Sharing is (not) caring: strategic information disclosure and platform's business model

VERY PRELIMINARY! DO NOT CIRCULATE NOR CITE IT WITHOUT THE AUTHORS' PERMISSION!

Federico Navarra* Flavio Pino† Luca Sandrini‡

March 17, 2024

Abstract

We analyse how data and data sharing influence the strategic interactions between a monopolistic platform and the many sellers operating within the digital marketplace. We investigate the effects of data sharing on market outcomes when data represent information about users preferences and can be used to price discriminate them. First, we show that data sharing has two effects: i) it intensifies competition if the platform decides to enter the market and compete against the active seller and ii) it allows sellers to fully extract consumer surplus when the platform operates as an intermediary. Thus the platform would strategically share data only with those sellers it does not want to compete with (i.e. the efficient ones) as it can exploit their efficiency through its revenues sharing fee. Second, data sharing increases theplatform's profit and decreases consumer surplus, such that the overall effect of data sharing on total welfare is ambiguous. Finally, when considering the possibility for sellers to offer their products directly to consumers, bypassing the platform, data sharing is highly pro-competitive since the platform is pushed to lower its prices in order to retain all consumers that would be tempted to buy products through sellers' direct channel.

*Charles River Associates; fnavarra@crai.com

†Polytechnic of Turin (Italy); flavio.pino@polito.it

‡Leibniz Centre for European Economic Research (ZEW), Mannheim (Germany); luca.sandrini@zew.de

1 Introduction

The relevance of data in modern economies has constantly been increasing during the past years, primarily due to the importance of digital markets. Every interaction on digital platforms and websites is tracked and registered by companies, and large amounts of data are traded every moment. Recent estimates from the European Commission show that:

> "The value of the data economy for the EU27 has been estimated to have reached almost €400 billion in 2019 and €440 billion in 2021, with a year-on-year growth rate of 4.9% in 2021. The estimated share of overall impacts on GDP in the EU27 ranges from 3.1% in 2019 to 3.6% in 2021" (DATA Market Study 2021–2023, pg. 116)."[1]

Data is a core input factor for production processes, logistics, targeted marketing, smart products, and services. Also, they are fundamental for training Artificial Intelligence and refining algorithms. On top of that, data drive interoperability in interconnected environments and are expected to drastically impact specific sectors such as mobility and healthcare.

Data owners have a large competitive advantage over their market rivals. Hence, data are very relevant to competition and privacy authorities. Digital platforms may have the incentive to adopt potential anti-competitive practices, such as self-preferencing (Padilla et al., 2022) and bias-recommendation (Bourreau and Gaudin, 2022; Gambato and Sandrini, 2023), or, more generally, they may abuse their dominant position.

To stay competitive, firms increasingly depend on timely access to relevant data and their ability to use them to develop new applications, services, and products. For these reasons, a widespread debate has emerged on whether – and under which conditions and legal bases – public intervention is required to ensure adequate and timely access to data. One of the proposed remedies is to mandate digital companies to share with complementors and rivals all or part of the consumers' data they together generate.

Data sharing is one of the pillars of the European strategy for data.[2] It is at the core of the Digital Market Act (DMA hereafter), which is the recently introduced EU competition law regulating large digital platforms' (gatekeepers) business conduct.[3] Even more relevantly, data sharing is central in the *Regulation on harmonised rules on fair access and use of data* (Data Act, hereafter). The new regulation, in force since January 11, 2024, aims to

> "Facilitate access to and the use of data by consumers and businesses, while preserving incentives to invest in ways of generating value through data. This includes increasing legal certainty around the sharing of data obtained from or generated by the use of

[1] See https://digital-strategy.ec.europa.eu/en/library/results-new-european-data-market-study-2021-2023. Accessed on January 13, 2024.

[2] See https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066&from=EN. Accessed on January 13, 2024.

