“Sherlocking” and Platform Information Policy*

Jay Pil Choi† Kyungmin Kim‡ Arijit Mukherjee§

March 7, 2024

Abstract

Platform-run marketplaces may exploit third-party sellers' data to develop competing products, but potential for future competition can deter sellers' entry. We explore how this trade-off affects the platform's referral fee and its own entry decision. We first characterize the platform's optimal referral fee under full commitment on entry decision and study its economic implications. We then analyze the extent to which the platform's own information sharing policy substitutes for its commitment to entry. We characterize the platform's optimal information policy and examine how it interacts with the platform's fee structure. Our findings highlight the importance of considering the platform's fee structure as a regulatory response in the policy debates on marketplace regulation.

JEL Classification Numbers: D82, D42, L1

Keywords: Platform marketplace, Referral fee, Entry, Information design

*For their helpful comments and suggestions, we thank Joseph Harrington, Justin Johnson, Jeanine Miklós-Thal, Volker Nocke, Martin Peitz, Yossi Spiegel, Tat-How Teh, Chengsi Wang, Julian Wright, and the seminar/conference participants at IIFT (Kolkata), Michigan State University, Seoul National University, Antitrust Economics & Competition Policy Workshop at Simon Business School (University of Rochester), 20th Annual Berkeley-Columbia-Duke-MIT-Northwestern IO Theory Conference (Northwestern University), Asian Meeting of the Econometric Society, and NUS Platform Workshop Program.

†Department of Economics, Michigan State University. Email: choijay@msu.edu.

‡Department of Economics, Emory University, Email: kyungmin.kim@emory.edu.

§Department of Economics, Michigan State University. arijit@msu.edu.

1 Introduction

Platform-run marketplaces facilitate product discovery by enabling consumers to find obscure niche offerings that match their preferences.[1] However, it is common practice for such platforms to adopt a hybrid business model where they not only earn revenue from charging a referral fee from their third-party sellers but also compete with the sellers by introducing their own private-label products. Often accused of being "both player and referee," the hybrid marketplace platform can be an uneven playing field. The platform can potentially use the third-party sellers' market data to design its own products and promote its private-label products over its competitors' offerings.[2]

For example, consider Apple's App Store, a gatekeeper marketplace for iPhone and iPad platforms. The App Store is the only channel through which app developers can distribute their apps to the end users. However, Apple is also a provider of apps that potentially compete against third-party apps, and has been accused of engaging in anti-competitive conduct often referred to as "Sherlocking": it allegedly uses market data to target and copy profitable third-party apps, rendering them obsolete and driving the third-party developers out of business.[3] Similar concerns were also raised for the leading e-commerce platform, Amazon Marketplace, as it is alleged to have improperly shared third-party sellers' data with the division in charge of private-level product developments (Mattiolli, 2020).

This type of predatory behavior by dominant platforms and its associated concerns have led to a variety of policy proposals to limit such exploitative conducts and ensure fair competition. For example, in the U.S., policy makers have proposed structural separation that would prohibit hybrid business models by dominant platforms (Warren, 2019). The proposal would designate large tech platforms as "Platform Utilities" whereby Amazon Marketplace and Basics, and Google's ad exchange and businesses on the exchange would be split apart into separate companies. In contrast, the EU Digital Markets Act calls for behavioral re-

[1]The rapid increase in the availability of product variety and expansion of market share of niche products with the emergence of online retailers is often referred to as the "long tail" effect (Anderson, 2006), and has been explored by several scholars in both economics and management literature (Brynjolfsson et al., 2011; Yang, 2013; Goldfard and Tucker, 2019).

[2]To quote Margrethe Vestager, European Competition Commissioner and Vice-President of the European Commission, "the decisions that gatekeepers take, about how to rank different companies in search results, can make or break businesses in dozens of markets that depend on the platform. And if platforms also compete in those markets themselves, they can use their position as player and referee to help their own services succeed, at the expense of their rivals." Speech by Executive Vice-President Margrethe Vestager: Building trust in technology, 29 October, 2020. (Available at: https://ec.europa.eu/commission/commissioners/2019-2024/vestager/announcements/speech-executive-vice-president-margrethe-vestager-building-trust-technology_en)

[3]The term was coined in early 2000 when Apple updated its own app "Sherlock," a search tool on its desktop operating system, to subsume all features that a third-party app named "Watson" was offering on its platform.

strictions on the use of proprietary data generated through activities by the sellers and end users on the gatekeeper platforms. It bans using third-party data when the gatekeepers compete with them on their own platform. In addition, the hybrid platform gatekeepers are prohibited from ranking their own products or services in a more favorable manner compared to those of third parties.

However, even though "Sherlocking" may be ex-post optimal to the platform, it may deter sellers' entry, which is not only socially sub-optimal but also detrimental to the platform. Indeed, the platform may want to limit the extent of its own entry if it can commit to one, so as to balance its ex-post gains from imitating the sellers' product against ex-ante loss from reduced entry by the third-party sellers.[4] The choice of its referral fee also accounts for this trade-off, and any policy regulation on the platform's behavior must account for its regulatory response in fee structure.

The goal of this paper is to analyze the interplay among the platform's fee structure, entry decision, and data usage policy in the face of this trade-off. We first analyze the optimal fee structure of the platform when it can commit to its entry policy. Next, we explore how and when the platform can use its data sharing/usage policy to achieve the outcome under entry commitment (when it cannot directly commit to its entry decision). Finally, we draw out the implications of our findings for some of the key policy proposals on platform regulation.

We develop a tractable model of a platform-run marketplace where the platform charges a referral fee to the sellers for access to the marketplace and may also subsequently launch its own private-label product by copying the seller. The game unfolds as follows: first, a third-party seller privately observes his "type" and, given the referral fee, decides whether to enter the marketplace by incurring a fixed cost. The seller's type determines the profitability of his product. After the seller's entry, the platform may subsequently observe the type, and decide whether to enter the market by imitating the seller's product.

Our first set of results explore how the optimal referral fee varies with the platform's commitment power over its entry decision. If the platform operates only as a marketplace (instead of operating in hybrid mode) its optimal fee trades off extracting more from the entering sellers against encouraging sellers' entry. How does the fee structure change if the platform adopts a hybrid mode? The answer depends on whether or not the platform can commit to its entry policy.

If the platform can enter but cannot commit to its entry policy, then it sets an exceedingly high referral fee which can serve as a commitment device. In the absence of any entry commitment, the platform would enter whenever it is profitable to do so. Anticipating that,

[4] For example, Gower and Henderson (2007) find that Intel tends to avoid competing directly with third-party developers of devices built on its microprocessors so as to stimulate entry of potential developers.

a seller, particularly the lucrative types, may stay out of the marketplace as he expects the platform to subsequently imitate his product and put him out of business. But a high referral fee acts as a commitment device for the platform not to enter the market because it raises the platform's own opportunity cost of entry. By raising the referral fee, the platform loses the relatively less profitable types of the seller, but the resulting loss is more than compensated as the high fee would encourage the more lucrative seller types to enter (ones who were staying out due to imitation threat). We show that the platform may raise its fee to a sufficiently high level—one that is at least as large as the optimal fee when the platform operated only as a marketplace—so that product imitation is never optimal regardless of the seller's type.

However, if the platform can use, and commit to, a nuanced (type-contingent) entry policy, then it would set a fee that is even lower than the one it would have set if it were to operate only as a marketplace. When the platform can commit to its entry decision, it can induce even a high-type seller—ones that are more vulnerable to imitation—to enter by limiting its own likelihood of entry, and extract (a part of) the rent the seller earns on the marketplace. In particular, the platform extracts all rents from the high-type sellers who are worth imitating, and have a stronger incentive to induce more seller types to enter. Consequently, it further lowers its referral fees.

While the platform is better off when it can commit to its entry decision, in many scenarios it may lack such commitment power, especially when the platform can capture a third-party seller's market share by introducing a close substitute that may not be a direct imitation of the seller's product. For example, when Apple allegedly "sherlocked" Watson, it did not develop a new app by copying Watson's, but updated one of its existing tools (named "Sherlock") to offer the same functionality that Watson offered. Such possibility of "inventing around" may pose a challenge in enforcing the platform's commitment to its entry decision due to ex-post verifiability of imitation. In fact, the "doctrine of equivalents" in the enforcement of patents extends the patentee's rights beyond the literal limits of the written claims to regulate imitations that deliberately design around an invention.[5] As a result, this doctrine creates significant legal uncertainty and makes patent disputes inevitable.

However, it might be possible for the platform to commit to its data usage policy. Such a policy stipulates the extent to which the platform's marketplace division shares third-party sellers' data with its product division, and the nature of the shared data may be discoverable

[5] According to the U.S. Supreme Court, the doctrine of equivalents is needed because "to permit imitation of a patented invention which does not copy every literal detail would be to convert the protection of the patent grant into a hollow and useless thing. Such a limitation would leave room for—indeed encourage—the unscrupulous copyist to make unimportant and insubstantial changes." (Graver Tank & Mfg Co v Linde Air Prods Co, 339 US 605, 607 (1950).)

and verifiable by the court of law. Indeed, various regulatory proposals on data usage, e.g., the EU Digital Markets Act, that seek to regulate the level of aggregation and anonymity in third-party data shared between the platform divisions rely on the verifiability of the shared data. One may also interpret the platform's data usage policy as a "privacy policy" offered to third-party sellers. Privacy policies specify what data is collected by the platform for its own commercial purposes, and are routinely used in e-commerce.[6]

If the platform cannot commit to entry, can it use its data sharing policy to implement the entry-commitment outcome? To explore this question, we model the platform's data sharing policy as an information design problem. Instead of directly observing the seller's type, the platform commits to observe only a signal on the seller's type that is generated from a pre-specified signal structure. We assume that the platform chooses its referral fee and the signal structure at the beginning of the game.

We derive a general condition under which the platform can secure its entry-commitment payoff through its information sharing policy. The condition holds if the cost of entry is relatively large and it takes relatively longer for the platform to imitate the seller's product. If this condition fails, then the platform must exclude some of the intermediate types of the seller. The optimal referral fee now balances a trade-off between the gains from enhanced entry by the seller (particularly of the intermediate types) and the loss due to distortions from the optimal fee under entry commitment.

Our findings have sharp implications for key policy recommendations on regulation of hybrid marketplaces, such as a ban on hybrid mode and a ban on the use of third-party sellers' information for the launch of private-level products. Due to the regulatory response of the platform in its choice of referral fee, the welfare implications of such policies are often ambiguous, and under some settings, they could be welfare reducing.

For example, as discussed earlier, under no regulation on the mode of operation and information usage, in equilibrium, the platform may be able to implement the entry-commitment outcome through its information policy. Moreover, the associated referral fee is lower than its counterpart when the platform operates only as a marketplace. Thus, banning hybrid mode may result in a much higher referral fee that stifles sellers' entry and reduces the welfare for both the sellers and the consumers.

