DATA, COMPETITION, AND DIGITAL PLATFORMS

By

Dirk Bergemann and Alessandro Bonatti

April 2023

COWLES FOUNDATION DISCUSSION PAPER NO. 2343R

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Data, Competition, and Digital Platforms*

Dirk Bergemann† Alessandro Bonatti‡

April 12, 2023

Abstract

We analyze digital markets where a monopolist platform uses data to match multiproduct sellers with heterogeneous consumers who can purchase both on and off the platform. The platform sells targeted ads to sellers that recommend their products to consumers and reveals information to consumers about their values. The revenue-optimal mechanism is a managed advertising campaign that matches products and preferences efficiently. In equilibrium, sellers offer higher qualities at lower unit prices on than off the platform. Privacy-respecting data-governance rules such as organic search results or federated learning can lead to welfare gains for consumers.

KEYWORDS: Data, Privacy, Data Governance, Digital Advertising, Competition, Digital Platforms, Digital Intermediaries, Personal Data, Matching, Price Discrimination, Automated Bidding, Algorithmic Bidding, Managed Advertising Campaigns, Showrooming.
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 Bank of Canada, Berkeley, Indiana, Northwestern, NYU, Paris-Panthéon, Toulouse, as well as participants in the CEPR Virtual IO Seminar, the Luohan Academy Webinar, the FTC Microeconomics Conference, the Asia-Pacific Industrial Organization Conference, the 2022 Market Innovation Workshop, and the 2022 NBER Economics of Privacy Conference. We are grateful to our discussants Andrei Hagiu, Alessandro Pavan, and Vasiliki Skreta for helpful comments and suggestions, and to Lucas Barros, Andrew Koh, 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

1.1 Motivation

The role of data in shaping competition in online markets has become a critical issue in both economics and policy. 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 sellers. They monetize their data by selling sponsored content and targeted advertising. The quality of a platform's data allows better pairing of viewers and sellers. Digital platforms not only serve as gatekeepers of information online but also act as competition managers.[1]

Regulators fear that platforms may leverage their gatekeeper position to increase merchants' market power, thereby raising their willingness to pay for advertising.[2] 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 sellers? How do they depend on the mechanisms for collecting and sharing consumer data?

In this paper, we develop a model of an intermediated online marketplace and trace how a data-rich platform creates and distributes surplus among market participants. Our goal is to provide a tractable and flexible framework to study digital markets where different privacy regimes can be compared. Our model captures three ubiquitous features of digital platforms. First, nearly every platform leverages the informational advantage to personalize the *sponsored content* at the individual consumer level through managed advertising campaigns.[3] Second, while price discrimination is rare, targeted advertising and personalized recommendations amount to *product steering*.[4] Third, most sellers have parallel sales channels, i.e., consumers can buy their products both on and off digital platforms.

[1]The European Union's Digital Markets Act establishes a set of narrowly defined objective criteria for qualifying a large online platform as a so-called "gatekeeper". These criteria will be met if a company: (a) has a strong economic position in the EU market, (b) has a strong intermediation position, (c) has (or is about to have) an entrenched and durable position in the market.

[2]The report by Crémer et al. (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."

[3]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.

[4]Donnelly et al. (2022) document the effect of personalized recommendations on a retail platform, and Raval (2020) illustrates a recent shift in eBay policy.

To capture these features, we consider a market with differentiated quality-pricing sellers. Consumers have heterogeneous preferences for the sellers' product lines but are imperfectly informed about their own values. The platform's data identify the most valuable consumer-seller pair and the most valuable product within that seller's product line. However, these data also create the potential for 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 sellers' advertising campaigns. Managed campaigns are emerging as the predominant mode of selling advertisements in real-world digital markets: sellers set a fixed advertising budget, specify high-level objectives for their campaigns, and leave the task of bidding to "auto-bidders" offered by the platform.[5] In our model, the digital platform receives an advertising budget from each seller. The advertising budget then generates sponsored listings for each specific product of that seller. In particular, the platform advertises to each consumer the product generating the highest social surplus for that consumer. This mechanism is akin to the managed advertising campaigns prevalent in sponsored search and other forms of digital advertising. In turn, these are frequently implemented through an automated bidding mechanism by platforms such as Google, Facebook, or Tiktok.

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

1.2 Results

The platform's informational advantage over consumers and the search frictions on the platform, no matter how small, give the digital platform significant bargaining power over sellers. Our first main result shows that the platform can completely control consumers' shopping

[5] Most of online advertising is traded programatically, often through a mechanism that maximizes the total number of clicks or conversions subject to a return on ad spend (ROAS) constraint (https://support.google.com/google-ads/answer/6268637). For recent work on programmatic or algorithmic bidding, see Aggarwal et al. (2019), Balseiro et al. (2021), and Deng et al. (2021).

behavior and steer them away from sellers who do not submit an advertising budget (Proposition 2). Consumers understand the managed-campaign mechanism and expect that in equilibrium, advertised products generate the highest value. In the presence of search costs, consumers only consider buying from the advertised brand, whether on-platform or off-platform. As a result, the platform restricts competition among sellers, as each seller only faces consumers who are most interested in their products and competes with their own off-platform offers only (Proposition 3).

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

Next, we examine the platform's dual gatekeeper role by considering two sources of the platform's bargaining power: its information advantage over consumers and sellers, and consumers' search costs off the platform. We first show that it would be against the platform's interest to provide consumers with information about non-sponsored products. We assume that consumers on the platform observe their values perfectly, so the platform cannot steer their behavior away from their favorite seller even if that seller rejects the platform's offer. Complete information for on-platform consumers does not change the equilibrium prices or products, but it reduces the platform's fees (Proposition 5). We then assume that the platform offers organic links that advertise all off-platform prices and products to all on-platform consumers. We show that the provision of price information introduces menu competition among sellers, reducing all equilibrium prices, both on and off the platform, as well as the platform's fees (Proposition 6).

We also investigate how data governance, the rules governing how consumer data can be collected and deployed, influences the creation and distribution of social surplus. In particular, we discuss the implications of cohort-based advertising, as the outcome of federated learning, compared to personalized advertising. We show that cohort-based advertising, which protects privacy and allows consumers to retain an information advantage over sellers on the platform, improves consumer surplus (Proposition 7).

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

Finally, we examine the effects of the platform size. We show that the distortions in off-platform quality become more severe as the fraction of on-platform consumers grows (Proposition 10). Holding prices fixed, consumers benefit from shopping on a better-informed platform and obtaining higher-quality matches. However, as more consumers join the platform, prices rise both on and off the platform.

1.3 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), Condorelli and Szentes (2022), De Corniere and Taylor (2019), Gomes and Pavan (2016), Gur et al. (2022), Hagiu and Jullien (2011), Inderst and Ottaviani (2012a), Inderst and Ottaviani (2012b), Ke et al. (2022), Rayo and Segal (2010), and Shi (2022)—analyze the steering role of platforms that strategically modify search results, e.g., to match consumers with the sellers that pay the largest commissions.[6]

The provision of information by a digital platform is central to the model of de Cornière and de Nijs (2016), who examine a platform's incentives to provide match-value information to differentiated sellers in a second-price auction model. More recently, Teh and Wright (2022) study the signaling role of ranking the search results, and Zhu et al. (2022) study the impact of a platform's privacy policies on the downstream competition within and across product categories. Relative to these papers, we allow the platform to provide information directly to the consumer, e.g., through product reviews. Moreover, the multiproduct sellers in our model 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, sellers pay fees (or bids) that do not vary with the prices of their products.[7]

A recent body of work including Choi et al. (2019), Acemoglu et al. (2022), Ichihashi (2021), Kirpalani and Philippon (2021), and Bergemann et al. (2022) documented the *data*

[6] The trade-off between value creation through personalization and consumer surplus extraction is also central in Hidir and Vellodi (2021) and Ichihashi (2020).

[7] In Gomes and Pavan (2022), a monopolist platform elicits sellers' willingness to pay for qualified consumer eyeballs through a nonlinear tariff for advertising space, which is analogous to our managed campaign. In contrast, we explicitly model consumer search and determine the equilibrium prices for the consumers.

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 the 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 the private label products.[8] Unlike in these papers, the sellers in our model are concerned about showrooming because the opportunity to sell on the platform benefits them through the added value of making personalized offers.

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 the ability to monopolize the market since the sellers have access to an outside option. Our setting shares some of the same features, but in an oligopoly environment where sellers compete for heterogeneous consumers. Furthermore, the sellers choose their product menus understanding that customers arrive through two different channels and that they have distinct information in each channel.

At a broad level, this paper relates to information structures in advertising auctions, e.g., Bergemann et al. (2021), and to nonlinear pricing, market segmentation, and competition, e.g., Bergemann et al. (2015), Bonatti (2011), Elliott et al. (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 et al. (2022), Kang and Muir (2021), Lam (2021), Lee (2021), Lee and Musolff (2021), and Padilla et al. (2020).

Finally, in parallel work, Bergemann et al. (2023) compare auto-bidding through managed advertising campaigns and data-augmented second-price auctions for online advertising. They show that the managed campaign mechanism increases the revenue of the digital platform and, with sufficient competition among advertisers, it also increases consumer surplus.

[8] Gutierrez (2022) estimates an equilibrium model of Amazon's hybrid retail platform.

2 Model

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

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

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

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

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

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

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

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

with equality for $v = 0$. If F is a mean-preserving spread of G, we write $F \succ G$. (Conversely, G is referred to as a mean-preserving contraction of the given distribution F.)

On Platform A fraction $\lambda \in [0, 1]$ of all consumers uses the platform to find a product.[9] The platform has access to extensive data and knows each consumer's value profile θ, while the sellers only know the corresponding prior distribution F. The platform offers a single

[9] We can endogenize the fraction of on-platform consumers λ by introducing heterogeneity in the cost and benefits of using the platform (e.g., in the loss of privacy due to leaving "footprints" online).

sponsored link for a product q_j sold by a seller j. The platform uses a *managed (advertising) campaign* mechanism to select which seller and which product to sponsor to each consumer. This mechanism captures all the salient features of automated bidding on real-world digital advertising platforms.