[3] In particular, article 6 of the DMA states that: *"The gatekeeper shall provide business users and third parties authorized by a business user, at their request, free of charge, with effective, high-quality, continuous and real-time access to, and use of, aggregated and non-aggregated data, including personal data, that is provided for or generated in the context of the use of the relevant core platform services or services provided together with, or in support of, the relevant core platform services by those business users and the end users engaging with the products or services provided by those business users"* (DMA, Art. 6.10). Available at https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32022R1925&from=EN. Accessed on January 13, 2024.

products or related services, as well as operationalising rules to ensure fairness in data sharing contracts.[4]"

This paper analyses how data and data sharing influence the strategic interactions between a monopolistic platform and the many sellers operating within the digital marketplace. We investigate the effects of data sharing on market outcomes when data represent information about users preferences and can be used to price discriminate them. We focus on a setting where a digital platform mediates between sellers and consumers. Each seller produces and offers a good in monopoly and must pay an ad-valorem transaction fee to the platform in order to use the marketplace, whereas consumers do not pay any admission fee.

The platform decides, in order, 1) whether to share data with third-party sellers, 2) the size of the fee, 3) its business model in each market — i.e., whether to operate as a simple intermediary or to enter a market with an exact copycat of the final good — and 4), in case of entry, the price of its final good with which it will compete against the seller. We further assume that, absent any regulation, the platform has a data advantage, meaning that it knows the distribution of consumers' valuation and can use this information to target them with tailored prices. We believe that this setup is a consistent, though stylized representation of a large family of digital hosting platforms who decide to operate as first-party sellers and compete with their suppliers. Amazon Marketplace is an obvious reference for our analysis, but it is not the only one. For example, digital stores such as Apple's App Store and Google's Play Store can be represented by our model.[5] Within this game-theoretical framework, we analyze the welfare effect of a policy that mandates data-sharing from the platform to the sellers.

First, we show that the platform has incentives to share data only with some sellers. This stems from the fact that data sharing has two effects: i) it intensifies competition if the platform decides to enter the market and compete against the active seller and ii) it allows sellers to fully extract consumer surplus when the platform operates as an intermediary; hence, the platform would strategically shares data only with those sellers it does not want to compete with (i.e. the efficient ones) as it can receive part of the extracted surplus through its revenues sharing fee.
Second, it follows from the previous point that data sharing increases platform's profit and decreases consumer surplus such that the overall effect of data sharing on total welfare (here defined as the sum of profits and surplus) is ambiguous. In particular, the overall effect of data sharing on total welfare is positive (negative) when the platform has high (low) marginal costs.
Last, when considering the possibility for sellers to offer their products directly to consumers, bypassing the platform, data sharing is highly pro-competitive. This stems from the fact that the platform is pushed to lower its fee and prices in order to retain all consumers that would be tempted to buy products through sellers' direct channel.

The rest of the paper is organized as follows; the rest of this section analyzes the contribution of the paper to the relevant literature. Section 2 presents the benchmark model and assumptions; Section 3 illustrates the effect of mandating a transfer of data from the platform to the seller. Section 4 extends the result to the case in which data do not reveal the consumers' valuation of the goods, but their contact information. In this alternative model, data-sharing allows firms

[4] Available at https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52022PC0068. Accessed on January 13, 2024

[5] For an analysis with horizontally differentiated goods, see Navarra et al. (2023).

to disintermediate and directly redirect consumers towards their web-shop. Finally, Section 5 concludes.

Literature review. This paper contributes to four main strands of literature. The first one focuses on the effect of consumer data in digital economics. The use of data is widespread across every sector, thanks to their versatility; typical uses include improving products or services quality and efficiency, personalisation, matching, and discriminating between different consumers groups or individuals. Recent surveys (Goldfarb and Tucker, 2019; Bergemann and Bonatti, 2019; Pino, 2022) have thus focused on categorizing both the types and uses of data, trying to extract broader insights that hold across different models. Two typical data functions are those of allowing price discrimination on consumers and of increasing the vertical differentiation between firms (either allowing for an improvement of the products or a reduction in their marginal cost of production, see Campbell et al. (2015); Bhargava and Choudhary (2008)).