A policy that only prohibits information usage (i.e., allows hybrid mode as long as the platform does not use its proprietary data for its own entry decision) may also have a negative welfare implication. In this case, the platform would infer the sellers' type from their entry decision, and we show that depending on the parameters, one of two equilibria may be played. In one, the marketplace may cease to exist as no seller type enters, and in the other

[6]For a recent survey of data markets, see Bergmann and Bonatti (2019).

some types of the seller enter but the platform never finds it optimal to imitate the seller. If the former equilibrium is played, the policy is clearly welfare reducing. And even if the latter equilibrium is played, the welfare implication of such a policy remains ambiguous, as the optimal referral fee under marketplace mode and hybrid mode cannot be ranked a priori.

Related literature. Our paper is related to several recent papers that address various issues associated with platform marketplaces. Madsen and Vellodi (forthcoming) analyze implications of platforms' use of proprietary marketplace sales data to target the introduction of private-level products. They explore how innovation incentives in digital markets can be shaped by various policies/regulations on the platforms' data usage. Even though there are some parallels between their paper and ours in terms of motivations and questions addressed, our paper differs from theirs in three major respects. First, they focus on the case where ad-valorem referral fee is fixed and do not characterize the optimal referral fee. In contrast, the interplay among the optimal referral fee, entry decisions, and data usage is one of the main foci of our paper. Second, they consider the seller's ex ante entry incentive (without knowing his type), while we focus on the seller's interim incentive. This drives several significant differences in the results and analysis. For example, in their paper the seller's (entrepreneur's) entry is always characterized by a cutoff type (cost), which is no longer the case in our model (see Section 5). Finally, they consider a regulator's problem regarding when to allow the platform to observe (what) marketplace data, while we focus on the platform's own data policy. In other words, they study a regulator's information design problem, while ours is the platform's self information design problem.

In another recent paper, Hagiu et al. (2022) build a model of platform that can choose to offer a marketplace to innovative third-party sellers and convenience benefits to consumers for transactions. They consider three types of platform business models: marketplace mode, seller mode, and dual mode where a platform sells on its own marketplace.[7] Given antitrust concerns about dominant platforms' adoption of the dual mode along with product imitation and self-preferencing, they analyze the effects of various policy options including an outright ban on the dual mode. They show that the policy outcomes crucially depend on the platform's policy-induced endogenous choice of business models.

However, the setup and focus of their model are very different from ours. For instance, the dual mode platform in their model can still sell its own existing product in the absence of entry by an innovative seller, and in their "exploitative" equilibrium a referral fee is set high

[7]Platforms in dual mode are also called "hybrid platforms" (Anderson and Bedre-Defolie, *forthcoming*), "Hybrid marketplaces" (Etro, 2023) or "retailer-led marketplaces" (Hervas-Drane and Shelegia, 2021). For papers that study the choice of business models, but without the possibility of dual mode, see Hagiu and Wright (2015, 2019) and Johnson (2017).

to shield its own product from the innovative seller. In our model, entry by a third party is essential for the platform to sell its own product because we focus on the product discovery aspect of marketplaces; we envision a situation with so many potential products that platforms are not expected to possess information about each individual product's market demand and it is not in the interest of the platform to sell a product without first observing its demand through a marketplace. Thus, the role of a referral fee is very different in our model in that it serves as a commitment device not to imitate the third-party seller's product to induce more entry. In an extension of their baseline model, they also consider the possibility of product imitation and self-preferencing by marketplace platforms. In particular, they consider a case in which the platform is able to imitate with an *exogenously* given parameter and show that third-party sellers' innovation level can increase with constrained imitation. In contrast, we derive the full commitment solution and investigate how such a solution can be achieved with an internal information policy in the absence of full commitment power by the platform.

Anderson and Bedre-Defolie (forthcoming) develop an analytical framework to investigate interactions between monopolistically competitive third-party sellers and a hybrid platform. As in this paper, the determination of a referral fee is a central question. However, the referral fee in their model is used as a mechanism of "insidious steering" by the hybrid platform to steer demand to its own products. In addition, there is no information elicitation from third-party sellers in their model.

In focusing on the product discovery aspect of marketplaces and addressing the issue of free-riding on the information provided by third-party sellers, our paper is related to Hervas-Drane and Shelegia (2022). They also assume the existence of "unobserved" products whose existence can be identified only through running a marketplace with dispersed information. In such a framework, they investigate incentives for a retailer to switch from a pure seller to a dual mode and the competitive effects from third parties associated with such a switch. One feature of their model is that the entry decision depends only on the referral fee, but is independent of the platform's ex post entry decision because they assume no fixed cost of entry. Thus, the issue of commitment does not arise in their model. In addition, their paper does not address the relevant issues from the information design perspective.

Finally, on the empirical side, Zhu and Liu (2018) investigate Amazon's entry pattern into third-party sellers' product spaces and show that Amazon is more likely to target successful product spaces, which empirically validates our modeling approach. Raval (2022) shows that Amazon steers consumers to a first-party offer sold by its retail arm ("Amazon Retail") over third-party offers. In a similar vein, Chen and Tsai (2021) analyze Amazon's "Frequently Bought Together" recommendation system and demonstrate that hybrid platforms have biased incentives for product recommendation.

This paper is structured as follows. Section 2 presents our model. We present two benchmark analyses in Section 3 that highlight how the potential of platform's entry (or lack thereof) affects the seller's entry decision. Section 4 explores the case where the platform has full commitment power over its entry decision. The optimal information policy is analyzed in Section 5. Section 6 discusses the welfare implications of some key policy recommendations in light of our findings. A few extensions of our baseline model are considered in Section 7. The final section, Section 8, presents a conclusion. All proofs are given in Appendix A.

2 Model

We describe the model below by elaborating on its three key components: *players*, *market interactions*, and *payoffs*.

Players: A platform owner P runs a marketplace where a (representative) seller S may bring his product to reach potential customers.

Market interactions: The marketplace operates for two periods, where the time lengths of the first and second periods are $1-\delta$ and δ, respectively.[8] At the beginning of the game, the platform posts an ad valorem referral fee $r \in [0,1]$ on the seller's profit that remains fixed across the two periods. The seller incurs a fixed cost $K > 0$ to enter the marketplace, and once he enters, he earns a profit of $\theta \in [0,1]$ per unit time if he stays as a monopolist on the marketplace. Crucially, θ is privately known to the seller before he enters the marketplace, and therefore, influences his entry decision. For simplicity, we assume that the seller does not incur any variable cost, and θ is drawn from a uniform distribution over $[0,1]$.

If the seller enters then in the first period, he earns $(1-r)(1-\delta)\theta$, while the platform owner P obtains $r(1-\delta)\theta$ as referral fees. At the beginning of the second period, P observes θ (by virtue of running the marketplace with access to the seller's proprietary data) and decides whether to launch its own product by imitating the seller's item. Should P decide to imitate, it faces the same cost as the seller (i.e., only pays a fixed cost K) and earns monopoly profit $\delta\theta$ by steering consumers to its own product with self-preferencing.[9] We assume that S does not have access to any alternative marketplace, and the platform cannot develop the product on its own.

[8] We do not consider any time discounting.

[9] Our results do not crucially depend on the assumption that P and S have the same entry cost K. See Section 7 for a relevant discussion.

Payoffs: If the seller S does not enter, all players earn 0. If S enters, his (ex-post) aggregate payoff is

$$\pi_S(r,\theta) := (1-\delta)(1-r)\theta - K + (1-\mathbf{1}_E)\delta(1-r)\theta,$$

and P's (ex-post) payoff is

$$\pi_P(r,\theta) := (1-\delta) r\theta + (1-\mathbf{1}_E)\delta r\theta + \mathbf{1}_E(\delta\theta - K),$$

where the value of the indicator function $\mathbf{1}_E$ is 1 if the platform enters the market in period 2 and 0 otherwise.

Timeline: The timeline of the game is summarized below:

- *Period* 1. The platform (P) posts a referral fee $r(\geq 0)$. The seller observes his "type" θ and the referral fee r, and decides whether or not to enter. Period-one payoffs are realized.

- *Period* 2. P observes the realization of θ and decides whether to launch its own product. If P enters, it steers all consumers to its own product. Period-two payoffs are realized.

To streamline our analysis, we impose the following parametric restriction.

Assumption 1 $1-\delta < K < \delta$.

The first inequality implies that even the highest type seller ($\theta = 1$) does not enter the marketplace if he expects the platform to steal his business for sure. Notice that if the platform enters for sure then the entering seller's payoff is equal to $(1-\delta)(1-r)\theta - K \leq (1-\delta) - K$. Now, if the platform enters in the second period then its payoff is equal to $\delta\theta - K$. Thus, the second inequality ensures that in the absence of any referral fee (i.e., when $r = 0$), the platform finds it profitable to enter if θ is sufficiently large. Note that this assumption implies $\delta > 1/2$, i.e., for both parties, the future (period 2) payoff is more important than their current (period 1) payoff.

We use (weak) Perfect Bayesian Equilibrium as the solution concept.

We conclude this section with the following remarks on our modeling specifications. First, one may interpret the seller's type in terms of the demand state of his product. For example, suppose that the demand function for the seller's product is invariant over time and is given by $D(p,\theta) := \theta d(p)$ where p is the unit price and the parameter θ captures the market thickness. The seller, if selling on the marketplace, would charge the monopoly price $p^m := \arg\max_p \; pd(p)$ that is independent of the market size and the referral fee, and the same holds for the platform if it launches its own product in the second period. One obtains our modeling setup by normalizing the monopoly profit $p^m d(p^m)$ to 1. This foundation also allows us to formally measure the consumer surplus: Let $\mathbf{Pr}_S(\theta)$ denote the probability that a seller of type θ enters the marketplace. Since the optimal price p^m is independent of whether the platform enters or not, *ex-ante* consumer surplus is given by

$$\begin{aligned} CS &:= \int_0^1 \theta \left[\int \max\{d^{-1}(q) - p^m, 0\} dq \right] \mathbf{Pr}_S(\theta) d\theta \\ &= \int \max\{d^{-1}(q) - p^m, 0\} dq \cdot \int \theta \mathbf{Pr}_S(\theta) d\theta. \end{aligned}$$

In what follows, whenever applicable, we refer to this measure as consumer surplus.

Second, one may interpret the relative "lengths" of the two periods as the duration of learning and monetizing phases of a product development process. Specifically, $1 - \delta$ can be interpreted as the relative length of time it takes for the platform to acquire relevant information on the profitability of the product, and δ reflects the time span over which the platform gets to exploit this information.

Third, in our setting, ad valorem fee levied on the seller's profit is quantitatively equivalent to levying it on the price, as we have assumed the seller's marginal cost of production to be zero. A strictly positive marginal cost compromises the algebraic tractability of our model as it leads to a double marginalization problem.