In a managed campaign mechanism, the platform requests an advertising budget $t \in \mathbb{R}_+$ from each seller and announces a *selection rule*. Any seller who pays the required budget submits quality and price functions $q_j(\theta)$ and $p_j(\theta)$ representing the product and price that seller j intends to advertise to each consumer θ. For each consumer value profile θ, the platform then advertises a single product $(q_j(\theta), p_j(\theta))$ according to the selection rule. The platform then chooses a sponsored seller according to the selection (or steering) rule:

$$\sigma : \Theta \times \mathbb{R}_+^\Theta \times \mathbb{R}_+^\Theta \to [J], \tag{1}$$

and advertises the selected seller j's product and price $q_j(\theta)$ and $p_j(\theta)$.

Until Section 6, we shall maintain the assumption that, upon advertising product q_j to consumer θ through the sponsored link, the platform directly provides additional information to the consumer that fully reveals their value θ_j for the advertised product.

Off Platform The remaining $1-\lambda$ consumers buy off the platform, e.g., from the merchants' own websites or physical stores. Off the platform, the consumers have their expectation m and the sellers know the corresponding distribution G. On the platform, there is extensive data and the platform knows each consumer's value profile θ.

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

$$\{(\widehat{p}_j(m_j), \widehat{q}_j(m_j))\}_{m_j \in [\theta_L, \theta_H]} \tag{2}$$

as in Mussa and Rosen (1978) and Maskin and Riley (1984). Throughout the paper, we thus use the circumflex to distinguish off-platform variables from on-platform variables. Importantly, the goods being sold are not experience or inspection goods: to learn the vector θ, consumers and sellers must gain access to the platform's data.

After receiving seller j's offer on the platform and learning their value θ_j for the advertised product q_j, each consumer can search off the platform and use the newly gained information to select a product from any seller. For example, the consumer can buy from the off-platform schedule (2) posted by the advertised seller j. Figure 1 summarizes the interaction between the agents and their actions in our model.

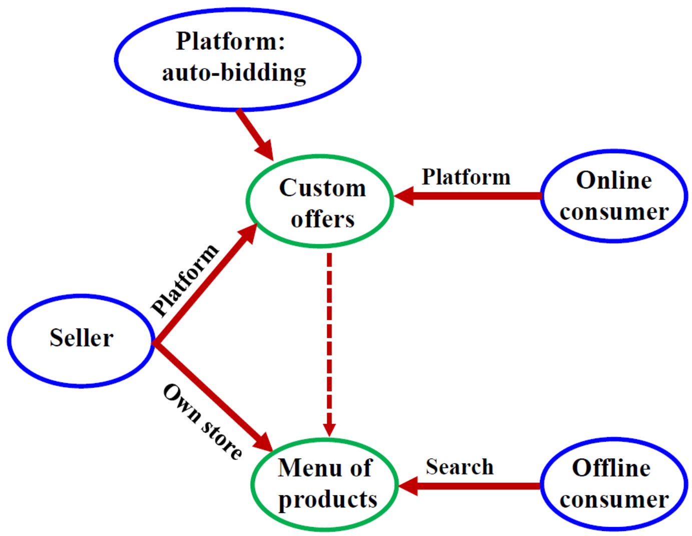

Figure 1: On- and Off-Platform Matching

Timing and Equilibrium We consider simultaneous choices of (on- and off-platform) prices to capture the great flexibility that algorithmic pricing offers the sellers both on and off the platform. The timing of our game is as follows:

1. The platform announces a selection rule σ and requests an advertising budget t_j from each seller j.

2. Sellers simultaneously set off-platform products $\widehat{q}_j(m)$ and prices $\widehat{p}_j(m)$, choose whether to submit the budget, 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 according to the announced selection rule.

4. The on-platform consumers learn their value θ_j for the advertised seller; they can purchase the advertised product on the platform or search off the platform.

Definition 1 (Symmetric Perfect Bayesian Equilibrium)
We consider symmetric Perfect Bayesian Equilibria. The consumers have symmetric beliefs over the sellers' off-platform menus both on and off the equilibrium path.

Thus, consumers expect sellers to play symmetric strategies on the equilibrium path, and they continue to hold symmetric (though not necessarily passive) beliefs over every seller's prices and qualities even when they observe a deviation (either on or off the platform).[10]

[10]Symmetric beliefs off the path of play facilitate our exposition of the intuition around Propositions 2 and 5. However, the results also hold under the passive beliefs refinement.

Discussion We briefly comment on a few central aspects of the model. The managed campaign mechanism has several important properties. First, sellers do not acquire the platform's data, but they condition products and prices on the platform's information about the consumer. In other words, the consumer's value 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). Second, fixed payments for advertising slots are similar to automated bidding in ad auctions. Sellers submit a budget and upload the ads for the products they wish to show to select consumers. Third, because each seller can tailor the product offer to each consumer value, the platform creates an opportunity for surplus extraction through *product steering*, without personalizing prices.

Several assumptions in the above mechanism can be easily relaxed. In particular, we allow the sellers' schedules $q_j(\theta)$ and $p_j(\theta)$ to condition on the entire value θ, and not just θ_j. However, this additional flexibility will be redundant in equilibrium. Our model of a single sponsored link is also simple in that the platform sells both information and recognition to a single seller and a specific product. In Section 6, we extend our model to allow for all brands and products to be present on the platform through organic search results that advertise their off-platform offers. Finally, the direct revelation of information to consumers captures the rich contextual detail that some retail platforms provide to their users. Here, we assume that the platform fully utilizes the informational advantage to generate surplus through efficient product-consumer matching. In Section 6.4, we relax the assumption of perfect revelation of θ_j, and we study information design by the platform.

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 seller (rather than many sellers) and binary values (rather than a multidimensional continuum of values). In addition, the distribution of consumer values and expected values is identical on and off the platform; thus, $F = G$. The platform retains an informational advantage over the sellers because it learns the value of the consumer that remains private information off the platform.

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

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

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

$$\{(q(\theta), U(\theta))\}_{\theta\in\{\theta_L,\theta_H\}}, \tag{3}$$

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

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

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

$$\{(\widehat{q}(\theta), \widehat{U}(\theta))\}_{\theta\in\{\theta_L,\theta_H\}}. \tag{4}$$

Throughout the paper, we thus use the circumflex to distinguish off-platform variables from on-platform variables. The seller's profit is:

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

The seller maximizes profit 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, if the seller wants a consumer to accept their targeted offer, the seller must induce the consumer not to buy off the platform. The seller then faces the following new "showrooming" constraints:

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

In other words, each consumer 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 seller should offer the socially efficient quality levels on the platform and that the showrooming constraint should bind,

$$q(\theta) = \theta \text{ and } U(\theta) = \widehat{U}(\theta), \ \theta \in \{\theta_L, \theta_H\}. \tag{5}$$

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

$$\max_{q,U} \Big[\lambda(f(\theta_L)\theta_L^2/2 + f(\theta_H)(\theta_H^2/2 - \widehat{U}(\theta_H))) \qquad (6)$$
$$+ (1-\lambda)(f(\theta_L)\left(\theta_L\widehat{q}(\theta_L) - \widehat{q}(\theta_L)^2/2\right) + f(\theta_H)(\theta_H^2/2 - \widehat{U}(\theta_H)))\Big]$$

subject to the constraint

$$\widehat{U}(\theta_H) = (\theta_H - \theta_L)\widehat{q}(\theta_L). \qquad (7)$$

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

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

$$\begin{aligned} \widehat{q}(\theta_L) &= \max\left\{0, \theta_L - \frac{f(\theta_H)}{f(\theta_L)}(\theta_H - \theta_L)\left(1 + \frac{\lambda}{1-\lambda}\right)\right\}, \\ \widehat{q}(\theta_H) &= \theta_H. \end{aligned} \qquad (8)$$

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

$$\begin{aligned} \widehat{q}(\theta_L) &= \max\left\{0, \theta_L - \frac{f(\theta_H)}{f(\theta_L)}(\theta_H - \theta_L)\right\}, \\ \widehat{q}(\theta_H) &= \theta_H. \end{aligned} \qquad (9)$$

Proposition 1 indicates an additional opportunity cost of serving the low value off the platform. Indeed, it is not difficult to find parameters (e.g., λ large enough) for which the quality level on-platform as in (9) is strictly positive but the quality off-platform as in (8) is zero. Thus, without a platform, the seller would offer a low-quality product to the low value. However, for a sufficiently large platform, the low value is only offered a product on the platform, where the seller can make a different personalized offer to the high value. This is not the case off the platform, where the seller 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 $\widehat{q}(\theta_L) = 0$, no consumer value receives any rent on or off the platform.

Finally, to determine the optimal advertising budget t^*, we need to consider what the on-platform consumers would do if the seller did not advertise. If these consumers can buy off the platform, the optimal advertising budget extracts the seller's extra profit relative to offering the menu in (9) to all consumers. If these consumers were not to buy at all, the seller's outside option is scaled by a factor $1 - \lambda$, and the optimal advertising budget is correspondingly higher.

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

4 Managed Campaign and Showrooming

We now analyze the environment with many sellers and a multidimensional continuum of values as introduced in Section 2. Our objective is to establish the equilibrium patterns of consumer search induced by the informational advantage of the platform. (This advantage is captured by the distinction between values θ_j with distribution F on the platform and expected values m_j with distribution G off the platform.)

We uncover the following tradeoff: off the platform, each seller j faces those consumers who value their product the most based on their expected value m. However, trade takes place under asymmetric information, i.e., the seller must elicit the consumers' willingness to pay. In contrast, the platform enables consumers and sellers to interact under symmetric information. Thus, the sellers are willing to pay for the right to make a personalized offer to each consumer they are matched to under the platform's managed campaign mechanism. However, if the winning seller wants a consumer to accept their personalized offer, this seller must induce the consumer not to buy from the off-platform store, i.e., not to use the platform for "showrooming." Thus, the seller's ability to product steer and price discriminate on the platform is limited by the presence of the off-platform channel.

4.1 Managed Campaign

As we introduced in Section 2, the platform designs a managed advertising campaign to select a sponsored product. This mechanism is defined formally as follows.

Definition 2 (Managed Advertising Campaign)
In a managed campaign mechanism:

1. *the platform requests an advertising budget $t \in \mathbb{R}_+$ from each seller as a take-it-or-leave-it offer;*

2. *each participating seller j submits schedules $q_j : \Theta \to \mathbb{R}_+$ and $p_j : \Theta \to \mathbb{R}_+$, which represent the quality and price seller j intends to advertise to each consumer θ;*

3. *for each consumer value profile $\theta \in [\theta_L, \theta_H]^J$, the platform chooses a sponsored seller according to the selection mapping*

$$\sigma : \Theta \times \mathbb{R}_+^\Theta \times \mathbb{R}_+^\Theta \to [J],$$

and advertises the selected seller j's product and price $q_j(\theta)$ and $p_j(\theta)$.

