrm{d}G^{J-1}(\theta_j) g(\theta_j)}{\mathrm{d}\theta_j} F^J(\theta_j)\right),\end{aligned}$$

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

References

Acemoglu, D., A. Makhdoumi, A. Malekian, and A. Ozdaglar (2022): "Too Much Data: Prices and Inefficiencies in Data Markets," *American Economic Journal: Microeconomics*, forthcoming.

Admati, A. R., and P. Pfleiderer (1990): "Direct and indirect sale of information," *Econometrica*, 58(4), 901–928.

Aggarwal, G., A. Badanidiyuru, and A. Mehta (2019): "Autobidding with Constraints," in *International Conference on Web and Internet Economics*, pp. 17–30. Springer.

Anderson, S. P., and Ö. Bedre-Defolie (2021): "Hybrid platform model," Discussion Paper 16243, CEPR.

Anderson, S. P., and R. Renault (1999): "Pricing, product diversity, and search costs: A Bertrand-Chamberlin-Diamond model," *RAND Journal of Economics*, 30(4), 719–735.

Armstrong, M., and J. Zhou (2011): "Paying for prominence," *The Economic Journal*, 121(556), F368–F395.

Balseiro, S. R., Y. Deng, J. Mao, V. S. Mirrokni, and S. Zuo (2021): "The Landscape of Auto-Bidding Auctions: Value Versus Utility Maximization," in *Proceedings of the 22nd ACM Conference on Economics and Computation*, pp. 132–133.

Bar-Isaac, H., and S. Shelegia (2020): "Search, Showrooming, and Retailer Variety," Discussion paper, CEPR Discussion Paper No. DP15448.

Baye, M., and J. Morgan (2001): "Information gatekeepers on the internet and the competitiveness of homogeneous product markets," *American Economic Review*, 91(3), 454–474.

Bergemann, D., and A. Bonatti (2019): "Markets for Information: An Introduction," *Annual Review of Economics*, 11, 85–107.

Bergemann, D., A. Bonatti, and T. Gan (2022): "The Economics of Social Data," *RAND Journal of Economics*, 53(2), 263–296.

Bergemann, D., B. Brooks, and S. Morris (2015): "The Limits of Price Discrimination," *American Economic Review*, 105, 921–957.

——— (2021): "Search, Information, and Prices," *Journal of Political Economy*, 129(8), 2275–2319.

Blackwell, D. (1951): "Comparison of Experiments," in *Proceedings of the Second Berkeley Symposium in Mathematical Statistics and Probability*, pp. 93–102. University of California Press, Berkeley.

Bonatti, A. (2011): "Brand-Specific Tastes for Quality," *International Journal of Industrial Organization*, 29, 562–575.

Calzolari, G., and V. Denicolo (2015): "Exclusive contracts and market dominance," *American Economic Review*, 105(11), 3321–51.

Choi, J., D. Jeon, and B. Kim (2019): "Privacy and Personal Data Collection with Information Externalities," *Journal of Public Economics*, 173, 113–124.

Cremèr, J., Y.-A. de Montjoye, and H. Schweitzer (2019): "Competition policy for the digital era," Discussion paper, European Commission.

De Corniere, A., and G. Taylor (2019): "A model of biased intermediation," *The RAND Journal of Economics*, 50(4), 854–882.

Deng, Y., J. Mao, V. Mirrokni, and S. Zuo (2021): "Towards Efficient Auctions in an Auto-Bidding World," in *Proceedings of the Web Conference 2021*, pp. 3965–3973.

Diamond, P. A. (1971): "A model of price adjustment," *Journal of Economic Theory*, 3(2), 156–168.

Donnelly, R., A. Kanodia, and I. Morozov (2022): "Welfare Effects of Personalized Rankings," *Available at SSRN 3649342*.

Elliott, M., A. Galeotti, and A. Koh (2020): "Market Segmentation through Information," Discussion paper, Cambdrige University.