Finally, it is also worth noting that even though we modeled the referral fee as an ad-valorem fee, our analysis remains qualitatively unchanged if one considers a specific fee (per transaction) instead of ad-valorem fee. A fixed ad-valorem fee reflects the practice of many large online marketplaces such as Apple and Amazon. Apple, for example, has maintained a consistent 30% fee since the inception of the App Store (with the exception of a reduced fee of 15% for sellers whose annual revenue is less than 1 million USD through its App Store Small Business Program). Amazon, on the other hand, establishes fees based on product categories, reflecting varying market demand parameters and entry costs across these categories. The fees for each category, however, have remained unchanged over time. These empirical observations suggest practical limitations on fee adjustments over time, notwithstanding the theoretical feasibility. One potential explanation is the ongoing presence of market entry, which imposes constraints on fee modifications.[10]

[10] See Madsen and Vellodi (forthcoming) for additional discussion on the justification of the stationary fee.

3 Benchmarks on platform's entry

We begin by analyzing two benchmark cases, one in which the platform never imitates the seller's product and the other in which the platform imitates the seller's profit whenever it is profitable to do so. Note that the first case can be interpreted as the case in which the platform commits to not imitate the seller's product, while the second case reflects the scenario where the platform has no commitment power over its own entry decision.

3.1 No entry by the platform

If the seller does not face any threat of entry from the platform, the platform's action affects the seller's entry only through the referral fee. Given r, the seller (S) enters if and only if

$$(1-r)\,\theta \geq K \Leftrightarrow \theta \geq \theta_S(r) := \min\left\{\frac{K}{1-r}, 1\right\}. \tag{1}$$

Therefore, the platform's problem is

$$\mathcal{P}_{NE}: \max_{r} \quad \Pi_P^{NE}(r) := \int_{\theta_S(r)}^{1} r\theta d\theta,$$

and its interior solution, denoted by r^{NE}, satisfies

$$\frac{1}{2}\left(1-\theta_S\left(r^{NE}\right)^2\right) = r^{NE}\theta_S\left(r^{NE}\right)\theta_S'\left(r^{NE}\right).$$

This is a standard monopoly pricing problem. Intuitively, an increase of r allows the platform to extract more from the seller conditional on entry (i.e., $\theta \in (\theta_S, 1]$). However, it reduces the seller's entry incentives and the platform loses the referral fee from the marginal seller type θ_S. In the above equation, the left-hand side captures the former marginal benefit, while the right-hand side represents the latter marginal cost. As the optimal fee is strictly positive, there is too little entry compared to the socially efficient level (note that it is efficient for all $\theta \geq K$ to enter).

3.2 Entry by the platform whenever profitable

Next, consider the case where the platform can imitate the seller's product and enter the market whenever profitable. Given θ, the platform prefers to enter if and only if

$$\delta\theta - K \geq \delta r\theta \Leftrightarrow \theta \geq \theta_P(r) := \min\left\{\frac{K}{\delta\,(1-r)}, 1\right\}. \tag{2}$$

So, given r, the type-θ seller's payoff from entry is:

$$\pi_S(r;\theta) = \begin{cases} (1-r)\,\theta & if\ \theta \le \theta_P(r) \\ (1-\delta)(1-r)\,\theta & otherwise. \end{cases}$$

By Assumption 1, no seller type above θ_P wishes to enter. This implies that the seller enters if and only of $\theta \in [\theta_S(r), \theta_P(r))$ where the cutoff $\theta_S(r)$ is as defined in (1). Since $\theta_S(r) \le \theta_P(r)$ for any r, the platform's problem is

$$\mathcal{P}_{NC}:\ \max_r \quad \Pi_P^{NC}(r) := \int_{\theta_S(r)}^{\theta_P(r)} r\theta d\theta.$$

Let r^{NC} be the solution to $\mathcal{P}_{NC}$. Then we have the following result.

Proposition 1 $r^{NC} \ge r^{NE}$.

In other words, the optimal referral fee set by the platform when it has no commitment power over its own entry decision is at least as large as the optimal fee it would set when it commits to not enter by imitating the seller's product.

To see the intuition for this result, recall that a lack of commitment by the platform causes the most lucrative seller types ($\theta \ge \theta_P$) not to enter. Raising r helps reduce associated losses, because it lowers the platform's own entry incentive, thereby encouraging third-party sellers' entry; that is, a higher r helps the platform mitigate its own (no-)commitment problem.[11] Specifically, as r rises, θ_P increases faster than θ_S, so the absolute length of the interval $[\theta_S(r), \theta_P(r)]$ expands (provided that $\theta_P(r) < 1$). This expansion is profitable: even if the average entering type is constant, it would increase the platform's profit. But, the expansion even raises the average entering type as it replaces lower types (around $\theta_S(r)$) with higher types (around $\theta_P(r)$). Consequently, the platform would raise r up to the point where $\theta_P(r) = 1$.

An important implication of this result is that the "no-entry" case explored earlier (in Section 3.1) Pareto dominates the "no-commitment" benchmark. The comparison of the seller's payoff is immediate from Proposition 1. When the platform cannot commit to its entry policy, in equilibrium it chooses a higher referral fee and stifles the seller's entry. The platform's payoffs in the two cases also have the same ranking. As no seller type above θ_P enters, in equilibrium the platform does not benefit from having the option to enter the marketplace: the platform never enters, and fewer types of the sellers enter the marketplace

[11] This observation is reminiscent of the "Arrow Replacement effect" (Arrow, 1962), namely, that a monopolist has a weaker incentive to innovate for fear of cannibalizing her current products.

relative to the "no-entry" case. Finally, consumes are better off because the seller is more likely to enter (i.e., $\theta_S\left(r^{NE}\right) \leq \theta_S\left(r^{NC}\right)$).[12]

4 Full commitment to entry

Section 3 highlights the importance of the platform's entry commitment. This section studies the platform's optimal strategy when it could exert full commitment power over its own entry. In particular, we assume that the platform can choose, and commit to, a function $\alpha : [0,1] \times \mathbb{R} \to [0,1]$ where $\alpha(\theta, r)$ represents the probability that it enters conditional on the seller's type being θ.

Given r and $\alpha(\cdot)$, the seller with type θ enters if, and only if,

$$[(1-\delta) + \delta(1-\alpha(\theta,r))](1-r)\theta \geq K. \tag{E}$$

This implies that the platform's problem can be written as:

$$\max_{r,\alpha(\cdot)} \int_{\Theta_E} [(1-\alpha(\theta,r))r\theta + \alpha(\theta,r)((1-\delta)r\theta + \delta\theta - K)]\,d\theta,$$

where $\Theta_E = \{\theta \in [0,1] \mid (E) \text{ is satisfied}\}$.

As before, given r, the platform has an incentive to enter if and only if $\theta \geq \theta_P(r)$.[13] So, for $\theta < \theta_P(r)$, it is optimal for the platform to set $\alpha(\theta, r) = 0$ (i.e., not to enter); in this case, the seller enters as long as $\theta \in [\theta_S(r), \theta_P(r))$. For $\theta \geq \theta_P(r)$ it is optimal for the platform to raise $\alpha(\theta, r)$ as much as possible subject to the seller's entry constraint (E). This implies that the optimal α is the value that makes (E) bind, that is,

$$\alpha(\theta; r) = \frac{1}{\delta}\left(1 - \frac{K}{(1-r)\theta}\right).$$

It is worth noting that $\theta_P(r)$ and $\theta_S(r)$ are increasing in r, whereas $\alpha(\theta; r)$ is decreasing in r and increasing in θ. Also, it is routine to check that $0 \leq \alpha(\theta; r) < 1$ (given Assumption 1). This implies that the platform's optimal commitment problem reduces to:

$$\mathcal{P}_C : \max_{r\in[0,1]} \quad \Pi_P^C(r) := \int_{\theta_S(r)}^{\theta_P(r)} r\theta d\theta + \int_{\theta_P(r)}^{1} \begin{bmatrix} (1-\delta)r\theta + \\ [\alpha(\theta;r)(\delta\theta - K) + (1-\alpha(\theta;r))\delta r\theta] \end{bmatrix} d\theta.$$

[12] Recall that in our setup, the consumer surplus (as defined in Section 2) only depends on the seller's likelihood of entry.

[13] In other words, the integrand $(1-\alpha)r\theta + \alpha((1-\delta)r\theta + \delta\theta - K)$ is increasing α if and only if $\theta \geq \theta_P$. Note that this implies that the integral is maximized when $\alpha = 0$ for $\theta < \theta_P$ and α is maximized for $\theta \geq \theta_P$.

Let r^C be the solution to $\mathcal{P}_C$. Notice that

$$\Pi_P^C(r) - \Pi_P^{NE}(r) = \int_{\theta_P(r)}^{1} \alpha\left(\theta; r\right)\left[\delta\left(1-r\right)\theta - K\right] d\theta \quad \begin{cases} > 0 & if\ r < \overline{r} \\ = 0 & otherwise, \end{cases}$$

where $\theta_P\left(\overline{r}\right) = 1$. This observation leads to the following result.

Proposition 2 $r^C \leq r^{NE}$.

The result above states that the optimal referral fee set by the platform when it has full commitment power over its own entry decision is no larger than the optimal fee it would set when it commits to not enter the marketplace.

In the no-entry benchmark case, the seller obtains positive rents whenever $\theta \geq \theta_S(r)$, whereas in the commitment case, the seller receives no rents whenever $\theta \geq \theta_P(r)$. This implies that the marginal gain of increasing r is higher under the no-entry case (extracting more from seller types in $[\theta_S(r), 1]$) than under the case where the platform can commit to its entry decision (extracting more from the seller types in $[\theta_S(r), \theta_P(r))$ only). The result that $r^C \leq r^{NE}$ follows, because the corresponding marginal cost of increasing r—$\theta_S(r)$ type not entering—is identical between the two cases. Notice that under full commitment, the seller's entry incentives are stronger as the referral fee is lower. Also the threat of business stealing does not thwart the seller's entry as, in equilibrium, the platform's likelihood of entry is set at the level where the seller breaks even by deciding to enter.

Compared to the two benchmark cases (in Section 3), with commitment to entry the platform and the consumers are both better off. It is trivial that commitment to entry would benefit the platform as it can always implement the market equilibrium under the "no entry" or "no commitment" case. The consumers are better off as there is more entry due to lowered referral fee (and they pay the same price regardless of platform's subsequent entry decision). However, the seller's payoff vis-a-vis the benchmark cases cannot be ranked as it is affected by two countervailing effects: a lower referral fee induces more entry by the seller (θ_S is lower) but it also makes the platform's entry more likely. Interestingly, this observation opens up the possibility that, relative to the case when the platform cannot enter, some types of the seller may be better off when the platform can commit to its entry decision. In particular, the seller types $\theta \in [\theta_S(r^C), \min\{\theta_S(r^{NE}), \theta_P(r^C)\}]$ now earn a positive rent following entry whereas they would have earned 0 in the scenario where the platform cannot enter the marketplace (and chooses its referral fee r^{NE} accordingly).

5 Optimal information policy

We now turn to the situation in which the platform cannot fully commit to its entry decision, but can optimally choose its "information policy", i.e., how much information about the profitability of the seller's product it may generate or share with its product division.[14] To evaluate the full potential of the platform's information policy, as in the recent literature on information design, we endow the platform with full flexibility in its choice of information. The following result, however, shows that in our environment, it suffices to restrict attention to binary signals.