Price discrimination has been observed in various markets: a typical example involves the use of geolocalization to tailor prices to different consumers (Mikians et al., 2012; Aparicio et al., 2021). The literature has mostly focused on competition between informed firms, stemming from Thisse and Vives (1988) seminal work, and on the vertical relations between firms and a data seller (Montes et al., 2019; Bounie et al., 2021; Delbono et al., 2021; Abrardi et al., 2024). The common insight of these models is that allowing all firms to obtain data benefits consumers while it harms both the firms and the data seller. Our main contribution highlights how, when vertical integration is introduced, mandating data sharing can benefit those actors, as firms can retain more profits due to their increase in competitiveness while the platform benefits from the overall increase in market efficiency.

Our contribution to the literature is twofold. First, by isolating the direct effect of data-sharing on price competition and abstracting away the indirect incentives to invest in innovation or differentiate, we show that the policy implies two main effects on competition intensity. On the one hand, data-sharing induce more competition when the platform compete against the seller. This is so because the informed seller price more efficiently than the uninformed one. On the other hand, if the platform operates in pure agency mode, the monopolist seller can extract all the consumer surplus. We show that the tension between these two forces is ultimately dominated by the latter, as the platform increases the number of markets in which it operates as an intermediary.

Second, we explore an alternative interpretation of data as contact information. Arguably, this environment fits some prominent cases such as *Epic vs. Apple* and *Spotify vs. Apple*, where both Epic Games and Spotify complained that Apple did not allow app-developers on the AppStore to communicate with their consumers directly to promote alternative channels to subscribe. We show that the welfare effect of data-sharing in this case vary substantially from the traditional case where data allow price-discrimination. Indeed, disclosing consumers' contact information (like their emails) increase consumers surplus, as it induces the platform to enter the market more often and intensify competition to keep the prices as low as possible and retain more consumers.

The second strand of literature focuses instead on information sharing. Information sharing has been extensively studied in the literature: Raith (1996) describes a general model that summarizes many existing models, to show the determinants of when and how firms are incentivized to disclose private information. The recent literature on digital economics is also gaining interest with regard to information sharing, with particular attention on consumer data. Krämer and Schnurr (2022) focus on market contestability with regard to data-rich incumbents, and explore the possible effects of policy interventions such as data siloing, data sharing and data portability. Regarding to e-commerce, they stress the importance of sellers' data portability, as this policy would allow sellers to grow without having to lock-in on a specific platform. Prüfer and Schottmüller (2021) study competition in data-driven market where data reduce the cost of quality production. Their model shows how mandated data sharing does not reduce the dominant firm's incentive to innovate and also eliminates the risk of market tipping. Krämer and Shekhar (2022) expand on this topic by analyzing how the aforementioned policy interventions impact competition between platforms, modeling the effect of data as an improvement in the user experience on the platforms and allowing platforms to compete as well as set their investment levels. In particular, they show that mandated data sharing can reduce innovation investment by platforms, which in turn can hurt consumers when data externalities are large. De Corniere and Taylor (2020) analyze the effects of data sharing by using a competition in utilities approach, finding sufficient conditions under which data sharing would be unambiguously pro-competitive. Liu et al. (2021) focus instead on a retail platform that hosts sellers and can strategically disclose information to them: they find that the platform has the incentive to disclose information only to a subgroup of sellers. (Magnani and Navarra, 2023) study the incentives of a monopolistic hybrid platform in sharing its superior market information with a third-party seller hosted on its marketplace: they find that, despite platform duality, the platform has incentives to share information with the seller in order to relax downstream price competition. While our work focuses on an exogenous shock that mandates complete information sharing instead of allowing for a strategic decision by the platform, as far as we know we are the first to allow the platform to vertically integrate, entering the downstream market and competing with sellers.