Throughout the paper, we focus on a specific selection rule, namely the one that matches consumers and products efficiently. Indeed, we establish the revenue optimality of this mechanism in Proposition 4.

Definition 3 (Efficient Steering)
For each value profile θ, the platform chooses the seller that maximizes the social value of the match

$$\sigma^*(\theta, q(\theta), p(\theta)) = \arg\max_j [\theta_j q_j(\theta) - c(q_j(\theta))],$$

among all the sellers that participate in the mechanism.

We now derive the equilibrium choice patterns when the platform steers consumers to products efficiently.

4.2 Choice Patterns

We begin with the off-platform consumers. These consumers (which have mass $1-\lambda$) face positive search costs beyond the first seller. As a result, in any symmetric equilibrium, a consumer with expected value m visits only the seller who offers the highest expected value:

$$j^{(1)} = \arg\max_j m_j.$$

This result does not depend on the magnitude of the search costs as established famously by Diamond (1971). Moreover, if the platform has a strict informational advantage ($F \succ G$), the on-platform consumers (which have mass λ) infer that the advertised seller j^* maximizes their willingness to pay, i.e., $\theta_{j^*} = \max_j \theta_j$. Because these consumers expect symmetric menus off the platform and the information rent function associated with those menus is strictly increasing, these consumers consider products offered by the advertised seller j^* only.

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

The platform augments the expectation of each consumer with additional data that lead to a revision from the expected value m to the (true) value θ. Figure 2 below illustrates the choice behavior by a consumer whose expected value m rank the sellers differently relative to (true) value θ. The consumer has two possible choices, thus $J = 2$ and the expected value m suggests that seller 1 offers a higher value, thus $m_1 > m_2$. Now suppose that on the platform the consumer is shown an advertisement by seller, $j = 2$. Indeed, the platform reveals to the consumer the value θ_2, and thus they will infer that $\theta_2 > \theta_1$. Therefore, by the above Proposition, the consumer either accepts seller 2's offer or shops off the platform from seller 2, but now with full knowledge of their value θ_2.

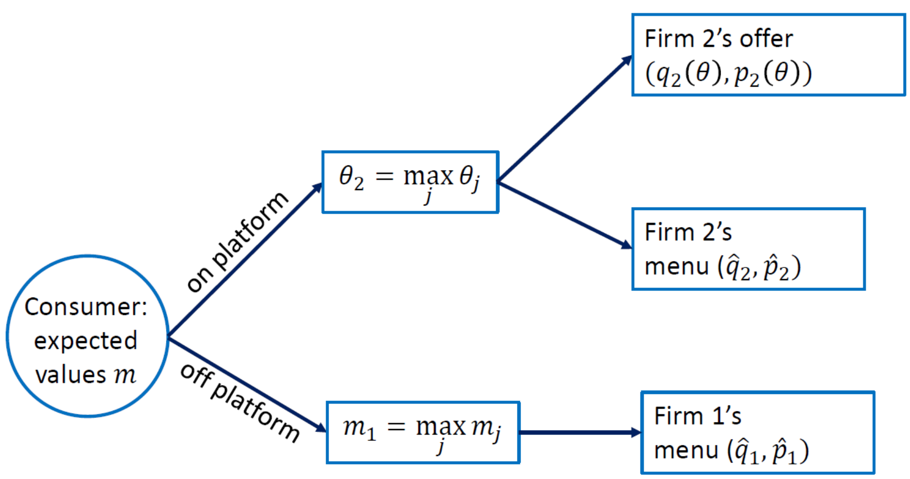

FIGURE 2: Possible Consumer Consideration Sets

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

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

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

4.3 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 seller, the seller can extract a substantial share of the created social surplus. To wit, the extraction of the surplus does not occur through personalized price discrimination but through product steering (thus a form of second-degree price discrimination). The only limit on surplus extraction by the advertising seller is given by the "showrooming constraint," which is a necessary condition for seller j to make a sale on the platform:

$$\theta_j \cdot q_j(\theta) - p_j(\theta) \geq \max_{m_j} [\theta_j \cdot \widehat{q}_j(m_j) - \widehat{p}_j(m_j)] \text{ for all } \theta_j. \tag{10}$$

Because seller j offers an incentive-compatible menu off the platform, each on-platform consumer would also report their value truthfully if shopping off-platform. By Proposition 2, we know the consumer chooses between two products by the same seller, and the showrooming constraint (10) reduces to

$$U_j(\theta) := \theta_j q_j(\theta) - p_j(\theta)) \geq \theta_j \widehat{q}_j(\theta_j) - \widehat{p}_j(\theta_j)) =: \widehat{U}_j(\theta). \tag{11}$$

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

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

[11] Mekonnen et al. (2023) consider the impact of exogenous information at each stage of a consumer's sequential search process, while Chen (2022) studies the evolution of a consumers' consideration set under the equilibrium advertising levels.

The socially efficient quality provision maximizes both the profit from the ad and the probability of being chosen by the platform. Similarly, it is optimal for each seller to offer the consumer a discount that satisfies the showrooming constraint (11) with equality. Thus, despite the flexibility awarded by the platform, the quality and utility offered on-platform by firm j are a function of θ_j only.

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

$$\pi_j(\theta, \widehat{U}_j) = \begin{cases} \theta_j^2/2 - \widehat{U}_j(\theta_j), & \text{if } \theta_j > \max_{k \neq j} \theta_k; \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

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 steering mechanism is the revenue-maximizing managed campaign for the platform.

By Proposition 2, in any symmetric equilibrium of our model, no consumer (off-platform or on-platform) searches past the first seller on the equilibrium path. Combining the off-platform profit with the on-platform profit (12), each seller's maximization problem can be written as follows.

$$\begin{aligned} \Pi_j^* = \max_{\widehat{q}_j, \widehat{U}_j} (1-\lambda) & \int_{\theta_L}^{\theta_H} [m_j \widehat{q}_j(m_j) - \widehat{q}_j(m_j)^2/2 - \widehat{U}_j(m_j)] G^{J-1}(m_j) g(m_j) \mathrm{d}m_j \qquad (13) \\ & + \lambda \int_{\theta_L}^{\theta_H} [\theta_j^2/2 - \widehat{U}_j(\theta_j)] F^{J-1}(\theta_j) f(\theta_j) \mathrm{d}\theta_j, \\ \text{s.t.} \qquad & \widehat{U}_j(m_j) \geq 0, \qquad \text{(IR)} \\ & \widehat{U}_j'(m_j) = \widehat{q}_j(m_j). \qquad \text{(IC)} \end{aligned}$$

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

Because we have assumed $m_j \in [\theta_L, \theta_H]$, we state all our results with θ_j as the argument, letting the distributions F and G indicate whether we refer to on- or off-platform variables. Maximizing (13) over rent and quality functions $\widehat{U}_j$ and $\widehat{q}_j$, we obtain the following characterization of the optimal menus.[12]

Proposition 3 (Equilibrium Menus with Efficient Steering)

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

$$q_j^*(\theta_j) = \theta_j, \tag{14}$$

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

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

$$U_j^*(\theta_j) = \widehat{U}_j^*(\theta_j) = \int_{\theta_L}^{\theta_j} \widehat{q}_j^*(x)\,dx. \tag{16}$$

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

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

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

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

[12] As we have ignored the monotonicity constraint in problem (13), Proposition 3 applies to the case where the distributions F and G are sufficiently regular. If they are not, i.e., if the function $\hat{q}_j^*(\theta_j)$ in (15) is not monotone, then the equilibrium quality schedule is given by the "ironed" version of (15).

Because of the shadow cost of showrooming, the off-platform quality schedule $\widehat{q}$ is further distorted downward. In particular, we can rewrite the equilibrium off-platform qualities (15) in Proposition 3 as

$$\widehat{q}_j^*(\theta_j) = \underbrace{\theta_j - \frac{1 - G^J(\theta_j)}{JG^{J-1}(\theta_j)g(\theta_j)}}_{\text{Mussa and Rosen (1978) quality}} - \frac{\lambda}{1-\lambda}\frac{1 - F^J(\theta_j)}{JG^{J-1}(\theta_j)g(\theta_j)}, \tag{17}$$

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

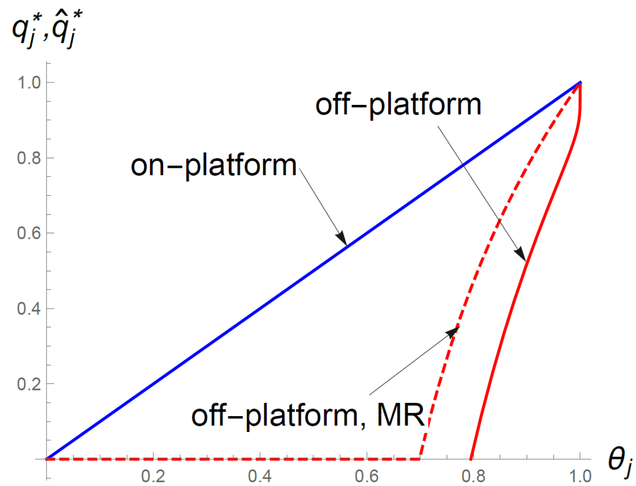

FIGURE 3: Match Values and Qualities On-Platform vs Off-Platform, $\lambda = 1/2, J = 5, G(m_j) = m_j, F(\theta_j) = \text{Beta}(\theta_j, 1/4, 1/4)$

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

The formulation of the optimal off-platform menu (17) allows us to establish several intuitive properties of the equilibrium. First, each value θ receives a higher quality level, namely the socially efficient allocation and pays a higher price on the platform than off the platform. However, while each value receives a better product at a higher price, each quality level q is sold at a lower price on the platform. Thus, let us define the equilibrium price for a given quality q on- and off-platform as $p(q)$ and $\widehat{p}(q)$, respectively. We then find that $p(q) \leq \widehat{p}(q)$ for all q. In other words, each seller is forced to introduce "on-platform only" discounts due to the threat of showrooming.

Figure 4 displays the nonlinear pricing schedules, namely the price $p_j(q_j)$ for every offered quality q_j under the same parameters as in Figure 3. Note that for a set of low values, namely those below 8/10, the nonlinear tariff is equal to the gross surplus generated by the efficient quality (i.e., $p_j(q_j) = q_j^2$). By contrast, values above 8/10 receive a positive rent off-platform, and hence on the platform the price is below the gross surplus.