Lemma 1 *Consider any set X of signal realizations and a (measurable) signal $\sigma : [\theta_S, 1] \to \Delta(X)$, where $\Delta(X)$ denotes the set of all probability distributions over X. There exists a binary signal $\hat{\sigma} : [\theta_S, 1] \to \Delta(\{0, 1\})$ that implements the same equilibrium outcome as the signal σ.*

To understand this result, notice that the platform's payoff is linear in θ, so its entry decision depends only on the conditional expectation of θ—in particular, whether it exceeds θ_P or not. All signal realizations leading to the posterior mean below (or above) θ_P can be pooled. Therefore, the outcome by any signal can be replicated by a binary signal, whose realization can be interpreted as an action (entry or not) recommended to the platform.

5.1 Implementing full commitment outcome

We first explore if and when it may be possible for the platform to implement the full commitment outcome characterized in Section 4.

To implement the full commitment outcome, it is necessary and sufficient that the platform enters with probability $\alpha(\theta, r)$ when the seller's type is θ. With a binary signal, this outcome can be induced if and only if the platform uses the following signal structure:

$$\Pr(x = 1 \mid \theta) = \begin{cases} \alpha(\theta, r) & if\ \theta \geq \theta_P(r) \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

In other words, the platform should observe $x = 1$, and so enters, with probability $\alpha(\theta, r)$ whenever $\theta \geq \theta_P(r)$; otherwise, the platform should observe $x = 0$ and not enter. By construction, this signal induces the full commitment outcome in Section 4 if and only if the

[14] A common regulatory policy proposal of completely banning internal information sharing is analyzed later in Section 6.

following two obedience constraints hold:

$$\mathbb{E}[\theta \mid x = 0;\ \theta \geq \theta_S(r)] \leq \theta_P(r), \tag{OC_0}$$

and

$$\mathbb{E}[\theta \mid x = 1;\ \theta \geq \theta_S(r)] > \theta_P(r). \tag{OC_1}$$

That is, conditional on $x = 0$, the expectation of θ should be less than $\theta_P(r)$ so that the platform has no incentive to enter. On the contrary, conditional on $x = 1$, the expectation of θ should exceed $\theta_P(r)$ so that the platform wishes to enter.

Since $\alpha(\theta, r) > 0$ (and so one may observe $x = 1$) only when $\theta > \theta_P(r)$, the constraint (OC_1) is always satisfied under the signal structure (3). This implies the following result.

Proposition 3 *The full commitment outcome is implementable by an information policy if and only if the constraint (OC_0) holds for $r = r^C$, which is equivalent to:*

$$\int_{\theta_P(r^C)}^{1} (\theta - \theta_P\left(r^C\right))(1 - \alpha(\theta, r^C))d\theta \leq \int_{\theta_S(r^C)}^{\theta_P\left(r^C\right)} (\theta_P\left(r^C\right) - \theta)d\theta. \tag{4}$$

Figure 1 shows the parameter region under which (OC_0) holds—so, the commitment outcome can be implemented by an information policy. If δ is sufficiently small, full commitment outcome is always implementable for any K (satisfying Assumption 1), whereas for large δ, K should exceed a threshold (that itself varies with δ).

In order to understand the pattern, it is useful to interpret the right-hand side of (4) as the platform's "budget for (no entry) obedience" while the left-hand side as the corresponding "spending". Recall that for a given r, θ_P is increasing in K and decreasing in δ, whereas α is decreasing in both K and δ. If K is large and δ is relatively small then $\theta_P(r)$ is close to 1 and α remains moderate. In this case, the left-hand side ("spending") is small, so (OC_0) holds. Similarly, if δ is small and close to $1/2$ then $\theta_P(r) - \theta_S(r)$ is large and so is α. Consequently the right-hand side ("budget") of (OC_0) becomes large whereas the left-hand side ("spending") remains moderate, and (OC_0) continues to hold. But (OC_0) fails when K is relatively small and δ is relatively large. The left-hand side of (OC_0) becomes large (as θ_P and α are both relatively small) and the right-hand side becomes small (as $\theta_P(r) - \theta_S(r)$ is small). As a result, (OC_0) is violated.

5.2 Optimal signal when full commitment outcome is infeasible

What is the optimal information policy for the platform when the the full commitment outcome cannot be implemented, i.e., what if (4) fails? We explore this case in two steps.

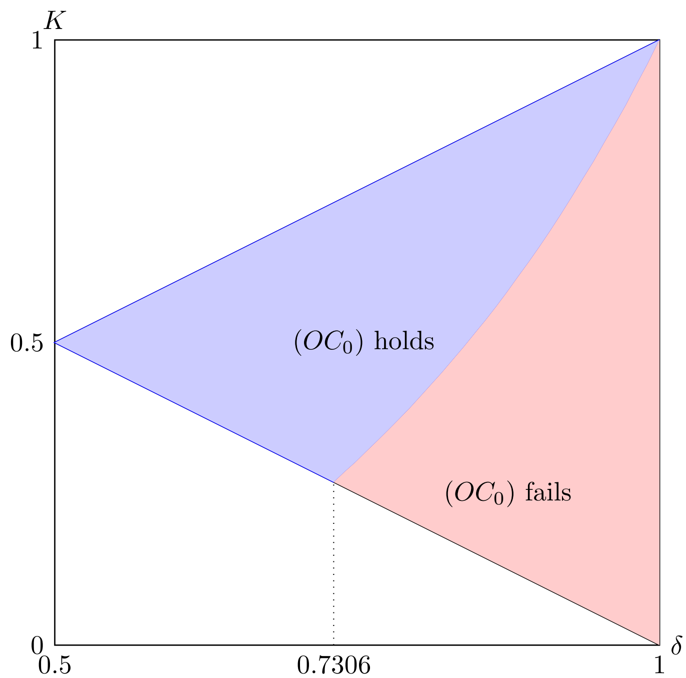

Figure 1: Set of K and δ where full commitment outcome can be implemented via information design (i.e., (OC_0) holds under $r = r^C$).

First, we study the optimal information policy for a given referral fee, and then explore the profit maximizing fee for the platform.

Fix a value of r. For any seller type θ who enters, let $\gamma(r, \theta)$ denote the probability that the platform receives $x = 1$, and therefore also enters, conditional on θ. For the optimal signal, we must have $\gamma = \alpha$ for all $\theta > \theta_P(r)$; if $\gamma(r, \theta) > \alpha(r, \theta)$, type θ would not enter, and if $\gamma(r, \theta) < \alpha(r, \theta)$, raising γ would relax (OC_0) and yield a higher payoff for the platform. This implies that whenever (OC_0) is violated under the signal structure (3), the optimal signal should dissuade entry for some seller types above θ_P. The following result shows that it is optimal for the platform to deter entry by intermediate seller types.

Proposition 4 *Given a referral fee r such that (OC_0) fails, the optimal binary signal is given as follows: there exist θ_* $(> \theta_P)$ and $\theta^* \in [\theta_*, 1]$ such that*

$$\Pr(x = 1 \mid \theta) = \begin{cases} \alpha(\theta, r) & \text{if } \theta \in [\theta_P(r), \theta_*) \cup [\theta^*, 1] \\ 1 & \text{if } \theta \in [\theta_*, \theta^*) \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

Facing this signal, the seller enters if and only if $\theta \in [\theta_S, \theta_*] \cup [\theta^*, 1]$.

To understand this result, let $\phi : [\theta_P, 1] \to \{0, 1\}$ denote the entry decision of the seller with type above θ_P induced by the platform's information policy. Given this, (OC_0) can be written as

$$\int_{\theta_P}^{1} \phi(\theta)(\theta - \theta_P)(1 - \alpha(\theta, r))d\theta \leq \int_{\theta_S}^{\theta_P} (\theta_P - \theta)d\theta.$$

Notice that this inequality necessarily holds if $\phi(\theta) = 0$ for all $\theta \geq \theta_P$ and fails if $\phi(\theta) = 1$ for all $\theta \geq \theta_P$. The platform should find the set of seller types for which $\phi(\theta) = 0$, and the platform's decision depends on how much a seller type would contribute to its profit as well as the implicit cost of including this type via (OC_0). Proposition 4 shows that the platform should exclude intermediate types $[\theta_*, \theta^*)$, that is, induce entry for seller types below θ_* or above θ^*. The seller types that are just above θ_P cost little in terms of (OC_0)—$(\theta - \theta_P)(1 - \alpha(\theta, r))$ is close to 0—but yield non-negligible profits to the platform. And the seller types that are sufficiently high (i.e., above θ^*), though costly to include, are particularly lucrative to the platform.

Next, consider the optimal referral fee r^I (say) when the platform uses the associated optimal information policy. The platform's optimal referral fee solves (recall that θ_S, θ_P, and α also depend on r):

$$\mathcal{P}_I : \max_{r\in[0,1]} \quad \Pi_P^I(r) := \int_{\theta_S}^{\theta_P} r\theta d\theta + \int_{[\theta_P, \theta_*]\cup[\theta^*, 1]} \left[(1 - \alpha(\theta))r\theta + \alpha(\theta)\left((1 - \delta)r\theta + \delta\theta - K\right)\right] d\theta.$$

Let r^I be the solution to $\mathcal{P}_I$. Notice that

$$\Pi_P^I(r) = \Pi_P^C(r) - \int_{\theta_*(r)}^{\theta^*(r)} \left[r\theta + \alpha(\theta; r)\left[(1 - r)\,\delta\theta - K\right]\right] d\theta. \tag{6}$$

As noted earlier, if (OC_0) does not bind at r^C, the optimal fee is the same as that under the full commitment case, i.e., $r^I = r^C$. In this case, $\theta_*(r^C) = \theta^*(r^C) = 1$ and the platform implements the full commitment outcome. Otherwise, if the platform sets $r = r^C$, in order to meet the (OC_0) constraint it must exclude entry for some of the intermediate types of the seller (i.e., for all types $\theta \in [\theta_*(r), \theta^*(r)]$). Thus, the optimal fee, r^I, may differ from r^C. As (6) indicates, r^I balances the trade-off between the loss relative to the full-commitment payoff and gains from relaxing (OC_0) that may allow for more entry from the intermediate types.