Related to the vertical integration aspect of our model, the third strand of literature concerns the classical questions regarding access pricing and sabotage. Indeed, our model resembles the typical setup of an upstream monopolist that controls an infrastructure and can choose to integrate downstream. Economides (1998) shows how an integrated monopolist has the incentive to degrade the quality of the downstream input, as to raise the costs of its rivals until they are driven out of the market. Beard et al. (2001) expand on this topic by showing that the upstream monopolist is always willing to expand downstream, but has the incentive to sabotage only when input price regulation is introduced. Our model presents a similar result: the platform can strategically use the per-transaction fee to increase downstream costs, allowing it to better compete against sellers. Moreover, in the case of price discrimination with heterogeneous goods, the platform sets the fee such that sellers opt to price as monopolists, increasing surplus extraction from consumers and, in turn, the platform's profits.

Finally, we contribute to the literature on hybrid marketplaces — i.e., platforms that allow transactions between sellers and buyers and where the platform can become a seller's competitor. This literature is becoming pivotal in policy discussion, as tech giants such as Amazon and Apple are themselves hybrid marketplaces (see Hagiu et al., 2022; Teh, 2022; Casner and Teh, 2023). Empirical evidence suggests that the downstream entry of the platform, sometimes referred to as *dual mode*, usually takes place in successful markets and leads sellers to reduce their growth efforts in the platform (Zhu and Liu, 2018). Moreover, sellers tend to increase the prices in the markets where the platform enters, while shifting their innovation investments elsewhere (Wen and Zhu, 2019). In particular, evidence suggests that complementary goods become the focus of innovation, as the platform entry usually expands the demand for that good (Foerderer et al., 2018). From a theoretical perspective, the effects of a platform operating in dual mode are ambiguous. On the one hand, platform downstream entry could reduce sellers' market power and increase competition, benefiting consumers (Dryden et al., 2020; Etro, 2021b). Platform entry could also induce it to reduce its commission fees to further expand the market's reach (Etro, 2021a). On the other hand, a higher quality (or lower cost) of the platform's goods can incentivize it to increase its commission fees, ultimately harming consumers (Anderson and Bedre-Defolie, 2021).

2 General model

The market environment. We consider a monopolistic platform which hosts many markets. In each of them, the platform can decide to operate in agency or dual mode. Agency mode is a business model that describes a situation in which the platform acts as pure intermediary between the sellers and the consumers, collecting revenue sharing fees. Instead, dual mode describes a situation where the platform intermediates between sellers and consumers while also selling directly to consumers a proprietary version of the final good.

We consider a mass of independent sellers ranked by their marginal costs of production c^s which we assume uniformly distributed between 0 and ∞. With a slight abuse of notation, we identify marginal costs with a technology. Each technology is allocated to only one seller, and each seller sells independent goods. In each market, the platform decides whether to enter and operate in dual mode, or to stay out and operate in agency mode.

There is a unitary mass of consumers for every product space (namely for every seller) and consumers are split into two groups according to their willingness to pay $v \in \{\underline{v}, \bar{v}\}$ with $\underline{v} < \bar{v}$. Groups have a mass of $\underline{n}$ and $\bar{n}$ respectively, such that $\underline{n} + \bar{n} = 1$. Every consumer buys at most one unit of any given good. Both the platform and sellers know the distribution of the willingness to pay but sellers cannot always price discriminate consumers — i.e., they cannot *identify* the willingness to pay of each consumer at the individual level.