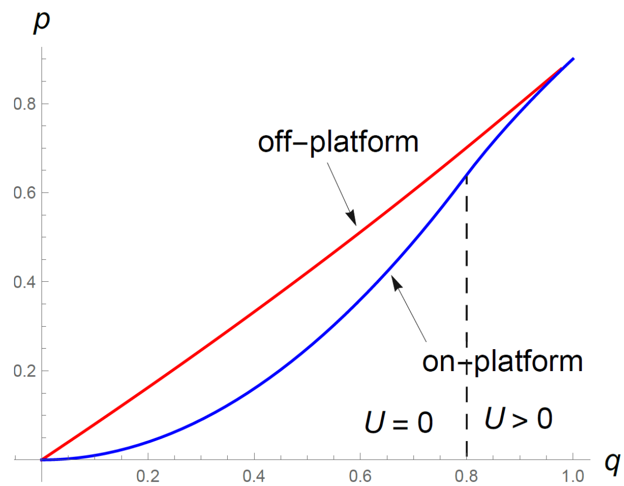

FIGURE 4: Nonlinear tariffs: every variety q is sold at a lower price on-platform.

Our results illustrate how advertising platforms that run managed campaigns face very different incentives than retail platforms that charge proportional transaction fees. In the case of Amazon, for example, merchants may want consumers to showroom to avoid the platform's variable fees. In the case of our advertising platform, ex-ante, fixed-price contracts with the sellers eliminate the need for 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 value θ, we have

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

However, we also know that $F \succ G$. This implies

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

because the highest order statistics satisfy $\mathbb{E}_{F^J}[\theta_j] > \mathbb{E}_{G^J}[\theta_j]$ and incentive compatibility requires the function $\widehat{U}_j^*$ to be increasing and convex. Thus, at the equilibrium prices, every consumer would rather be on the platform (ex ante) than off the platform. A stronger result is that, holding prices fixed, the consumer would like the platform to have as precise information as possible about their value, which enables better matching of products to preferences. However, the consumer does not necessarily benefit from the presence of the platform *in equilibrium*. Indeed, Proposition 10 considers the effects of a larger platform (λ) and finds that *all* consumers are worse off as the share λ of consumers who shop on the platform increases.

Platform Revenue To examine the implications for the sellers' profit and the platform's revenue, we characterize the advertising budgets the platform can demand in equilibrium under a managed campaign mechanism.

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

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

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

$$\overline{\Pi}_j(\sigma^*) = \max_{\widehat{q},\widehat{U}} (1-\lambda) \int_{\theta_L}^{\theta_H} \left[\theta_j \widehat{q}(\theta_j) - \widehat{q}(\theta_j)^2/2 - \widehat{U}(\theta_j)\right] G^{J-1}(\theta_j) \mathrm{d}G(\theta_j). \tag{18}$$

Given the equilibrium profit levels $\Pi_j^*(\sigma^*)$ defined in (13), the platform then requests the following advertising budget:

$$t^*(\sigma^*) = \Pi_j^*(\sigma^*) - \overline{\Pi}_j(\sigma^*). \tag{19}$$

Another way to interpret the advertising budget is the following: the sellers are willing to give up all the on-platform profit to participate, but need to be compensated for distorting the off-platform menus away from the monopoly benchmark.

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

$$\max_{\sigma} \left[t^*(\sigma)\right],$$

and we show that the efficient-steering mechanism indeed solves this problem.[13]

[13] Consistent with this result, Bergemann et al. (2023) show that the optimal managed campaign improves the platform's revenue relative to running auctions for targeted advertising slots. Furthermore, the experimental analysis of Decarolis et al. (2023) shows that, when a platform provides less detailed information to the bidders' algorithms, its revenues are "substantially and persistently higher."

Proposition 4 (Optimality of Efficient Steering)
The efficient-steering managed campaign mechanism maximizes the platform's profit.

The proof of this result establishes that the efficient-steering managed campaign attains an upper bound on the platform's profit. In particular, the equilibrium profits of the sellers $\Pi_j^*(\sigma^*)$ coincide with the vertically integrated benchmark where the platform controls all sellers' menus on both sales channels. Thus, this mechanism maximizes the sellers' and the platform's joint profits across the on- and off-platform markets. Moreover, the seller's outside option $\overline{\Pi}_j(\sigma)$ is bounded from below, for all managed campaigns σ, by the profits $\overline{\Pi}_j(\sigma^*)$ a seller can obtain through the optimal menu for the off platform consumers only. Because the advertising budget extracts each seller's surplus over and above their exogenous outside option, this mechanism maximizes the platform's revenue. Importantly, the efficient-steering managed campaign relies only on advertising and steering through sponsored products. In particular, the optimal revenue can be attained without levying commission or transaction fees on sellers or consumers.[14]

6 Value of Information and Privacy

In this section, we explore the platform's bargaining power by examining the role of its informational advantage and the implications of privacy policies. We start by removing the informational advantage of the platform in Section 6.1. Instead, we assume that every on-platform consumer learns their entire value profile θ, not just their value for the sponsored seller. One possible reason for this could be that reviews and recommendations are available on the platform and online more generally.

Next, in Section 6.2, we examine the role of price information. We consider the provision of organic search links by the platform that enable consumers to learn about all off-platform prices and products.

In Section 6.3, we introduce privacy policies that safeguard the consumers' information from the sellers. We consider cohort-based privacy protection where the platform informs the sellers only about the consumer's ranking of the sellers, while disclosing the exact value for the sponsored seller to the consumer. Consequently, the platform targets ads at the level of a *cohort* of consumers, and each consumer within a cohort have the same preference ranking over the J sellers.[15]

[14] In the presence of complete information by the platform, it suffices that the platform offers a single sponsored product rather many sponsored product slots. We suspect that in richer environments where consumers have some independent private information, many sponsored links would optimally screen for this additional information.

[15] This is in line with the recent Google Privacy Sandbox proposals to replace third-party cookies.

Lastly, in Section 6.4, we examine whether revealing the full value for the sponsored seller to the consumer maximizes the platform's revenue. We introduce information design in our managed campaign mechanism and provide conditions under which full or partial information revelation is optimal.

6.1 Symmetric Information

To assess the value of the platform's information advantage, we now assume all consumers who visit the platform learn their *entire value profile* θ (i.e., not just their value for the sponsored seller). The consumers off the platform remain imperfectly informed with expected value profile m. In Proposition 5, we establish that the ensuing symmetric information limits the platform's ability to steer the consumers' search behavior and reduces the advertising budget the platform can request from the sellers.

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

In equilibrium, both on- and off-platform consumers have the same information as in Section 5, where the platform has initially exclusive information about θ (henceforth, the "baseline" model). Thus, any seller who participates in the managed campaign mechanism offers the optimal menu in Proposition 3. Facing informed consumers, however, changes the seller's value of turning down the platform's offer, because consumers who know their values visit their favorite seller off-platform regardless of the identity of the sponsored seller on the platform. Suppose a consumer sees a product by a seller they did not expect on the platform. Under our symmetry refinement, this consumer continues to believe that all sellers offer symmetric menus off the platform.

Therefore, each seller j can choose not to participate in the managed campaign and poach any consumer to whom they offer the highest value: $\theta_j = \max_k \theta_k$. Seller j can achieves this by offering the consumers an off-platform information rent $\widehat{U}_j(\theta_j)$ above the level $\widehat{U}_j^*(\theta_j)$ in (16) which is offered by the competitors in equilibrium. When contemplating such a menu, the deviating seller j solves the following problem:

$$\hat{\Pi} \triangleq \max_{\widehat{q},\widehat{U}} \int_{\theta_L}^{\theta_H} [\theta_j \widehat{q}(\theta_j) - \widehat{q}(\theta_j)^2/2 - \widehat{U}(\theta_j)] \begin{bmatrix} (1-\lambda)G^{J-1}(\theta_j)g(\theta_j) \\ +\lambda F^{J-1}(\theta_j)f(\theta_j) \end{bmatrix} \mathrm{d}\theta_j \qquad (20)$$

$$\text{s.t. } \widehat{U}(\theta_j) \geq \widehat{U}_j^*(\theta_j). \qquad (21)$$

The equilibrium rent function (16) in the baseline model satisfies the constraint (21) and yields a strictly larger profit. Therefore, the sellers' outside option with known values exceeds the outside option $\overline{\Pi}$ of the baseline model characterized in (18).

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

$$\widehat{q}(\theta_j) = \max \left\{ 0, \theta_j - \frac{1 - \lambda F^J(\theta_j) - (1-\lambda) G^J(\theta_j)}{\lambda J F^{J-1}(\theta_j) f(\theta_j) + (1-\lambda) J G^{J-1}(\theta_j) g(\theta_j)} \right\}. \tag{22}$$

The equilibrium quality in (22) is larger for every value than the equilibrium $\widehat{q}_j^*(\theta_j)$ in (15) and yields higher utility to the consumers. Thus, constraint (21) does not bind in the optimal deviation—the best off-platform menu for seller j offers a higher utility level to j's favorite consumers than all other sellers' menus.[16]

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

To quantify the value of the platform's steering power, fix the value distribution F and consider the distribution G of expected values generated by an imperfectly informative signal that each (on- and off-platform) consumer observes about their value. Denote by $t^*(G)$ the equilibrium advertising budgets in (19) under distribution G. Now let the consumer's signal become arbitrarily precise, so that the distribution G of expected converges to the distribution F of values. The equilibrium menu for the limit case can be obtained by setting $G = F$ in (15). Proposition 5 then implies the following observation.

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

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

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

Corollary 1 has important implications for a platform's choice of information design to which we turn in Section 6.4. In particular, the equilibrium advertising budgets jump up as soon as the platform has any informational advantage relative to the consumers. This suggest that some degree of information asymmetry—revealing some additional information to consumers—is always part of the optimal design.