Fuchs, W., and A. Skrzypacz (2015): "Government interventions in a dynamic market with adverse selection," *Journal of Economic Theory*, 158, 371–406.

Gomes, R., and A. Pavan (2016): "Many-to-Many Matching and Price Discrimination," *Theoretical Economics*, 11(3), 1005–1052.

Gur, Y., G. Macnamara, I. Morgenstern, and D. Saban (2022): "Information Disclosure and Promotion Policy Design for Platforms," Discussion paper, Stanford University.

Hagiu, A., and B. Jullien (2011): "Why do intermediaries divert search?," *RAND Journal of Economics*, 42(2), 337–362.

Hagiu, A., T.-H. Teh, and J. Wright (2022): "Should platforms be allowed to sell on their own marketplaces?," *RAND Journal of Economics*, forthcoming.

Ichihashi, S. (2021): "The Economics of Data Externalities," *Journal of Economic Theory*, 196, 105316.

Idem, B. (2021): "Coexistence of Centralized and Decentralized Markets," Discussion paper, Pennsylvania State University.

Inderst, R., and M. Ottaviani (2012a): "Competition through commissions and kickbacks," *American Economic Review*, 102(2), 780–809.

——— (2012b): "How (not) to pay for advice: A framework for consumer financial protection," *Journal of Financial Economics*, 105(2), 393–411.

Jullien, B. (2000): "Participation Constraints in Adverse Selection Models," *Journal of Economic Theory*, 93(1), 1–47.

Kang, Z. Y., and E. Muir (2021): "Contracting and vertical control by a dominant platform," Discussion paper, Stanford University.

Ke, T. T., S. Ling, and M. Y. Lu (2022): "Information Design of Online Platforms," Discussion paper, Chinese University of Hong Kong.

Kirpalani, R., and T. Philippon (2021): "Data Sharing and Market Power with Two-Sided Platforms," Discussion Paper 28023, NBER.

Lam, H. T. (2021): "Platform Search Design and Market Power," Discussion paper, Northwestern University.

Lee, C. (2021): "Optimal Recommender System Design," Discussion paper, University of Pennsylvania.

Lee, K.-H., and L. Musolff (2021): "Entry into Two-Sided Markets Shaped By Platform-Guided Search," Discussion paper, Princeton University.

Maskin, E., and J. Riley (1984): "Monopoly with Incomplete Information," *Rand Journal of Economics*, 15(2), 171–196.

Miklós-Thal, J., and G. Shaffer (2021): "Input price discrimination by resale market," *RAND Journal of Economics*, 52, 727–757.

Mussa, M., and S. Rosen (1978): "Monopoly and Product Quality," *Journal of Economic Theory*, 18(2), 301–317.

Padilla, J., J. Perkins, and S. Piccolo (2020): "Self-preferencing in markets with vertically-integrated gatekeeper platforms," *Available at SSRN 3701250*.

Philippon, T., and V. Skreta (2012): "Optimal interventions in markets with adverse selection," *American Economic Review*, 102(1), 1–28.

Rayo, L., and I. Segal (2010): "Optimal Information Disclosure," *Journal of Political Economy*, 118(5), 949–987.

Shi, P. (2022): "Optimal match recommendations in two-sided marketplaces with endogenous prices," Discussion paper, University of Souther California.

Teh, T.-H., and J. Wright (2022): "Intermediation and steering: Competition in prices and commissions," *American Economic Journal: Microeconomics*, 14(2), 281–321.

Tirole, J. (2012): "Overcoming adverse selection: How public intervention can restore market functioning," *American Economic Review*, 102(1), 29–59.

Wang, C., and J. Wright (2020): "Search platforms: Showrooming and price parity clauses," *RAND Journal of Economics*, 51(1), 32–58.

Yang, K. H. (2022): "Selling Consumer Data for Profit: Optimal Market-Segmentation Design and Its Consequences," *American Economic Review*, 112(4), 1364–93.