The platform intermediates between sellers and consumers and impose on the former group a revenue sharing fee f. The fee is unique for all markets, meaning that the platform does not price discriminate the sellers.[6] If the platform enters a market, it competes in prices *à la Bertrand* against the seller active in that product space. As the other sellers, the platform incurs a marginal production cost $c^a \in [0, \infty)$ which is equal across product spaces.[7]

Platform cost c^a is known by the platform and revealed to sellers when the platform enters their markets. In order to capture the fact that the platform exploits its data advantage to avoid retail risk, we impose that the platform can only produce homogeneous copycat versions of existing goods. Hence, the platform cannot *activate new markets*, nor personalize their goods.[8] Instead, it can enter those markets where a seller is already active.[9] In other words, we never consider the platform to be a monopolist on its marketplace for a given product space.

[6] Our modeling choice is consistent with Amazon's practice of setting a specific fee per product group. This type of fee is labeled by Amazon as a "referral fee": *Amazon charges a referral fee on each item sold. The amount varies depending on the product group.* See https://sell.amazon.com/pricing#referral-fees. Our model thus aptly represents one of Amazon's product groups, which is composed of multiple sub-markets. Our analysis could be easily replicated in any one of Amazon product groups to find the optimal fee for each of them.

[7] The assumption of a single technology that the platform can use to produce every good allows us to generate asymmetries between the platform and the sellers. We are interested in analyzing a situation in which the platform could be more or less efficient than the relevant seller, depending on the sub-market they are operating in. Furthermore, this assumption allows us to keep the model as simple and clear as possible.

[8] See Navarra et al. (2023) for a detailed analysis of the problem with horizontally differentiated goods.

[9] See https://www.reuters.com/investigates/special-report/amazon-india-rigging/ and Madsen and Vellodi (2022)

The timing of the game. First, the platform decides whether to share data or not at time $t = 0$; the platform then sets the revenue sharing fee f at time $t = 1$. Accordingly, the platform also decides its business model in each market (agency or dual mode). Finally, at time $t = 2$ price competition ensues in markets where the platform operates in dual mode, whereas the sellers set monopoly prices in markets where the platform operates in agency. The solution concept is the Sub-game Perfect Nash Equilibrium (SPNE).

2.1 Baseline set-up (no data sharing)

We begin our analysis by considering data as an essential tool for price discrimination. In other words, if sellers could access data they were able to distinguish consumers and their actual willingness to pay, thus setting different prices for the same good. Otherwise, denied the access to data, each seller sets a uniform price for all consumers. In our model, the platform is endowed with the data and always price discriminates if it enters the market. When the platform shares its data, sellers price discriminate as well.

2.1.1 Platform data advantage (no data sharing)

Starting from the last stage of the game, consumers in each market buy a good if and only if the price they pay is lower than their willingness to pay. Consider a market i. Given the binomial distribution of consumers' willingness to pay, an individual j buy the good if $p^k_{ji} \leq v_{ji}$ — i.e., if the price offered by firm $k = a, s$, where a and s stand for the platform and the seller, respectively, to consumer j in market i is not greater than j's individual valuation of the good. We define p^s_i the uniform price charge by the uninformed seller in the market. Instead, $p^k_{i,j}$ is the personalized price that the platform or the informed seller offers to each consumer. In this benchmark, the seller has no access to data, hence it can only set a uniform price.

Monopolist seller. First, let's investigate those markets where the platform decided to operate in pure agency mode. There, the seller is a monopolist and can decide its price independently of the platform reaction. Because the seller is uninformed, it cannot identify the willingness to pay of each consumer. However, it knows that $\underline{n}$ and $\bar{n}$ consumers have willingness to pay $\underline{v}$ and $\bar{v}$, respectively. Thus, in a market i, the seller will choose to either set a high price and serve only "high-"type consumers, or to set a low price and to serve the entire product market. More in detail, in deciding whether to serve all consumers or just $\bar{n}$ of them, the seller compares:

$$\underline{v}(1-f) - c^s_i > \bar{n} \times (\bar{v}(1-f) - c^s_i)$$

From this condition, we derive the following Lemma:

Lemma 1. *A monopolist seller in market i sets*

$$p^s_i = \begin{cases} \underline{v} & if \quad c^s_i \leq \frac{(\underline{v}-\bar{v}\bar{n})(1-f)}{1-\bar{n}} \\ \bar{v} & if \quad c^s_i > \frac{(\underline{v}-\bar{v}\bar{n})(1-f)}{1-\bar{n}} \end{cases}$$

Accordingly, the seller's and platform's payoffs in this scenario are:

$$\pi_i^s = \begin{cases} \underline{v}(1-f) - c_i^s & \text{if} \quad c_i^s \leq \frac{(\underline{v}-\bar{v}\bar{n})(1-f)}{1-\bar{n}} \\ \bar{n}(\bar{v}(1-f) - c_i^s) & \text{if} \quad c_i^s > \frac{(\underline{v}-\bar{v}\bar{n})(1-f)}{1-\bar{n}} \end{cases} \qquad \pi_i^a = \begin{cases} \underline{v}\, f & \text{if} \quad c_i^s \leq \frac{(\underline{v}-\bar{v}\bar{n})(1-f)}{1-\bar{n}} \\ \bar{n}\, \bar{v}\, f & \text{if} \quad c_i^s > \frac{(\underline{v}-\bar{v}\bar{n})(1-f)}{1-\bar{n}} \end{cases}$$

Finally, consumers receive utility:

$$CS = \begin{cases} \bar{n}(\bar{v} - \underline{v}) & \text{if} \quad c_i^s \leq \frac{(\underline{v}-\bar{v}\bar{n})(1-f)}{1-\bar{n}} \\ 0 & \text{if} \quad c_i^s > \frac{(\underline{v}-\bar{v}\bar{n})(1-f)}{1-\bar{n}} \end{cases}$$

Competition in the market. Consider now the case where the platform decided to enter the market and compete in prices against the seller. Recall that, in this baseline model, the platform possesses a data advantage, meaning that it is able to identify which consumers belong to the two groups $\bar{n}$ and $\underline{n}$. Also, we assume that in case of equal prices, the consumers prefer the original good, i.e., the one offered by the seller.

With that in mind, let's investigate the strategies of the two agents. It is immediate to observe that, for $c^a \geq \bar{v}$, the platform can never profitably win the price competition, as no consumer is willing to pay a price that would earn the platform a margin. Therefore, entry in the market is only observed when $c^a < \bar{v}$. When this condition is verified, we can analyze the firms' strategy at the price competition stage. Both set uniform prices that are observable by anyone. Additionally, the platform offer secrets deals to consumers that cannot be observed by the sellers.

Competition between the platform and each seller can occur in many ways. Here, we distinguish three main features of competition:

Hybrid model: the platform may decide to enter only to cover a share of the demand (low-type consumers) that the seller is not willing to serve. To do so, the platform set tailored prices for low-type users. We define this strategy as *hybrid model*, because the platform is both operating an agency business model for high-type consumers, while selling the good directly to low-type consumers. The profit functions of the two firms, the platform and the seller, respectively, in a given market i are:

$$\pi_i^a = (1-\bar{n})(\underline{v} - c^a) + f\,\bar{n}\,\bar{v}, \qquad \pi_i^s = (1-f)\,\bar{n}\,\bar{v} - c_i^s$$

This strategy can be implemented if the platform is relatively efficient, i.e. if $c^a < \underline{v}$. Otherwise, serving the low-type consumers would represent a cost. Moreover, it requires that the seller is not independently serving the low-type workers — i.e., $c_i^s > \frac{(\underline{v}-\bar{v}\bar{n})(1-f)}{1-\bar{n}}$. If those condition applies, one can notice that the *hybrid model* dominates the pure agency mode, as it guarantees additional revenues without altering the revenues from intermediating the high-type consumers and the seller.