It is reasonable to conjecture that $r^I > r^C$. Raising r from r^C would have a second-order effect on Π_P^C (the first term in the expression of Π_P^I in (6)) but it increases θ_P and may relax (OC_0), consequently, having a first-order effect on the gains due to more entry of the

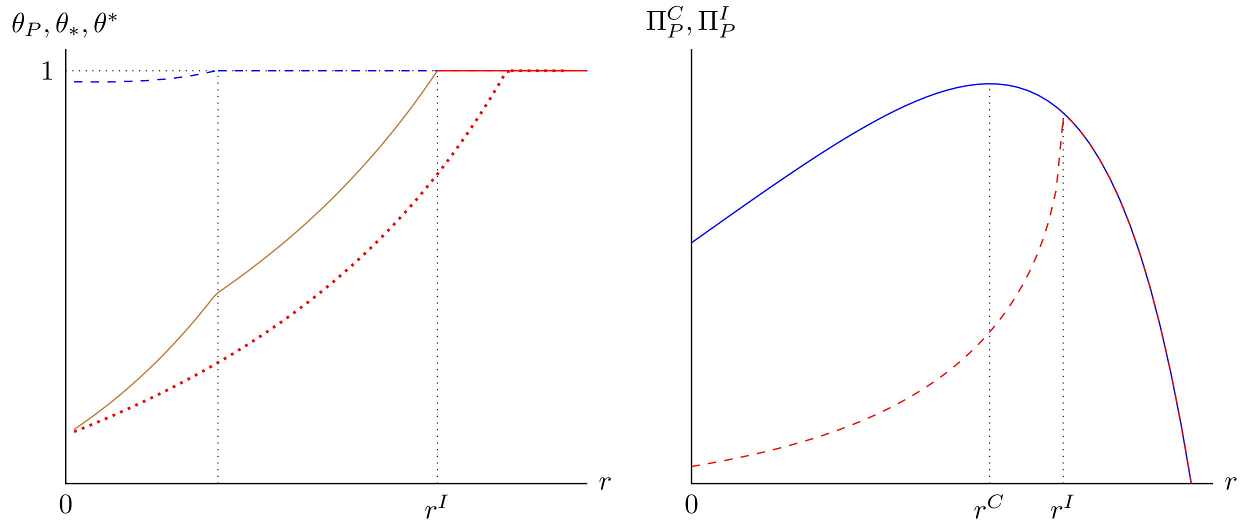

Figure 2: The left panel shows how θ_* (brown solid), θ^* (blue dashed), and θ_P (red dotted) depend on r, while the right panel compares P's commitment profit $\Pi_P^C(r)$ (blue solid) to its profit $\Pi_P^I(r)$ from the optimal information policy (red dashed). In both panels, $K = 0.4$ and $\delta = 0.85$.

intermediate types of the sellers (captured by a decrease in the value of the second term in (6)). Unfortunately, the comparative statics for the cutoffs $\theta_*(r)$ and $\theta^*(r)$ appear to be analytically intractable, and hence, a formal proof of this conjecture remains elusive. But numerical solution to the platform's program ($\mathcal{P}_I$) conforms to our conjecture.[15]

Increasing r tend to foster more entry as the set of excluded types $[\theta_*(r), \theta^*(r)]$ gets smaller with r. Moreover, the optimal fee r^I is the smallest value of r such that the obedience constraint (OC_0) does not bind. That is, at the optimum, the seller's entry decision continues to follow the cutoff rule where all types $\theta \geq \theta_S$ enter. While there is no exclusion of the intermediate types, the higher fee does restrict entry of the low types by raising the entry cutoff θ_S.

6 Implications for regulatory measures

In light of our discussion above we can now explore the welfare implications of two salient policy proposals for regulation of platform-run marketplaces: prohibition on imitating the third-party sellers' products, and prohibition on the use of proprietary data about sellers' products.

The implications for the former policy—ban on imitation—can be readily obtained from

[15]For a simpler case of binary types of sellers, one could indeed show that $r^I \geq r^C$ always holds. See Appendix B.

our analysis of the case where the platform operates only as a marketplace. The welfare implication of such a policy depends on the relative ranking of r^I and r^{NE}—the equilibrium referral fees when there are no regulations and when the platform commits not to enter. Clearly, when the full commitment case can be implemented through information policy (i.e., we have $r^I = r^C < r^{NE}$), prohibition on the platform's entry is welfare reducing for the consumers. But otherwise, r^I and r^{NE} cannot be ordered a priori, and when $r^I > r^{NE}$, a ban on imitation improves welfare.

The implications for the latter policy—ban on the use of (entry) decision-relevant information from the third-party sellers—is more subtle. Our model implies that the welfare implications of such a policy are again ambiguous, and may in fact lead to a welfare loss.

If information usage is ruled out by regulators but there is no restriction on imitation, the platform's entry decision would be only based on the seller's entry threshold. In what follows, we first characterize the platform's entry decision for a given referral fee. Let $\alpha \in [0,1]$ denote the probability that the platform launches its own product by imitating the seller. Given (r, α), a type-θ seller enters if and only if

$$(1-\alpha\delta)(1-r)\,\theta \geq K \Leftrightarrow \theta \geq \theta_S(r,\alpha) := \frac{K}{(1-\alpha\delta)(1-r)}.$$

As discussed earlier, for $r \geq \overline{r} := 1 - \frac{K}{\delta}$, $\theta_P(r) = 1$, i.e., the platform never enters. Therefore, for any such $r \geq \overline{r}$, the seller's entry threshold is

$$\hat{\theta}_S(r) = \min\{\theta_S(r,0), 1\},$$

and the principal's profit is given by

$$\Pi_P^{NI}(r) := \int_{\hat{\theta}_S(r)}^{1} r\theta d\theta = r\mathbb{E}\left[\theta \mid \theta \geq \hat{\theta}_S(r)\right].$$

For $r < \overline{r}$ (equivalently, $\theta_P(r) < 1$), there are multiple equilibria. One can construct a "futile" equilibrium where the seller never enters. In this equilibrium, given $r\,(<\overline{r})$ the seller believes that the platform will enter with probability 1 (i.e., $\alpha = 1$). If so, the seller would not enter regardless of the his type. And this outcome can be supported as a perfect Bayesian Equilibrium with the platform's belief that assigns probability 1 to $\theta = 1$ as the entering seller's type.

Now we consider the case in which (in equilibrium) the platform does not enter (i.e.,

$\alpha = 0$). This equilibrium exists if and only if

$$\mathbb{E}\left[\theta \mid \theta \geq \hat{\theta}_S(r)\right] \leq \theta_P(r) \Leftrightarrow \int_{\hat{\theta}_S(r)}^{1} (\theta - \theta_P(r))\, d\theta \leq 0. \tag{7}$$

This inequality reduces to

$$\frac{1}{2}(1-\theta_P)^2 - \frac{1}{2}\left(\hat{\theta}_S - \theta_P\right)^2 \leq 0 \Leftrightarrow r \geq \underline{r} := 1 - \frac{2-\delta}{\delta}K. \tag{8}$$

It is easy to see that $\underline{r} < \overline{r}$. Also note that

$$\underline{r} \geq 0 \Leftrightarrow K \geq \frac{\delta}{2-\delta}.$$

As $K \in (1-\delta, \delta)$, we therefore have the following two cases: if $K \in \left(\frac{\delta}{2-\delta}, \delta\right)$ then $\underline{r} \leq 0$, so this equilibrium exists for all $r < \overline{r}$.[16] Otherwise (i.e., $K \in \left(1-\delta, \frac{\delta}{2-\delta}\right)$), this equilibrium does not exist for $r < \underline{r}$.

This observation implies that if $r < \underline{r}$, then the above futile equilibrium is the only equilibrium. Therefore, with a prohibition on information sharing, either the market ceases to exist or the platform never enters. Consequently, the welfare implication of such a policy cannot be ascertained a priori. Even in the equilibrium where the seller enters, the optimal referral fee would be r^{NE} (as the platform does not enter), and as we have already argued, in general, r^{NE} and r^I cannot be ordered. In particular, if $r^I < r^{NE}$, prohibition on information sharing would reduce consumer welfare. It is also worth noting that when $r^{NE} \geq \underline{r}$, a ban on imitation and a ban on third-party data usage—the two seemingly different policies targeting different aspects of the platform's behavior—lead to the same equilibrium outcome, and hence, have exactly the same welfare consequences.

It is worth noting that apart from prohibition on imitation and data usage, the regulators may also require a platform to share its proprietary information with the third-party sellers. For instance, Amazon offered to share marketplace data with third-party sellers to settle EU antitrust investigations on its business practices of possible preferential treatment of its own retail offers.[17] This policy may have additional effects of limiting the market power of third-party sellers. However, third-party sellers would be averse to such information being

[16] One may wonder whether there exists an equilibrium where the platform enters with probability $\alpha \in (0,1)$. The answer is negative. For such a mixed strategy equilibrium to exist, the platform should be indifferent between entering and not entering, that is, $E[\theta \mid \theta \geq \theta_S(r,\alpha)] = \theta_P(r)$. But this equality cannot hold whenever $r < \underline{r}$. From (7) and (8) it follows that for any such r we would have $\mathbb{E}[\theta|\theta \geq \theta_S(r,0)] > \theta_P(r)$ and, since $\theta_S(r,\alpha)$ is increasing in α, for all $\alpha > 0$, we have $\mathbb{E}[\theta|\theta \geq \theta_S(r,\alpha)] > \mathbb{E}[\theta|\theta \geq \theta_S(r,0)] > \theta_P(r)$.

[17] https://www.reuters.com/technology/amazon-offers-share-data-boost-rivals-dodge-eu-antitrust-fines-sources-2022-06-13/

disclosed to potential rival firms beyond the platform itself because it can raise serious issues with firms' privacy and confidential business plans.[18]

7 Discussion

We made several simplifying assumptions in our model to improve its analytical tractability. However, some stylized aspects of our model are not essential for our findings, and the interplay between referral fee and information policy that we highlight in our paper would continue to hold even if we relax some of these assumptions. In this section, we discuss three such extensions of our model: (i) different entry costs for the seller and the platform, (ii) competition between seller and platform following the platform's entry in the second period, and (iii) alternative specifications for the prior distribution of the profitability of the seller's product.

Asymmetric entry costs. Our model assumes that the cost of entry is the same for both the seller and the platform. However, it is conceivable that they might face different costs to set up their operations, and the two parties may not have the same access to necessary production and distribution facilities. Our model and analysis can be readily adjusted to explore the case where the platform's entry cost, denoted by K_P, is different from the seller's entry cost of K.

In the spirit of Assumption 1, we may continue to assume that $1 - \delta < K$ (i.e., the first-period profit is not sufficient to induce the seller's entry) and $K_P < \delta$ (i.e., in the absence of any referral fee, the platform would prefer to enter if θ is sufficiently high). With varying entry costs, the profitability thresholds for entry by the seller and the platform, θ_S and θ_P, are given as:

$$\theta_S(r) = \min\left\{\frac{K}{1-r}, 1\right\} \text{ and } \theta_P(r) = \min\left\{\frac{K_P}{\delta\,(1-r)}, 1\right\}.$$

Now, there are two cases to consider: (i) $K_P/\delta > K$, and (ii) $K_P/\delta \leq K$.

Under case (i), $\theta_S(r) < \theta_P(r)$, and all our results remain qualitatively the same. Propositions 1 and 2 continue to hold as their proofs rely on the ranking of $\theta'_S(r)$ and $\theta'_P(r)$ (in the case of Proposition 1) and that of $\Pi_P^{C\prime}(r)$ and $\Pi_P^{NE\prime}(r)$ (in the case of Proposition 2), both of which continue to hold as long as $K_P/\delta > K$. The characterization of the optimal information policy also remains unaltered as (OC_0) can be met by using the same information policy as given in Proposition 4.