6.2 Organic Links

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

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

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

$$\widehat{U}_k(\theta_k^*(\theta_j)) = \widehat{U}_j(\theta_j). \tag{23}$$

With $\theta_k^* = \theta_k^*(\theta_j)$ defined as in (23), seller j's best-response problem is given by

$$\begin{aligned}
\max_{\widehat{q},\widehat{U}} (1-\lambda) &\int_{\theta_L}^{\theta_H} \underbrace{[\theta_j \widehat{q}(\theta_j) - \widehat{q}(\theta_j)^2/2 - \widehat{U}(\theta_j)]}_{\text{off-platform sales}} G^{J-1}(\theta_j) g(\theta_j) \mathrm{d}\theta_j \qquad (24)\\
&+ \lambda \int_{\theta_L}^{\theta_H} \underbrace{(\theta_j^2/2 - \widehat{U}(\theta_j))}_{\text{on-platform sales}} \min\{F^{J-1}(\theta_k^*(\theta_j)), F^{J-1}(\theta_j)\} f(\theta_j) \mathrm{d}\theta_j \\
&+ \lambda \int_{\theta_L}^{\theta_H} \underbrace{[\theta_j \widehat{q}(\theta_j) - \widehat{q}(\theta_j)^2/2 - \widehat{U}(\theta_j)]}_{\text{off-platform sales}} \max\{0, F^{J-1}(\theta_k^*(\theta_j)) - F^{J-1}(\theta_j)\} f(\theta_j) \mathrm{d}\theta_j.
\end{aligned}$$

The first term in (24) captures the off-platform consumers. The second term captures the sales to on-platform consumers for which seller j offers both the highest utility level $\widehat{U}_j$ and the highest marginal value θ_j. The third term, whenever positive, captures on-platform consumers with $\theta_j < \max_{k \neq j} \theta_k$ to whom seller j nonetheless offers the highest utility level $\widehat{U}_j$. Seller j is not advertised to these consumers, who instead showroom and buy from seller j off the platform.

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

Proposition 6 (Equilibrium with Organic Links)

1. *The equilibrium quality and utility levels $\widehat{q}_j^*(\theta_j)$ and $\widehat{U}_j^*(\theta_j)$ are weakly higher for all θ_j with organic links than without.*

2. *The sellers' profits are lower and their outside options are higher with organic links than without.*

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

Intuitively, the presence of organic information benefits consumers but reduces the platform's ability to restrain competition and extract surplus from sellers. In a symmetric equilibrium, each seller's market segment consists of all consumers who like the products the most. These market shares, unlike the baseline case, are endogenous to the choice of $\widehat{U}_j$. 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 profit of the sellers are lower, and consumer surplus is higher.

However, the sellers' profits net of the advertising budget are equal to the value of their outside option. With organic links, any seller can respond to competitors' prices without participating in the mechanism. Let θ_k^* be given by (23), with $\widehat{U}_k = \widehat{U}_j^*$. Because all of the sales necessarily happen off the platform, a deviating seller j can then obtain profit of

$$\widetilde{\Pi}_j \triangleq \max_{\widehat{q},\widehat{U}} \int_{\theta_L}^{\theta_H} \left[\theta_j \widehat{q}(\theta_j) - \widehat{q}(\theta_j)^2/2 - \widehat{U}(\theta_j)\right] \begin{bmatrix} (1-\lambda)\, G^{J-1}(\theta_j) g(\theta_j) \\ +\lambda F^{J-1}(\theta_k^*(\theta_j)) f(\theta_j) \end{bmatrix} \mathrm{d}\theta_j. \qquad (25)$$

Unlike in the baseline model, each deviating seller has the opportunity to win over some (but not necessarily all) on-platform consumers for which $\theta_j \geq \max_{k \neq j} \theta_k$. The outside option $\widetilde{\Pi}_j$ in (25) then exceeds the value $\overline{\Pi}_j$ defined in (18). In other words, the sellers' outside options are higher with organic links than without, and the equilibrium advertising budgets are correspondingly lower.

Finally, recall that with symmetric information and no organic links, the deviating seller wins *all* on-platform consumers for which $\theta_j \geq \max_{k \neq j} \theta_k$. Thus, the outside option $\hat{\Pi}_j$ defined in (20) is even higher than $\widetilde{\Pi}_j$ in (25). Because the equilibrium menus with organic links do not change if consumers know their values, it is possible that the platform might profitably raise the requested advertising budget by showing organic links if consumers are already fully informed about their values.

6.3 Privacy Protection

Up to this point, we have not limited the platform's ability to share information about the values of the buyers, θ, with the sellers. In reality, the extent of data sharing may be restricted by regulation or design choices made by the platform. Now, we assume that the value of the advertised product θ_j is shared with consumers on the platform, but not with sellers. Sellers are only allowed to base their offers on the *ranking* of consumer valuations θ_j within a *cohort* of consumers, where each consumer ranks the J sellers in the same way.

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

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

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

Proposition 7 (Cohort Targeting)
In the unique symmetric equilibrium, each seller offers quality levels

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

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 7 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 seller offers the same menu to consumers both on and off the platform. In this menu, the equilibrium quality schedule is the same as in (22), 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 value θ.

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

6.4 Information Design

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

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

We assume that the platform knows each consumer's value θ and their expected value m. This assumption simplifies the analysis, but it is also a reasonable approximation since if the platform knows every consumer's true preferences, it may also have information about their past *experiences*.[17] For instance, the platform may have access to the consumer's cookies and browsing history, which would enable it to estimate the consumer's expected value.

In a managed campaign, each seller j maximizes total profit by choosing prices and product qualities given the platform's designed information. The platform, in turn, maximizes the sellers' profits by choosing the distribution of expected values. Advertising budgets then extract the sellers' willingness to pay for this information. Hence, we can think of the platform as designing both the information and the prices to maximize the sellers' profits.

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

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

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

Large Platform We begin by considering the limit case where $\lambda = 1$. In other words, a measure one of consumers shop on the platform, and hence sellers have no reason to post off-platform menus. In this case, we show that the platform maximizes the advertising budgets

[17] See Liang and Madsen (2020) for a formal distinction.

by committing to the efficient managed campaign and by fully revealing θ_j to both consumer θ and to the sponsored seller j.

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

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

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

Because the off-platform consumers have no private information, each seller j offers just one product of quality $\widehat{q}_j \in \mathbb{R}_+$ off-platform at a price that extracts the consumer's expected willingness to pay, i.e., $\widehat{p}_j = \mu \widehat{q}_j$. As we know from our baseline model, the seller's choice of off-platform quality $\widehat{q}_j$ directly controls the information rent of all on-platform consumers. In particular, a consumer with an expected value θ_j (given the information revealed to them by the platform) obtains a rent

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

when buying from seller j. Consequently, seller j's on-platform profit as a function of the realized value θ_j are given by

$$\pi(\theta_j, \widehat{q}_j) = \theta_j^2/2 - U(\theta_j, \widehat{q}_j). \tag{26}$$

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

As a first step towards characterizing the optimal information design, we establish that the platform shows each consumer an ad by their favorite seller and reveals an informative

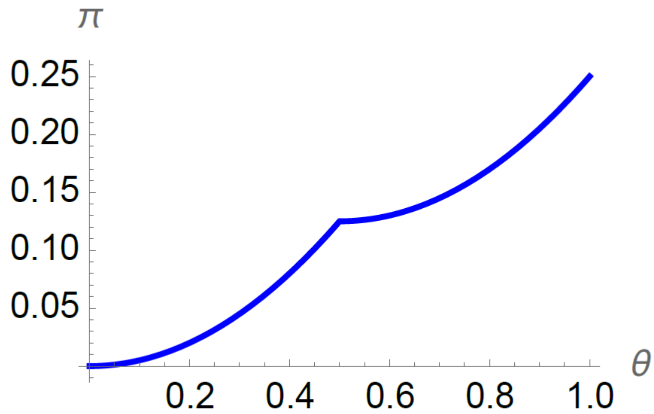

FIGURE 5: On-platform profit levels

signal about their value θ_{j^*}.

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

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

Therefore, we can write the platform's problem as

$$\max_{\widehat{q}_j, \widehat{F}} \left[(1-\lambda)(\mu\hat{q}_j - \widehat{q}_j^2/2) + \lambda \int_{\theta_L}^{\theta_H} (\theta_j^2/2 - \max\{0, (\theta_j - \mu)\widehat{q}_j\}) \mathrm{d}\widehat{F}(\theta_j)\right] \tag{27}$$
$$\text{s.t. } F^J \succ \widehat{F}.$$

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

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

1. *There exist two thresholds $x_1 \leq \mu \leq x_2$ such that the optimal distribution of posteriors $\widehat{F}^*(\theta_j)$ coincides with $F^J(\theta_j)$ on $[0, x_1]$ and $[x_2, 1]$ and has an atom at μ.*

2. The pair of optimal thresholds (x_1, x_2) *are the unique solution to*

$$\begin{aligned} x_1 + 2\widehat{q} &= x_2 \\ \mathbb{E}_{F^J}[\theta \mid x_1 \leq \theta \leq x_2] &= \mu. \end{aligned}$$

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

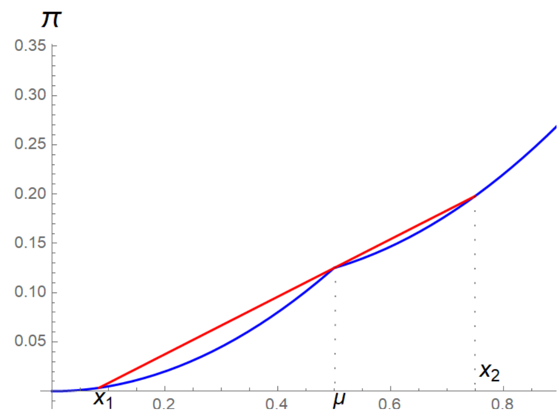

FIGURE 6: Optimal Separating and Pooling Intervals

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

$$\begin{aligned} \Pi(\widehat{q}) =& \lambda \int_0^{x_1(\widehat{q})} (\theta_j^2/2) \mathrm{d}F^J(\theta_j) + (\mu^2/2) \left(F^J(x_2(\widehat{q})) - F^J(x_1(\widehat{q}))\right) \\ &+ \lambda \int_{x_2(\widehat{q})}^1 \left(\theta_j^2/2 - \widehat{q}(\theta_j - \mu)\right) \mathrm{d}F^J(\theta_j) + (1-\lambda)\left(\mu\widehat{q} - \widehat{q}^2/2\right). \end{aligned}$$

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

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

7 Platform Size and Competition

Having examined the informational sources of the platform's bargaining power, we now return to our baseline setting of Section 4 to study the role of the size of the platform and the competition among the sellers. We first investigate how the market share of the platform λ affects the welfare and distribution of the social surplus. We then analyze how an increase in competition in terms of the number of competing sellers affects the welfare outcomes on and off the platform.

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 10.