Constrained agency model: the platform may decide to intervene more aggressively. It can do so by *faking* entry and quoting a uniform price that forces the active seller to lower its own price. We define this strategy as *constrained agency model*, because the platform is effectively operating an agency business model, while constraining the price of the seller to be such that the entire market is covered. In this case, the profits of the two firms, the platform and the seller, are, respectively:

$$\pi^a = \underline{v}\, f; \qquad \pi_i^s = \underline{v}(1-f) - c_i^s$$

This strategy can be profitable if the platform's costs of production are not particularly low, so that it prefers capturing a share of the seller per-transaction revenues rather than producing and selling the good directly to consumers. However, for the strategy to be credible, the platform must be able to profitably serve low-type consumers — i.e., $c^a < \underline{v}$. Furthermore, this strategy requires that the seller can lower its price to the point where low-type consumers activate — i.e., $c_i^s < (1-f)\underline{v}$. It also requires that the seller is not already covering the entire market independently — i.e., $c_i^s > \frac{(\underline{v}-\bar{v}\bar{n})(1-f)}{1-\bar{n}}$. If these conditions hold, the platform sets a price $p_i^a = \underline{v}$ which is matched by the seller. The market is then fully covered and there is no reason for the platform to push the seller to further reduce its price.

Competition model: the platform can enter the market and out-compete the active seller. We define this strategy *competition model*, because the platform aims to win price competition and serving the market directly. To win the price competition, a necessary requirement is that the platform faces a marginal cost of production lower than the one faced by the seller, adjusted by the per-transaction ad-valorem fee. Formally, *competition model* requires $c^a < \frac{c_i^s}{1-f}$. If the seller is sufficiently efficient to profitably serve the entire market ($c_i^s < (1-f)\underline{v}$), the platform's data advantage plays no role, as price competition pushes the price of the seller to its marginal costs, thus activating all the market. If instead the seller cannot profitably cover the entire market, then the platform adopting a *competition model* compete with the seller for high-type consumers and set a tailored price to the low-type ones (if profitable).

In this case, the profits of the platform is:

$$\pi^a = \begin{cases} \frac{c_i^s}{1-f} - c^a & \text{if} \quad c^a < \frac{c_i^s}{1-f} < \underline{v} \\ (1-\bar{n})\underline{v} + \bar{n}\,\frac{c_i^s}{1-f} - c^a & \text{if} \quad c^a < \underline{v} < \frac{c_i^s}{1-f} \\ \bar{n}\left(\frac{c_i^s}{1-f} - c^a\right) & \text{if} \quad \underline{v} < c^a < \frac{c_i^s}{1-f} \end{cases}$$

whereas the seller earns $\pi_i^s = 0$.

2.2 The choice of the business model

Starting from the condition stated in Lemma 1, we can now derive the optimal choice of the business model in different parameter regions. We distinguish two main scenarios: when the share of high-type consumers is relatively large — i.e., $\bar{n} \in (\underline{v}/\bar{v}, 1)$ — and when it is relatively small — i.e, $\bar{n}(0, \underline{v}/\bar{v}]$.

Large share of high-type consumers. When there is a large number of high-type users relatively to low-type ones, the seller has the incentives, absent any form of competition, to set a high price and leave low-type consumers uncovered — i.e., condition $c_i^s \leq \frac{(\underline{v}-\bar{v}\bar{n})(1-f)}{1-\bar{n}}$ cannot be satisfied. In this scenario, we have three sub-cases to account for. In particular, we identify two main areas in terms of platform's efficiency that are relevant for a proper identification of the optimal business strategy. The threshold that disentangles between these two areas is $\underline{v}$.

Consider first the case in which the platform faces a marginal cost of production $c^a < \underline{v}$. As we observed above, in this case, pure agency model is dominated by *hybrid model*. In fact, knowing the the seller would not cover the low-type consumers, the platform has a strong incentive to enter with an *hybrid model* and personally take care of the production and selling of the final good to the $\underline{n}$ share of low-type users. Additionally, one can see that in this scenario *constrained agency model* is also dominated by the *hybrid model*.