[18] This may be the reason why such information sharing was not included in Amazon's commitments which were accepted by the European Commission on December 20, 2022. (https://ec.europa.eu/ commission/presscorner/detail/en/ip_22_7777)

Under case (ii), however, $\theta_S(r) = \theta_P(r)$. That is, the platform would enter whenever it is profitable for the seller to enter, and therefore, when the platform cannot commit to entry, the seller would not enter in the first place. This observation also implies that a platform lacking commitment power over its entry decision cannot use its information policy as a substitute—(OC_0) would necessarily fail as $\mathbb{E}[x = 0 \mid \theta > \theta_S(r)] > \theta_S(r) = \theta_P(r)$.

Post-"Sherlocking" competition. In the main model, we have assumed that once the platform enters, it monopolizes the market with the third-party seller being deprived of opportunities to make any further sales. This is equivalent to an extreme form of self-preferencing by the platform.[19] We can easily accommodate the possibility of ex post competition between the platform owner and third parties with a reinterpretation of the platform's entry probability as the extent to which the platform's own products are featured with prominence (while the third-party products are featured with the complementary probability). With this interpretation, the design of the (product recommendation) algorithm plays the same role as information policy in our model in terms of providing entry incentives for third-party sellers.

However, there is one major difference between the two cases. With a self-preferencing algorithm that shares the market with the third-party seller, the platform needs to enter every market for which $\alpha(\theta, r) > 0$. In contrast, with information design, the platform enters only when its signal realization is 1 with probability $\alpha(\theta, r) > 0$. Thus, an information design policy is more efficient than the use of algorithms in terms of saving duplicative entry costs. The regulatory scrutiny on self-preferencing may induce the platform to shy away from the extreme form of self-preferencing and compel it to share the market with third parties when it enters. However, one may argue that such an intervention may entail inefficiency in terms of entry cost duplication.

Note that in our model we have taken the extent of consumer participation in the platform as fixed to focus on the effect of "Sherlocking" on the entry incentives for third-party sellers. However, one benefit from platform entry may be to discipline third-party sellers' pricing decisions, that, in turn, benefits the consumers. And the platform may care about such consumer benefit if it competes with other platforms or when its objective is to grow its

[19] As a channel of self-preferencing, we can imagine a situation in which consumers make purchase decisions based on the platform's recommendations. For instance, Amazon's Buy Box features a default seller and according to one estimate (Juul, 2020), over 80% of Amazon sales take place through the Buy Box. In addition, academic research shows that the Amazon platform substantially prioritizes its own private label products and third party offers shipped by its fulfillment arm, in the algorithm that determines which products are featured in its Buy Box (Devesh, 2022). Farronato et al. (2023) also find that amazon-branded products are ranked higher than observably similar products in consumer research results. For a theoretical treatment of self-preferencing and the effect of regulations on search neutrality, see Zou and Zhou (2023).

consumer base over time. We abstract from this aspect of the post entry competition which can be important in the initial phase of marketplace to get consumers on board.[20]

Delegated entry decision. In our analysis, we have assumed that the platform's referral fees and entry decisions are both set by its owner. Thus, while deciding on its entry or information policy, the platform weighs its profits from entry against the potential loss of referral fees. In contrast, one may assume that the platform owner delegates the entry decision to its product division manager (the manager could have better information about the platform's production capabilities) but can control the flow of information between the marketplace and the product division. That is, the product division may not readily observe the profitability of various products that are sold on the marketplace, but receives signals on the same as per the platform's information policy. Upon receiving such signals, the division decides whether to enter the marketplace by imitating a seller's product. As the product division does not internalize the opportunity cost of entry in terms of the foregone referral fees, it may enter as long as the expected profit from entry (conditional on the received signal) exceeds its entry cost. This observation has two major implications for our results.

First, when the platform cannot commit to entry and the seller's type (θ) becomes public following the seller's entry (as is the case studied in Section 3.2), the platform would enter if and only if $\delta\theta \geq K$, i.e., the platform's entry threshold becomes $\theta_P = K/\delta$ (that is independent of r), and so the platform's optimal referral fee solves:

$$\Pi_P^{NC}(r) := \max_r \int_{\theta_S(r)}^{\theta_P} r\theta d\theta. \tag{9}$$

Now, Proposition 1 no longer holds, i.e., we have $r^{NC} \leq r^{NE}$: the platform's optimal referral fee when it cannot commit to its entry policy would be lower than its optimal fee when it commits not to enter. The argument is simple. When the platform owner decides on both the referral fee and product entry but lacks commitment power over her entry decision, a high referral fee (r) serves as a commitment device. It assures the seller that the platform is less likely to enter as it has a lot to lose in terms of the foregone referral fees. But when the entry decision is delegated at the division level, the product division does not internalize the loss of referral fees and therefore, a higher value of r no longer deters the platform's entry. One may also readily see this from the first-order condition of the platform's problem (9):

[20]For an analysis of the platform's incentives to promote competition and limit the sellers' market power (as opposed to fully "expropriating" them), see Teh (2022) and Johnson *et al.* (2023).

If r^{NE} is an interior solution, then

$$\frac{d}{dr}\Pi_P^{NC}(r^{NE}) = \int_{\theta_S(r^{NE})}^{\theta_P} \theta d\theta - \theta_S'(r^{NE})r^{NE}\theta_S(r^{NE})$$
$$< \int_{\theta_S(r^{NE})}^{1} \theta d\theta - \theta_S'(r^{NE})r^{NE}\theta_S(r^{NE}) = 0.$$

Second, consider the issue of optimal information policy. One may again ask when the platform could implement its full commitment outcome through a suitably chosen information policy. To address this issue, one may follow the exact steps discussed in Section 5.1, and if entry occurs (by the product division) whenever it believes θ (conditional on the signal it receives) to exceed its entry threshold K/δ, then the only change to make is to replace $\theta_P(r^C)$ with $\theta_P = K/\delta$. Since $\theta_P < \theta_P(r)$ for any $r > 0$, condition (4) becomes harder to satisfy. This is intuitive because the product division manager's incentives are misaligned with the platform's, so the platform will do necessarily (weakly) worse than in our main model.

General distributions for θ. While our model assumes the uniform distribution for the profitability parameter θ, this assumption is not essential for our findings. In particular, if we consider a general distribution F for θ, both Propositions 1 and 2 as well as our characterization of the optimal information policy continue to hold; however, for Proposition 1 to hold we need the associated distribution function f to "not decrease too fast."[21] The comparison between the optimal fee under commitment and information policy (i.e., between r^C and r^I), of course, would remain analytically intractable as it remains so even under the uniform prior.

[21] A sufficient condition is $f(\theta_P(r))/f(\theta_S(r)) \geq \delta^2$ for any r. To see this, notice that r^{NE} maximizes $\int_{\theta_S(r)}^1 r\theta dF(\theta)$, so if it is an interior solution then it satisfies

$$\Pi_P^{NE}(r^{NE}) = \int_{\theta_S(r^{NE})}^{1} \theta dF(\theta) - r^{NE}\theta_S'(r^{NE})\theta_S(r^{NE})f(\theta_S(r^{NE})) = 0.$$

The result follows from the fact that, with no commitment to entry (as considered in Section 3.2),

$$\begin{aligned}\Pi_P^{NC}(r^{NE}) &= \int_{\theta_S(r^{NE})}^{\theta_P(r^{NE})} \theta dF(\theta) + r^{NE}\theta_P'(r^{NE})\theta_P(r^{NE})f(\theta_P(r^{NE})) - r^{NE}\theta_S'(r^{NE})\theta_S(r^{NE})f(\theta_S(r^{NE}))\\ &= \frac{\theta_P'(r^{NE})\theta_P(r^{NE})f(\theta_P(r^{NE}))}{\theta_S'(r^{NE})\theta_S(r^{NE})f(\theta_S(r^{NE}))}\int_{\theta_S(r^{NE})}^{1} \theta dF(\theta) - \int_{\theta_P(r^{NE})}^{1} \theta dF(\theta)\\ &\geq \left(\frac{1}{\delta^2}\frac{f(\theta_P(r^{NE}))}{f(\theta_S(r^{NE}))} - 1\right)\int_{\theta_S(r^{NE})}^{1} \theta dF(\theta).\end{aligned}$$

8 Conclusion

The use and misuse of proprietary data on third-party sellers by hybrid platforms have drawn considerable attention from the regulatory agencies. Leading platform-run marketplaces, such as Apple App Store and Amazon Marketplace, are alleged to use third-party sellers' product data to target and copy successful products forcing the third-party sellers to exit. Such a practice by the platforms, often termed as " Sherlocking," may stifle innovation and entry by third-parties. The anticompetitive implications of this practice have led several regulatory agencies in the U.S. and the European Union to propose limits on the platform's behavior, including ban on imitation and ban on data sharing between the platform's marketplace and product division.

However, the policy on third-party sellers' data usage is not the only strategic lever that a platform can use. The platform also chooses the referral fee for the sellers using their marketplace and decides on its own entry policy accordingly. Thus, in order to assess the platform's response to a regulatory limit on its data policy, one must analyze the interplay between its fee structure, data policy, and entry decision. In this article, we present a stylized model of hybrid platform to explore this issue. Our findings indicate that for a broad range of parameters, the platform's commitment to data policy can be a substitute for commitment to entry. And an outright ban either on entry by imitation or data sharing, in general, may be welfare reducing as the platform adjusts its referral fee in response to such policies.

Appendix

A Omitted Proofs

Proof of Proposition 1. Notice that

$$\Pi_P^{NE}(r) - \Pi_P^{NC}(r) = \int_{\theta_P(r)}^{1} r\theta d\theta \quad \begin{cases} > 0 & if\ r < \overline{r} \\ = 0 & otherwise \end{cases},$$

where $\overline{r}$ denote the value such that $\theta_P(\overline{r}) = 1$, that is, $\overline{r} = 1 - K/\delta$. It is immediate that $r^{NC} = r^{NE}$ if $r^{NE} \geq \overline{r}$.