Proposition 10 (Platform Size)

1. *The equilibrium quality levels $\widehat{q}_j^*(\theta_j)$ are decreasing in λ for all $\theta_j < \theta_H$, and the information rents $\widehat{U}_j^*(\theta_j)$ are decreasing in λ for all θ_j.*
2. *For every $\theta_j < \theta_H$, there exists $\bar{\lambda} < 1$ such that $\widehat{q}_j^*(\theta_j) = 0$ for all $\lambda \geq \bar{\lambda}$.*

In Figure 7, we illustrate how the off-platform quality provision changes as the market share λ of the platform increases.

The on-platform allocation remains unchanged and is given by the socially efficient quality provision. However, as the platform grows larger, each seller attempts to minimize the information rents on the platform and in turn renders the menu off the platform less attractive. Thus, for every value θ_j, the equilibrium quality-match off the platform $\widehat{q}_j^*(\theta_j)$ decreases, the price per unit of quality increases, and the consumer surplus $\widehat{U}_j^*$ off the platform decreases as the size of the platform increases.[18]

[18] Recent work by Valenzuela-Stookey (2022) considers many-to-many matching with congestion effects on each side in the Gomes and Pavan (2016) framework. In our model, congestion effects arise endogenously on the consumers' side, because the more consumers visit the platform the fewer options are available off-platform, which drives information rents down.

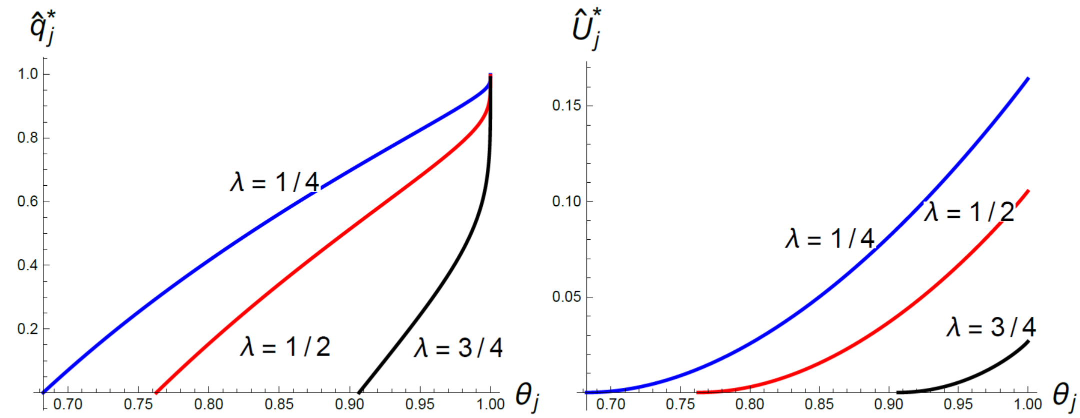

FIGURE 7: Off-Platform Menus, $J = 3, G(m_j) = m_j, F(\theta_j) = \text{Beta}(\theta_j, 1/4, 1/4)$.

Number of Sellers As the number of sellers increases, a larger number J of draws for each value 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 seller will know that every consumer who shops on their site, or receives their ads, has a value near θ_H with probability 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.

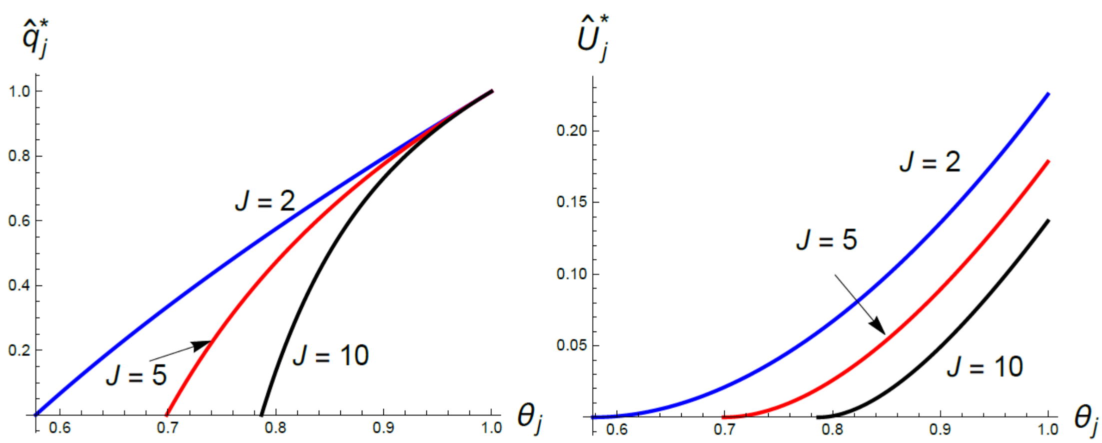

FIGURE 8: Off-Platform Menus, $\lambda = 0, G(m_j) = m_j, F(\theta_j) = \text{Beta}(\theta_j, 1/4, 1/4)$.

Relative to the Diamond (1971) model, quality distortions decrease faster in our setting for lower values and slower for higher values. This effect is due to the interaction of showrooming and the different distributions of values. In particular, one can show that the additional distortion term in the equilibrium quality (17), 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 θ_H. Thus, for a small number of sellers, high values of θ_j receive a higher quality as J increases while low values of θ_j receive a lower quality.

As Figure 9 illustrates, this effect may not be sufficient to generate a larger rent for any value. Furthermore, Proposition 11 shows that as J grows large, every value's quality allocation eventually decreases in the number of sellers.

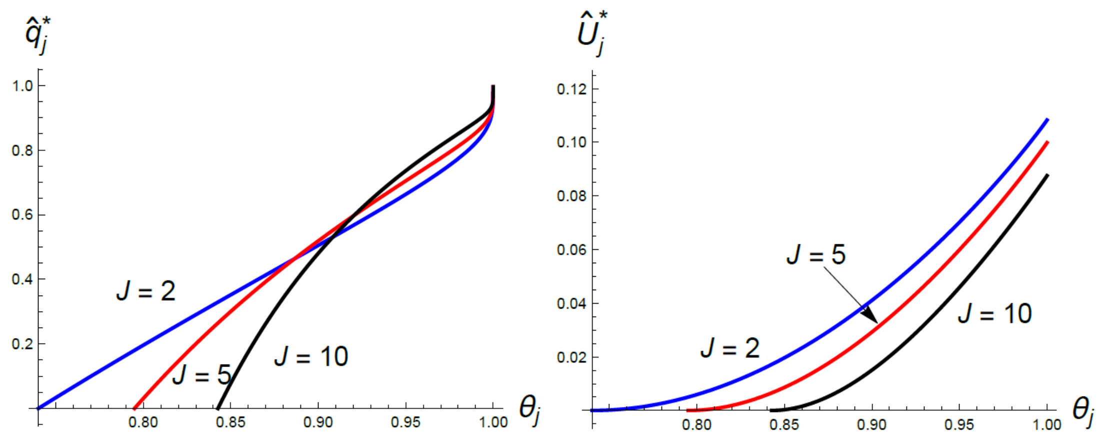

FIGURE 9: Off-Platform Menus, $\lambda = 1/2, G(m_j) = m_j, F(\theta_j) = \text{Beta}(\theta,1/4,1/4)$.

Proposition 11 (Number of Sellers)

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

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

We can then examine the impact of the size of the platform λ and of the number of sellers J on all parties' surplus levels. An immediate consequence of Propositions 10 and 11 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 value component is converging in probability to 1), and therefore sellers need not distort the off-platform menus when they participate in the platform's mechanism. Figure 10 illustrates the platform revenue, consumer surplus, seller surplus (i.e., the outside option $\overline{\Pi}_j$), and the profit generated on the platform for various numbers of sellers.

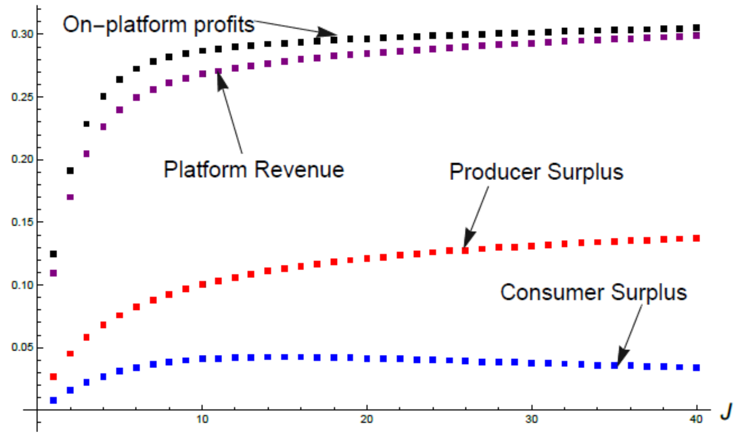

FIGURE 10: Surplus Levels, $\lambda = 2/3, G(m_j) = m_j, F(\theta_j) = \text{Beta}(\theta_j, 1/3, 1/3)$.

8 Conclusion

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

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

[19]The recent evidence in Mustri et al. (2022) is consistent with a mechanism like the one we outlined.

Appendix

This Appendix contains the proofs of all our results.

Proof of Proposition 1. To derive the low type's optimal quality, we substitute the expression for the binding incentive compatibility constraint for the high type (7) into both the on-platform and off-platform profit in objective (6). Differentiating with respect to $\widehat{q}\,(\theta_L)$ yields the result in (8). ■

Proof of Proposition 2. In any symmetric equilibrium, each consumer θ with expected value m learns their true value θ_{j^*} for the advertised seller j^* and believes that $j^* = \arg\max_j \theta_j$ with probability one. Moreover, each consumer expects identical menus to be posted off-platform and knows that the rent function $\widehat{U}_j(\theta_j)$ is strictly increasing in θ_j for all j. Therefore, the consumer searches for the advertised seller's off-platform prices. She does not search any further and does not learn any other seller's prices. Because the menu off the platform is incentive compatible, it is sufficient for consumer θ to compare the two items $q_{j^*}(\theta_{j^*})$ and $\widehat{q}_{j^*}(\theta_{j^*})$.