Thus, it suffices to show that $r^{NC} \geq \overline{r}$: if $r^{NE} \geq \overline{r}$ then $r^{NE} = r^{NC}$, while if $r^{NE} < \overline{r}$ then $r^{NE} \leq r^{NC}$. For any $r < \overline{r}$, we have $\theta_S(r) < \theta_P(r)$ and

$$\theta_S'(r) = \frac{K}{(1-r)^2} < \theta_P'(r) = \frac{K}{\delta(1-r)^2} = \frac{1}{\delta}\theta_S'(r).$$

Therefore,

$$\frac{d}{dr}\Pi_P^{NC}(r) = \int_{\theta_S(r)}^{\theta_P(r)} \theta d\theta + r\left[\theta_P'(r)\theta_P(r) - \theta_S'(r)\theta_S(r)\right] > 0.$$

This means that $\Pi_P^{NC}(r)$ is strictly increasing whenever $r < \overline{r}$, so $r^{NC} \geq \overline{r}$. ■

Proof of Proposition 2. Since the result is immediate if $r^C < \overline{r} \leq r^{NE}$, it suffices to consider the case where r^C, $r^{NE} < \overline{r}$. At any $r < \overline{r}$, we have

$$\frac{d}{dr}\left(\Pi_P^C(r) - \Pi_P^{NE}(r)\right) = \frac{d}{dr}\int_{\theta_P(r)}^{1} \alpha(\theta;r)\left[\delta(1-r)\theta - K\right] d\theta < 0,$$

where the inequality holds because both $\alpha(\theta;r)$ and $\delta(1-r)\theta - K$ are positive and decreasing in r. In other words, $\Pi_P^C(r)$ is strictly decreasing whenever $\Pi_P^{NE}(r)$ is decreasing. Assuming that $\Pi_P^C(r)$ is quasi-concave, this implies that $r^C < r^{NE}$. ■

Proof of Lemma 1. For each x, let $\beta(x)$ denote the probability that the principal enters. Clearly, (since the principal cannot commit to its entry policy)

$$\beta(x)\begin{cases} = 0 & \text{if } \mathbb{E}[\theta|x] < \theta_P \\ \in [0,1] & \text{if } \mathbb{E}[\theta|x] = \theta_P \\ = 1 & \text{if } \mathbb{E}[\theta|x] > \theta_P. \end{cases}$$

Notice that the probability that the principal enters conditional on θ is given by

$$\widehat{\beta}(\theta) = \int \beta(x) d\sigma(\theta).$$

Consider the following binary signal: $\widehat{X} = \{0,1\}$ and for each θ, $\sigma(\theta)$ assigns probability $1 - \widehat{\beta}(\theta)$ to 0 and probability $\widehat{\beta}(\theta)$. By construction, this signal induces the same outcome, provided that the principal's obedience constraint is satisfied (i.e., $\mathbb{E}[\theta|0] \leq \theta_P$ and $\mathbb{E}[\theta|1] \geq \theta_P$). To see that this last requirement automatically holds from our construction, divide the set X as follows:

$$\begin{aligned} X_< &:= \{x \in X : \mathbb{E}[\theta|x] < \theta_P\}, \\ X_= &:= \{x \in X : \mathbb{E}[\theta|x] = \theta_P\}, \\ \text{and } X_> &:= \{x \in X : \mathbb{E}[\theta|x] > \theta_P.\} \end{aligned}$$

By our construction, each $x \in X_<$ is mapped to 0 in our binary signal, each $x \in X_=$ is split between 0 and 1, and each $x \in X_>$ is mapped to 1. This implies that

$$\mathbb{E}[\theta|0] \leq \mathbb{E}[\theta|x \in X_< \cup X_=] \leq \theta_P \text{ and } \mathbb{E}[\theta|0] \geq \mathbb{E}[\theta|x \in X_= \cup X_>] \geq \theta_P.$$

■

Proof of Proposition 4. The proof is given by the following steps.

Step 1. Let $\phi : [0,1] \to \{0,1\}$ be the entry decision of the seller ($\phi = 1$ if the seller enters) induced by the information structure (and r). As argued in Section 5.2, we have $\Pr(x = 1 \mid \theta) = \alpha(r, \theta)$ for all θ such that $\phi(\theta) = 1$. In addition, the platform prefers the seller to enter whenever $\theta \in [\theta_S(r), \theta_P(r)]$, as it not only yields direct benefits to the platform, but also relaxes (OC_0) by raising the right-hand side in (4). This implies that the platform's problem can be written as[22]:

$$\max_{\phi(\cdot)\in\{0,1\}} \int_{\theta_P}^{1} \phi(\theta) \left[(1-\alpha(\theta,r))r\delta\theta + \alpha(\theta)(\delta\theta - K)\right] d\theta$$

$$s.t. \quad \int_{\theta_P}^{1} \phi(\theta)(\theta-\theta_P)(1-\alpha(\theta,r))d\theta \leq \int_{\theta_S}^{\theta_P} (\theta_P - \theta)d\theta. \tag{10}$$

Step 2. Consider the associated Lagrangian:

$$\mathcal{L} = \int_{\theta_P}^{1} \phi(\theta) \left[(1-\alpha(\theta))r\delta\theta + \alpha(\theta)(\delta\theta - K) - \lambda(\theta-\theta_P)(1-\alpha(\theta))\right] d\theta + \lambda \int_{\theta_S}^{\theta_P} (\theta_P - \theta)d\theta.$$

Since this is linear in $\phi(\theta)$, it is immediate that $\phi(\theta) = 1$ if

$$H(\theta,\lambda) := (1-\alpha(\theta))\left(r\delta\theta - \lambda(\theta-\theta_P)\right) + \alpha(\theta)(\delta\theta - K) > 0,$$

and $\phi(\theta) = 0$ if $H(\theta, \lambda) \leq 0$.

Step 3. Routine calculation yields,

$$\frac{\partial^2}{\partial\theta^2} H(\theta,\lambda) = \frac{2}{\theta^3}\left((1-r)\delta + \lambda\right)\left(\frac{K}{(1-r)\delta}\right)^2 > 0,$$

hence, H is convex in θ. Now $H(\theta_P, \lambda) = (1-\alpha(\theta_P))(r\delta\theta_P) + \alpha(\theta_P)(\delta\theta_P - K)$. Plugging

[22] To simplify notation, we do not explicitly note the dependency of θ_P, θ_S, and $\alpha(\theta)$ on r.

$\alpha(\theta_P) = \frac{1}{\delta} - 1$, we obtain

$$H(\theta_P, \lambda) = \frac{rK}{1-r} > 0.$$

Thus, if $H \leq 0$ for some θ, we have an interval $[\theta_*, \theta^*]$ where $\theta_* > \theta_P$ and $\theta^* \in [\theta^*, 1]$ such that $H \leq 0$ if and only if $\theta \in [\theta_*, \theta^*]$. And if $H > 0$ for all θ, we set $\theta_* = \theta^* = 1$. Thus, at the optimum, $\phi(\theta) = 1$ for all $\theta \in [\theta_P, \theta_*] \cup [\theta^*, 1]$ and $\phi(\theta) = 0$ for all $\theta \in [\theta_*, \theta^*]$. The signal structure (5) follows as having $x = 1$ with certainty for all $\theta \in [\theta_*, \theta^*]$ relaxes (10) and thwarts entry of the types in $[\theta_*, \theta^*]$. ■

B The Binary-Type Case

This appendix considers the case where the seller's type θ is either ℓ or h, where $0 < \ell < h$. We use μ to denote the probability that $\theta = h$.

We maintain the following assumptions on parameter values.

Assumption 2

$$(i)\ K < \ell,\ (ii)\ \max\{(1-\delta)h, \delta\ell\} < K < \delta h,\ \textit{and}\ (iii)\ \delta h > \ell.$$

The first assumption means that the low type is relevant; if the inequality fails then the low type has no incentive to pay K and enter the platform. The second assumption corresponds to Assumption 1 in Section 2 and encodes the following two economic assumptions.

- $(1-\delta)h < K$: This means that even the high-type seller will not enter if the platform (P) will enter (and so steal his business) for sure.
- $\delta\ell < K < \delta h$: The first inequality implies that P never enters if $\theta = \ell$, while the second suggests that P wishes to enter if $\theta = h$ and r is sufficiently small.

As shown below, the final assumption ensures that P has a non-trivial incentive to enter the platform when the seller's type is h.

B.1 No entry by platform

Consider the case where P does not (or cannot) enter. In this case, the optimal r should be the level at which one of the two seller types is indifferent between entering and not doing so. For each $\theta = \{\ell, h\}$, let r_θ be the value such that

$$(1 - r_\theta)\theta = K \Leftrightarrow r_\theta := 1 - \frac{K}{\theta}.$$

Then, P prefers r_ℓ to any $r < r_\ell$ (as both seller types will enter as long as $r \le r_\ell$) and r_h to any $r > r_\ell$ (as only the high type will enter if $r \in (r_\ell, r_h)$ and no type will if $r > r_h$). This implies that r_h is optimal to P if

$$r_\ell((1-\mu)\ell + \mu h) \le r_h \mu h \Leftrightarrow \mu \ge \mu^{NE} := \frac{\ell}{\ell + (r_h/r_\ell - 1)h}.$$

If $\mu < \mu^{NE}$ then r_ℓ is optimal.

B.2 Entry by platform whenever possible (or no commitment)

Consider the case where P enters whenever profitable (i.e., when P has no commitment power over his entry). (ii) in Assumption 1 implies that P will not enter if $\theta = \ell$. Therefore, it suffices to consider P's incentive to enter when $\theta = h$.

Suppose $r \le r_\ell$. In this case, the low-type seller will enter, while the high type will not; the latter is because P will enter if $\theta = h$; observe that

$$\delta h - K \ge r\delta h \Leftrightarrow r \le 1 - \frac{K}{\delta h} < 1 - \frac{K}{\ell} = r_\ell,$$

where the last inequality is due to (iii) in Assumption 1. This implies that if $r \le r_\ell$ then P's expected profit is $r(1-\mu)\ell$, which is strictly increasing in r and so maximized when $r = r_\ell$.

If $r > r_\ell$ then the low type will never enter. In this case, it is optimal for P to extract full surplus from the high-type seller by setting $r = r_h$. It follows that r_h is optimal if

$$r_\ell(1-\mu)\ell \le r_h \mu h \Leftrightarrow \mu \ge \mu^{NC} := \frac{\ell}{\ell + h r_h / r_\ell}.$$

It is easy to show that $\mu^{NE} > \mu^{NC}$, as depicted in the figure below. This means that if $\mu \in (\mu^{NC}, \mu^{NE})$ then P chooses r_ℓ with no entry but r_h with no commitment; otherwise, P has the same optimal referral fee in the two cases. Note that this result is consistent with Proposition 1 in Section 3; P has an incentive to raise his referral fee when he has no commitment power over entry.

B.3 Full commitment to entry

Consider the case where P can commit to its entry policy, as in Section 4. As before, it suffices to consider r_ℓ and r_h, and the latter yields $r_h \mu h$ to P. Unlike before, if $r = r_\ell$ then P can extract more surplus from the high type. In particular, P can set $r = r_\ell$ and commit to enter with probability α if $\theta = h$, where[23]

$$(1-\delta\alpha)(1-r_\ell)h = K \Leftrightarrow \alpha := \frac{1}{\delta}\left(1 - \frac{K}{(1-r_\ell)h}\right) = \frac{1}{\delta}\left(1 - \frac{\ell}{h}\right).$$

By construction, the high-type seller enters for sure, but does not receive any surplus. Therefore, this policy is optimal to P conditional on $r = r_\ell$. P's corresponding expected payoff is

$$(1-\mu)r_\ell\ell+\mu\left[(1-\alpha)r_\ell h + \alpha\left((1-\delta)r_\ell h + \delta h - K\right)\right] = (1-\mu)r_\ell\ell+\mu\left[r_\ell h + \alpha\left(\delta(1-r_\ell)h - K\right)\right].$$

Comparing this to $r_h\mu h$, it follows that r_h is optimal if[24]

$$\mu \geq \mu^C := \frac{\ell}{\ell + (r_h/r_\ell - 1)h - \alpha\frac{1-r_\ell}{r_\ell}(\delta h - \ell)}.$$

It is easy to see that $\mu^C > \mu^{NE}$, so P more likely chooses $r = r_\ell$ than in the no-entry case.

B.4 Information policy

For the same reasons as in Section 5, we can restrict attention to the case where P observes either 0 or 1, and the commitment outcome for $\mu < \mu^C$ is implementable if and only if P has no incentive to enter conditional on observing 0, that is,

$$\delta\frac{(1-\mu)\ell + \mu(1-\alpha)h}{1-\mu+\mu(1-\alpha)} - K \leq r_\ell\delta\frac{(1-\mu)\ell + \mu(1-\alpha)h}{1-\mu+\mu(1-\alpha)}.$$

[23] Under our maintained assumptions, α is well defined in $(0,1)$ because

$$\alpha = \frac{1}{\delta}\left(1-\frac{\ell}{h}\right) < 1 \Leftrightarrow 1-\delta < \frac{\ell}{h}$$

and

$$(1-\delta)h < K = (1-r_\ell)\ell \Rightarrow (1-\delta) < (1-r_\ell)\frac{\ell}{h} < \frac{\ell}{h}.$$

[24] For $\mu^C < 1$, it suffices to show that

$$\left(\frac{r_h}{r_\ell} - 1\right)h > \alpha\frac{1-r_\ell}{r_\ell}(\delta h - \ell) = \frac{1}{\delta}\frac{h-\ell}{h}\frac{1-r_\ell}{r_\ell}(\delta h - \ell) \Leftrightarrow (r_h - r_\ell)h\delta h > (1-r_\ell)(h-\ell)(\delta h - \ell).$$

This inequality holds because $\delta h > \delta h - \ell$ and $(r_h - r_\ell)h = (1-r_\ell)(h-\ell) \Leftrightarrow (1-r_\ell)\ell = K = (1-r_h)h$.

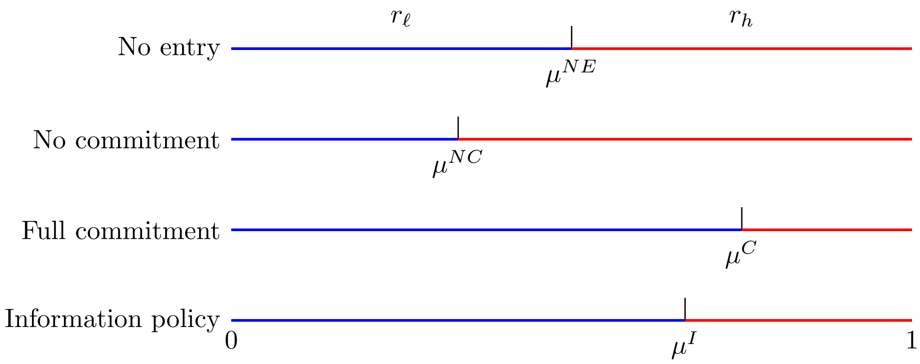

Figure 3: This figure shows when r_ℓ (or r_h) is optimal in each regime. The relative positions of μ^{NE}, μ^{NC}, and μ^{C} are always as described, and $\mu^I \leq \mu^C$ also always holds. However, μ^I is not necessarily larger than μ^{NE}.

Arranging the terms, the condition simplifies to

$$\mu \leq \mu^* := \frac{\ell}{\ell + \frac{(1-\alpha)(\delta h - \ell)}{1-\delta}}.$$

If $\mu^C \leq \mu^*$ then the above commitment outcome can be always implemented. Otherwise, as in Section 5, the optimal information policy requires certain distortions.

Consider the case where $\mu \in (\mu^*, \mu^C)$. If $r = r_\ell$ then it cannot be an equilibrium that the high-type seller enters with probability 1; if so, P will enter regardless of his signal, in which case no seller type has an incentive to enter. Let ϕ denote the probability that the high-type seller enters. For $\phi \in (0,1)$, the platform should enter with probability α if it receives a good signal; this is necessary for the high-type seller to be indifferent between entering and not entering. In addition, the following condition should hold, so that the platform has no incentive to enter after receiving a bad signal:

$$(1-r_\ell)\delta\frac{(1-\mu)\ell + \mu\phi(1-\alpha)h}{(1-\mu)+\mu\phi(1-\alpha)} \leq K \Leftrightarrow \phi \leq \phi^* := \frac{(1-\mu)(1-\delta)\ell}{\mu(1-\alpha)(\delta h - \ell)}.$$

The platform's maximized expected profit is then given by

$$(1-\mu)r_\ell\ell + \mu\phi^*\left[r_\ell h + \alpha(\delta(1-r_\ell)h - K)\right].$$

Clearly, this profit is lower than his profit under full commitment, because the latter case can be interpreted as when $\phi^* = 1$.

Recall that with full commitment to entry, P strictly prefers r_ℓ (together with entering with probability α if $\theta = h$) to r_h if $\mu < \mu^C$ and is indifferent between the two if $\mu = \mu^C$.

Combining this with the fact that $\phi^* = 1$ if $\mu = \mu^*$ and $\phi^* < 1$ if $\mu \in (\mu^*, \mu^C)$, it follows that with the optimal information policy, P strictly prefers r_ℓ to r_h if μ is close to μ^* and r_h to r_ℓ if μ is close to μ^C. Specifically, it can be shown that r_h is optimal to P if $\mu \geq \mu^I (\in (\mu^*, \mu^C))$, where μ^I is the value that satisfies

$$(1 - \mu^I) r_\ell \ell + \mu^I \phi^*(\mu^I) \left[r_\ell h + \alpha(\delta(1 - r_\ell)h - K) \right] = \mu^I r_h h.$$

References

[1] Anderson, C. (2006) "The Long Tail: Why the future of business is selling less of more." New York: Hyperion.

[2] Anderson, Simon, and Özlem Bedre-Defolie. "Hybrid platform model: Monopolistic competition and a dominant firm." *RAND Journal of Economics*, forthcoming.

[3] Arrow, K. (1962). "Economic Welfare and the Allocation of Resources for Invention." In *The Rate and Direction of Inventive Activity: Economic and Social Factors*, pp. 609–625. Princeton, NJ: Princeton University Press.

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

[5] Brynjolfsson, E., Y. Hu, and D. Simester. (2011) "Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales." *Management Science*, Vol. 57, pp. 1373–1386.

[6] Chen, Nan, and Hsin-Tien Tsai. (2021) "Steering via Algorithmic Recommendations." Working paper available at SSRN: https://ssrn.com/abstract=3500407

[7] Etro, Federico. (2023) "Hybrid Marketplaces with Free Entry of Sellers." *Review of Industrial Organization*, Vol. 62, pp. 119–148.

[8] Farronato, Chiara, Andrey Fradkin, and Alexander MacKay. (2023) "Self-Preferencing at Amazon: Evidence from Search Rankings." *AEA Papers and Proceedings*, Vol. 113, pp. 239–243.

[9] Gawer, A., and R. Henderson. (2007) "Platform owner entry and innovation in complementary markets: Evidence from Intel." *Journal of Economics & Management Strategy*, Vol. 16, pp. 1–34.

[10] Goldfarb, Avi, and Catherine Tucker. (2019) " Digital economics." *Journal of Economic Literature*, Vol. 57, pp. 3–43.

[11] Hagiu, A., T.-H. Teh, and J. Wright (2022). "Should platforms be allowed to sell on their own marketplaces?" *RAND Journal of Economics*, Vol. 53, pp. 297–327.

[12] Hagiu, A. and Wright, J. (2015) "Marketplace or Reseller?" *Management Science*, Vol. 61, pp. 184–203.

[13] Hagiu, A. and Wright, J. (2019) "Controlling vs. Enabling." *Management Science*, Vol. 65, pp. 577–595.

[14] Johnson, J. (2017) "The Agency Model and MFN Clauses." *The Review of Economic Studies*, Vol. 84, pp. 1151–1185.

[15] Johnson, J., Rhodes, A., and M. Wildenbeest. (2023) "Platform Design When Sellers Use Pricing Algorithms." *Econometrica*, Vol. 91, pp. 1841-1879.

[16] Juul, Matt. "The Amazon Buy Box Playbook for Sellers and Retailers." *Feedvisor*, July 20, 2021, available at https://feedvisor.com/resources/e-commerce-strategies/the-amazon-buy-box-playbook-for-sellers-and-retailers/.

[17] Hervas-Drane, Andres, and Sandro Shelegia. (2022) "Retailer-led marketplaces." Mimeo, Universitat Pompeu Fabra, Barcelona.

[18] Kamenica, Emir, and Matthew Gentzkow. (2011) "Bayesian Persuasion." *American Economic Review*, Vol. 101, pp. 2590–2615.

[19] Masden, Erik, and Nikhil Vellodi. "Insider imitation." *Journal of Political Economy*, forthcoming.

[20] Mattioli, D. "Amazon Scooped Up Data from its own Sellers to Launch Competing Products." *Wall Street Journal*, April 23, 2020. https://www.wsj.com/articles/amazon-scooped-up-data-from-its-own-sellers-to-launchcompeting-products-11587650015

[21] Mehta, Ivan. "All the things Apple Sherlocked at WWDC 2022." *TechCrunch*, June 13, 2022, available at https://techcrunch.com/2022/06/13/all-the-things-apple-sherlocked-at-wwdc-2022/

[22] Mueller, Florian. "Will Apple's 'Sherlocking' practice draw antitrust scrutiny? Alternative app stores would not solve–but in many cases alleviate–the problem," Foss Patents, June 16, 2022.

[23] Raval, Devesh. (2022) "Steering in One Click: Platform Self-Preferencing in the Amazon Buy Box." Unpublished manuscript.

[24] Rayo, Luis, and Ilya Segal. (2010) "Optimal Information Disclosure." *Journal of Political Economy*, Vol. 118, pp. 949–987.

[25] Roesler, Anne-Katrin, and Balazs Szentes. (2017) "Buyer-Optimal Learning and Monopoly Pricing." *American Economic Review*, Vol. 107, pp. 2072–80.

[26] Rosar, Frank. (2017) "Test Design under Voluntary Participation." *Games and Economic Behavior*, Vol. 104, pp. 632–655.

[27] Teh, Tat-How. (2022) "Platform Governance." *American Economic Journal: Microeconomics*, Vol. 14, pp. 213–54.

[28] Warren, Elizabeth. "Here's how we can break up Big Tech." March 8, 2019, available at https://medium.com/@teamwarren/heres-how-we-can-break-up-big-tech-9ad9e0da324c.

[29] Yang, H. (2013) "Targeted Search and the Long Tail Effect." *The RAND Journal of Economics*, Vol. 44, pp. 733–756.

[30] Zhu, F. and Q. Liu. (2018) "Competing with complementors: An empirical look at amazon. com." *Strategic Management Journal*, pp. 2618–42.

[31] Zou, Tianxin, and Zhou, Bo. (2023) "Self-preferencing and Search Neutrality in Online Retail Platforms." Working paper. Available at SSRN: https://ssrn.com/abstract=3987361.