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

Proof of Proposition 3. Seller j's gross profit (i.e., after paying the platform's required advertising budget) can be written as

$$\begin{aligned} \max_{\widehat{q},\widehat{U}} \,(1-\lambda) \int_{\theta_L}^{\theta_H} & [\theta_j \widehat{q}(\theta_j) - \widehat{q}(\theta_j)^2/2 - \widehat{U}(\theta_j)] G^{J-1}(\theta_j) \mathrm{d}G(\theta_j) \\ + \lambda \int_{\theta_L}^{\theta_H} & [\theta_j^2/2 - \widehat{U}(\theta_j)] F^{J-1}(\theta) \mathrm{d}F(\theta), \\ \text{s.t. } \widehat{U}'(\theta_j) = \widehat{q}(\theta_j), & \text{ and } \widehat{U}(\theta_j) \geq 0 \text{ for all } \theta_j \in [\theta_L, \theta_H]\,. \end{aligned} \tag{28}$$

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

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

At a symmetric equilibrium, the optimality conditions are given by

$$\begin{aligned} (1-\lambda)\left(\theta_j - \widehat{q}(\theta_j)\right) G^{J-1}(\theta_j) g(\theta_j) + \hat{\gamma}(\theta_j) &= 0, \\ -(1-\lambda) G^{J-1}(\theta_j) g(\theta_j) - \lambda F^{J-1}(\theta_j) f(\theta_j) + \hat{\gamma}'(\theta_j) &= 0, \\ \hat{\gamma}(1) &= 0. \end{aligned} \tag{29}$$

Integrating, we obtain

$$\hat{\gamma}(\theta_j) = \frac{1}{J}\left((1-\lambda) G^J(\theta_j) + \lambda F^J(\theta_j) - 1\right). \tag{30}$$

Therefore, the equilibrium quality level is given by

$$\widehat{q}_j^*(\theta_j) = \theta_j - \frac{1-(1-\lambda) G^J(\theta_j) - \lambda F^J(\theta_j)}{(1-\lambda) J G^{J-1}(\theta_j) g(\theta_j)} \tag{31}$$

if the right-hand side is nonnegative, and nil otherwise, as in (15). This ends the proof. ■

Proof of Proposition 4. We argue the optimality of the managed-campaign mechanism in two steps. First, by Proposition 2, any seller j that does not participate in the mechanism does not make any sales to the consumers on the platform. This is because every consumer will see an ad by a different seller $k \neq j$ and will only consider on- and off-platform offers by seller k. Therefore, every seller's outside option consists of offering the optimal Mussa and Rosen (1978) menu to their off-platform consumers. This yields the profit level $\overline{\Pi}_j$ in (18), which is a fixed outside option independent of the menus posted by the other sellers.

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

The solution to the coalition profit-maximization problem coincides with the equilibrium outcome of the managed campaign mechanism. In this mechanism, each seller maximizes profit by offering the socially efficient quality $q_j(\theta) = \theta_j$ to each on-platform consumer θ. As a result, the platform assigns each seller to the consumers that value their products the most. Thus, the sellers-platform coalition's profits are maximized by the equilibrium menus that solve problem (28). Because the platform extracts the entire seller surplus in excess of the fixed outside option (18), no mechanism generates greater revenue for the platform. ■

Proof of Proposition 5. Suppose on-platform consumers know their value θ. If seller j participates in the mechanism but offers an out-of-equilibrium menu off the platform, only consumers who search for seller j in equilibrium observe this deviation. Therefore, every seller that participates in the mechanism can do no better than to advertise the efficient quality levels and post the off-platform menus that solve (28). However, if seller j does not participate, it can match the competitors' information rents $\widehat{U}_{k\neq j}$ and attract all the consumers on the platform who value their products the most. (Under the symmetric beliefs refinement, these consumers search for seller j's off-platform offer regardless of the ads shown to them by the platform.) Furthermore, if a nonparticipating seller j maximizes profit with respect to $(\widehat{q},\widehat{U})$ over the combined off- and on-platform market segments, it solves the problem in (20). This is a standard second-degree price discrimination problem where consumer values are distributed according to $\lambda F^J + (1-\lambda)G^J$. The optimal quality provision in such a deviation is given by

$$\widehat{q}(\theta_j) = \max\left\{0, \theta_j - \frac{1-(1-\lambda)\,G^J(\theta_j) - \lambda F^J(\theta_j)}{\lambda F^{J-1}(\theta_j)f(\theta_j) + (1-\lambda)\,G^{J-1}(\theta_j)g(\theta_j)}\right\}. \tag{32}$$

Because the quality level $\widehat{q}$ in (32) is pointwise larger than $\widehat{q}_j^*$ in (31), the resulting information rent is correspondingly higher for each θ_j. Thus, the deviating seller's optimal choice of menu yields an outside option $\hat{\Pi}_j$ larger than $\overline{\Pi}_j$ in (18). In addition, because the on-path profit are unchanged relative to the case of asymmetrically informed consumers, the advertising budget requested by the platform must decrease. ■

Proof of Proposition 6. We characterize all symmetric equilibria with full participation of the subgame following the platform's announcement of the required advertising budget. We first rewrite problem (24) as follows:

$$\begin{aligned}
\max_{\widehat{q},\widehat{U}}\,(1-\lambda)&\int_{\theta_L}^{\theta_H}[\theta_j\widehat{q}(\theta_j) - \widehat{q}(\theta_j)^2/2 - \widehat{U}(\theta_j)]G^{J-1}(\theta_j)g(\theta_j)\mathrm{d}\theta_j\\
&+\lambda\int_{\theta_L}^{\theta_H}[\theta_j\widehat{q}(\theta_j) - \widehat{q}(\theta_j)^2/2 - \widehat{U}(\theta_j)]F^{J-1}(\theta_k^*(\theta_j))f(\theta_j)\mathrm{d}\theta_j\\
&+\lambda\int_{\theta_L}^{\theta_H}(\theta_j^2/2 - \theta_j\widehat{q}(\theta_j) + \widehat{q}(\theta_j)^2/2)\min\{F^{J-1}(\theta_k^*(\theta_j)), F^{J-1}(\theta_j)\}f(\theta_j)\mathrm{d}\theta_j,
\end{aligned}$$

where, as in (23), θ_k^* satisfies

$$\widehat{U}_k\left(\theta_k^*(\theta_j)\right) = \widehat{U}_j\left(\theta_j\right).$$

The associated Hamiltonian can be written as

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

Totally differentiating (23), we obtain

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

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

$$(1-\lambda)\left(\theta_j - \widehat{q}(\theta_j)\right) G^{J-1}(\theta_j) g(\theta_j) + \gamma(\theta_j) = 0, \tag{33}$$

$$\begin{aligned} -\lambda\frac{\left(J-1\right)F^{J-2}(\theta_j) f^2(\theta_j)}{\widehat{q}(\theta_j)} &\left(\alpha\theta_j^2/2 + (1-\alpha)\left(\theta_j\widehat{q}(\theta_j) - \widehat{q}(\theta_j)^2/2\right) - \widehat{U}(\theta_j)\right) \\ &+ (1-\lambda)\, G^{J-1}(\theta_j) g(\theta_j) + \lambda F^{J-1}(\theta_j) f(\theta_j) = \gamma'(\theta_j), \end{aligned} \tag{34}$$

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

Now recall the costate equation (29) in the baseline model. Because the last two terms on the left-hand side of (34) are 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

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

Together with (29) and (33), this implies that the quality and utility levels $\widehat{q}_j(\theta_j)$ and $\widehat{U}_j(\theta_j)$ are weakly higher for all θ_j *with* organic links than without.

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

$$\begin{aligned}
\Pi_j(\widehat{q}, \widehat{U}) &= (1-\lambda) \int_{\theta_L}^{\theta_H} \left[\theta_j \widehat{q}(\theta_j) - \widehat{q}(\theta_j)^2/2 - \widehat{U}(\theta_j)\right] G^{J-1}(\theta_j) \mathrm{d}G(\theta_j) \\
&\quad + \lambda \int_{\theta_L}^{\theta_H} (\theta_j^2/2 - \widehat{U}(\theta_j)) F^{J-1}(\theta_j) \mathrm{d}F(\theta_j).
\end{aligned} \tag{35}$$

The equilibrium menu in the baseline model $(\widehat{q}_j^*, \widehat{U}_j^*)$ maximizes (35), while the equilibrium menu maximizes (24) and hence achieves a weakly lower profit level. Now consider the deviating seller's profit. For any choice of $(\widehat{q}, \widehat{U})$ off path, the deviation profit are weakly larger with organic links than without. Without organic links, the deviating seller posts the Mussa and Rosen (1978) menu and makes no sales on the platform. With organic links and posting the same menu, the seller wins a fraction $F^{J-1}(\theta_k^*(\theta_j)) \in [0, 1]$ of on-platform values θ_j. Consequently, we have $\widetilde{\Pi}_j \geq \overline{\Pi}_j$, which implies *a fortiori* that the advertising budgets are lower. ■

Proof of Proposition 7. We construct an equilibrium where each seller j sets on- and off-platform menus to maximize profit, given that seller j expects to face all consumers that rank j the highest. Therefore, consider the joint optimization problem over menus (q, U) and $(\widehat{q}, \widehat{U})$ when facing distributions F^J and G^J, respectively, under the showrooming constraint. Seller j solves:

$$\begin{aligned}
&\max_{q, \widehat{q}, U, \widehat{U}} \left[\begin{array}{c} \lambda \int_{\theta_L}^{\theta_H} (\theta_j q(\theta_j) - q(\theta_j)^2/2 - U(\theta_j))\, F^{J-1}(\theta_j) \mathrm{d}F(\theta_j) \\ + (1-\lambda) \int_{\theta_L}^{\theta_H} \left(\theta_j \widehat{q}_j(\theta_j) - \widehat{q}(\theta_j)^2/2 - \widehat{U}(\theta_j)\right) G^{J-1}(\theta_j) \mathrm{d}G(\theta_j) \end{array} \right] \qquad (36) \\
\text{s.t.}\quad & U'(\theta_j) = q(\theta_j) \\
& \widehat{U}'(\theta_j) = \widehat{q}(\theta_j) \\
& U(\theta_j) \geq \widehat{U}(\theta_j) \geq 0.
\end{aligned}$$

We now show that the solution to (36) is given by

$$q_j^*(\theta_j) = \widehat{q}_j^*(\theta_j) = \theta_j - \frac{1 - (1-\lambda)\, G^J(\theta_j) - \lambda F^J(\theta_j)}{J \lambda F^{J-1}(\theta_j) f(\theta_j) + J\,(1-\lambda)\, G^{J-1}(\theta_j) g(\theta_j)} \tag{37}$$

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

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

The pointwise necessary conditions for this problem are the following:[20]

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

If $q(\theta_j) = \widehat{q}(\theta_j)$ as in (37), 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. ■

Proof of Proposition 8. When the platform becomes arbitrarily large ($\lambda \to 1$), rents off-platform vanish and the sponsored seller appropriates the entire surplus it generates.

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

The sponsored seller's profit under complete and symmetric information on each value θ_j is then given by the first-best surplus $\pi_j^*(\theta) = \theta_j^2/2$. Fix any matching mechanism, and let F^* denote the distribution of θ_j that are matched to seller j. Therefore, the information design problem of the platform is given by

$$\max_{\hat{F} \prec F^*} \int_{\theta_L}^{\theta_H} \pi_j^*(\theta_j) \, \mathrm{d}\hat{F}(\theta_j).$$

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

Proof of Proposition 9. Fix $\widehat{q}$ and let $\pi(\theta_j)$ denote the online profit function. By Dworczak and Martini (2019) (Theorem 1), if there exists a distribution $\hat{F} \prec F^J$ and a convex supporting function $y(\theta_j)$ such that

$$y(\theta_j) \geq \pi(\theta_j)$$
$$\int_{\theta_L}^{\theta_H} y(\theta_j)\mathrm{d}F^J(\theta_j) = \int_{\theta_L}^{\theta_H} y(\theta_j)\mathrm{d}\hat{F}(\theta_j)$$
$$\mathrm{supp}(\hat{F}) \subseteq \{\theta_j \in [\theta_L, \theta_H] : y(\theta_j) = \pi(\theta_j)\},$$

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

Because our function $\pi(\theta_j)$ satisfies the Dworczak and Martini (2019) regularity conditions, computing the supporting function and the associated distribution is relatively easy: by Dworczak and Martini (2019) (Proposition 1), the support of the optimal $\hat{F}$ can be found by solving

$$\min_y \int_{\theta_L}^{\theta_H} y(\theta_j)\mathrm{d}F^J(\theta_j)$$
$$\text{s.t. } y(\theta_j) \geq \pi(\theta_j) \; \forall \theta_j$$
$$y \text{ convex.}$$

Moreover, the optimal $\hat{F}$ is then supported only on points where $y^* = \pi$. This result allows us to compute the supporting function independently of the distribution.

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

$$x_2^2/2 - (x_2-\mu)\widehat{q} = \mu^2/2 + s(x_2-\mu)$$
$$\mu^2/2 - s(\mu - x_1) = x_1^2/2.$$

Therefore, solving we have

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

and we can write the objective as

$$\min_s \Bigg[\int_0^{x_1(s)} (\theta_j^2/2)\mathrm{d}F^J(\theta_j) + \int_{x_1(s)}^{x_2(s)} \left(x_1^2(s)/2 + s(\theta_j - x_1(s))\right) \mathrm{d}F^J(\theta_j) + \int_{x_2(s)}^1 \left(\theta_j^2/2 - \widehat{q}(\theta_j-\mu)\right) \mathrm{d}F^J(\theta_j)\Bigg].$$

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

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

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

Finally, note that

$$\Pi'(\widehat{q}) = -\lambda\left(\left(x_2^2 - \mu^2\right)/2 - \widehat{q}(x_2-\mu)\right) dF^J(x_2(s))x_2'(q) - \lambda \int_{x_2(s)}^1 (\theta_j-\mu)\mathrm{d}F^J(\theta_j) + (1-\lambda)(\mu - \widehat{q})$$

hence

$$\widehat{q} = \mu - \frac{\lambda}{1-\lambda}\left[\int_{x_2(s)}^1 (\theta_j - \mu)\mathrm{d}F^J(\theta_j) + \left(\frac{x_2^2-\mu^2}{2} - \widehat{q}(x_2-\mu)\right) JF^{J-1}(x_2) f(x_2)\, x_2'(q)\right]$$

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

Therefore $\widehat{q}$ is decreasing in λ and consequently x_2 is decreasing and x_1 increasing in λ. ■

Proofs of Proposition 10. These results are obtained by differentiating expression (15) for the equilibrium quality off-platform with respect to λ. In particular, whenever it is strictly positive, the equilibrium $\widehat{q}_j^*(\theta_j)$ is strictly decreasing in λ. Because the equilibrium quality provision is equal to the marginal information rent, the comparative statics of quality $\widehat{q}_j^*$ immediately extend to the information rent $\widehat{U}_j^*$. ■

Proof of Proposition 11. These results are obtained by differentiating expression (15) with respect to J. Whenever it is strictly positive, the equilibrium $\widehat{q}_j^*(\theta_j)$ is strictly decreasing in J for J large enough. Because the equilibrium quality provision is equal to the marginal information rent, the comparative statics of quality $\widehat{q}_j^*$ immediately extend to the information rent $\widehat{U}_j^*$. ■

References

Acemoglu, D., Makhdoumi, A., Malekian, A., and Ozdaglar, A. (2022). Too much data: Prices and inefficiencies in data markets. *American Economic Journal: Microeconomics*, 14(4):218–256.

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

Aggarwal, G., Badanidiyuru, A., and Mehta, A. (2019). Autobidding with constraints. In *International Conference on Web and Internet Economics*, pages 17–30. Springer.

Anderson, S. P. and Bedre-Defolie, Ö. (2021). Hybrid platform model. Technical Report 16243, CEPR.

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

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

Balseiro, S. R., Deng, Y., Mao, J., Mirrokni, V. S., and Zuo, S. (2021). The landscape of auto-bidding auctions: Value versus utility maximization. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, pages 132–133.

Bar-Isaac, H. and Shelegia, S. (2020). Search, showrooming, and retailer variety. Technical report, CEPR Discussion Paper No. DP15448.

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

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

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

Bergemann, D., Bonatti, A., and Wu, N. (2023). Managed campaigns and data-augmented auctions for digital advertising. Technical report, Cowles Foundation for Research in Economics.

Bergemann, D., Brooks, B., and Morris, S. (2015). The limits of price discrimination. *American Economic Review*, 105:921–957.

Bergemann, D., Brooks, B., and Morris, S. (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*, pages 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 Denicolo, V. (2015). Exclusive contracts and market dominance. *American Economic Review*, 105(11):3321–51.

Chen, D. (2022). The market for attention. Technical report, Stanford GSB.

Choi, J., Jeon, D., and Kim, B. (2019). Privacy and personal data collection with information externalities. *Journal of Public Economics*, 173:113–124.

Condorelli, D. and Szentes, B. (2022). Bargaining rents vs. efficient sorting in two-sided matching by an informed platform. Technical report, University of Warwick and LSE.

Crémer, J., de Montjoye, Y.-A., and Schweitzer, H. (2019). Competition policy for the digital era. Technical report, European Commission.

de Cornière, A. and de Nijs, R. (2016). Online advertising and privacy. *Rand Journal of Economics*, 47(1):48–72.

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

Decarolis, F., Rovigatti, G., Rovigatti, M., and Shakhgildyan, K. (2023). Artificial intelligence, algorithmic bidding and collusion in online advertising. Technical Report 18009, CEPR.

Deng, Y., Mao, J., Mirrokni, V., and Zuo, S. (2021). Towards efficient auctions in an auto-bidding world. In *Proceedings of the Web Conference 2021*, pages 3965–3973.

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

Donnelly, R., Kanodia, A., and Morozov, I. (2022). Welfare effects of personalized rankings. *Available at SSRN 3649342*.

Dworczak, P. and Martini, G. (2019). The simple economics of optimal persuasion. *Journal of Political Economy*, 127(5):1993–2048.

Elliott, M., Galeotti, A., and Koh, A. (2020). Market segmentation through information. Technical report, Cambdrige University.

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

Gomes, R. and Pavan, A. (2016). Many-to-many matching and price discrimination. *Theoretical Economics*, 11(3):1005–1052.

Gomes, R. and Pavan, A. (2022). Price customization and targeting in matching markets. *RAND Journal of Economics*, forthcoming.

Gur, Y., Macnamara, G., Morgenstern, I., and Saban, D. (2022). Information disclosure and promotion policy design for platforms. Technical report, Stanford University.

Gutierrez, G. (2022). The welfare consequences of regulating amazon. Technical report, University of Washington.

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

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

Hidir, S. and Vellodi, N. (2021). Privacy, personalization, and price discrimination. *Journal of the European Economic Association*, 19(2):1342–1363.

Ichihashi, S. (2020). Online privacy and information disclosure by consumers. *American Economic Review*, 110(2):569–595.

Ichihashi, S. (2021). The economics of data externalities. *Journal of Economic Theory*, 196:105316.

Idem, B. (2021). Coexistence of centralized and decentralized markets. Technical report, Pennsylvania State University.

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

Inderst, R. and Ottaviani, M. (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 Muir, E. (2021). Contracting and vertical control by a dominant platform. Technical report, Stanford University.

Ke, T. T., Ling, S., and Lu, M. Y. (2022). Information design of online platforms. Technical report, Chinese University of Hong Kong.

Kirpalani, R. and Philippon, T. (2021). Data sharing and market power with two-sided platforms. Technical Report 28023, NBER.

Kolotilin, A., Li, M., Mylovanov, T., and Zapechelnyuk, A. (2017). Persuasion of a privately informed receiver. *Econometrica*, 85(6):1949–1964.

Lam, H. T. (2021). Platform search design and market power. Technical report, Northwestern University.

Lee, C. (2021). Optimal recommender system design. Technical report, University of Pennsylvania.

Lee, K.-H. and Musolff, L. (2021). Entry into two-sided markets shaped by platform-guided search. Technical report, Princeton University.

Liang, A. and Madsen, E. (2020). Data and incentives. Technical report, Northwestern University.

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

Mekonnen, T., Murra-Anton, Z., and Pakzad-Hurson, B. (2023). Persuaded search. Technical report, Brown University.

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

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

Mustri, E. A. S., Adjerid, I., and Acquisti, A. (2022). Behavioral advertising and consumer welfare: An empirical investigation. Technical report, Carnegie Mellon University.

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

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

Raval, S. (2020). ebay makes search more efficient through personalization. *Available at https://tech.ebayinc.com/product/ebay-makes-search-more-efficient-through-personalization/*.

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

Shaked, M. and Shanthikumar, J. (1994). *Stochastic Orders and their Applications*. Academic Press, San Diego.

Shi, P. (2022). Optimal match recommendations in two-sided marketplaces with endogenous prices. Technical report, University of Souther California.

Teh, T.-H. and Wright, J. (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.

Valenzuela-Stookey, Q. (2022). Platform-mediated competition. Technical report, UC Berkeley.

Wang, C. and Wright, J. (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.

Zhu, R., Moorthy, S., and Shi, X. (2022). Advertising platforms and privacy. Technical report, University of Toronto.