Thus, what we are left with is to compare the *hybrid model* with the *competition model*. The next Lemma summarizes the analysis:

Lemma 2. *Assume $c_i^s > \frac{(\underline{v}-\bar{v}\bar{n})(1-f)}{1-\bar{n}}$ and $c^a < \underline{v}$. The platform chooses the hybrid model if the following condition is satisfied:*

$$c_i^s < \max\{\tau_{1,i};\ \tau_{1,ii}\} \equiv \tau_1 \quad \textit{with} \quad \tau_{1,i} > \tau_{1,ii} \quad \textit{if} \quad c^a < \underline{v} - f\,\bar{v}$$

with $\tau_{1,i} = (1-f)(\bar{n}(c^a + f\bar{v}) + (1-\bar{n})\underline{v})$ and $\tau_{1,ii} = (1-f)(c^a + f\bar{v})$. Otherwise, it chooses the competition model. Hence, the aggregate profit of the platform across all markets can be written as:

$$\Pi^a(f) = \int_0^{\tau_1} (1-\bar{n})(\underline{v}-c^a) + f\,\bar{n}\,\bar{v}\ dc^s + \int_{\tau_1}^{(1-f)\bar{v}} \begin{cases} \frac{c^s}{1-f} - c^a\ dc^s & \textit{if} \quad \tau_1 = \tau_{1,i} \\ (1-\bar{n})\underline{v} + \bar{n}\frac{c^s}{1-f} - c^a\ dc^s & \textit{if} \quad \tau_1 = \tau_{1,ii} \end{cases} \tag{1}$$

Let us now turn to the case where $c^a > \underline{v}$. Here, the platform cannot profitably serves the low-type consumers. Hence, both the constrained agency and the hybrid model cannot be adopted. We then compare the incentives of the platform to adopt an agency or a competition model. The next Lemma summarizes the analysis:

Lemma 3. *Assume $c_i^s > \frac{(\underline{v}-\bar{v}\bar{n})(1-f)}{1-\bar{n}}$ and $c^a > \underline{v}$. The platform chooses the competition model if the following condition is satisfied:*

$$c_i^s > (1-f)(c^a + f\bar{v}) \equiv \tau_2$$

Otherwise, it operates in pure agency. Hence, the aggregate profit of the platform across all markets can be written as:

$$\Pi^a(f) = \int_0^{\tau_2} f\,\bar{n}\,\bar{v}\ dc^s + \int_{\tau_2}^{(1-f)\bar{v}} \bar{n}\left(\frac{c^s}{1-f} - c^a\right) dc^s \tag{2}$$

Intuitively, results in Lemmata 2 and 3 reveal a simple decision mechanism by the platform: *compete against inefficient sellers and exploit the efficient ones*. In other words, the platform generates a barrier to entry that keep inefficient sellers out of business. At the same time, as efficient sellers are better off by serving only the share of demand with high-type consumers, the platform can enter the market and serve to low-type ones, while simultaneously capturing a cut of the seller revenues from trading with high-type consumers.

Small share of high-type consumers. When the share of high-type consumers is relatively small — i.e., $\bar{n} < \underline{v}/\bar{v}$. In this case, all scenarios are ex-ante possible, as serving the entire market might be profitable for some sellers.

Let's take a step-by-step approach. If we narrow our focus on the region of parameters in which a monopolist seller would not serve the entire market — i.e., $c_i^s > \frac{(\underline{v}-\bar{v}\bar{n})(1-f)}{1-\bar{n}}$ — then, Lemmata 2 and 3 apply. Indeed, the problem is the same as in the case analyzed just above and the incentives of the platform do not vary. Some condition must, however, be checked.

Now, we turn our attention at what happens if the active seller in a market is particularly efficient, and therefore willing to serve the entire market even absent any competitive pressure from the platform — i.e., when $c_i^s \leq \frac{(\underline{v}-\bar{v}\bar{n})(1-f)}{1-\bar{n}}$. As illustrated above in the description of the monopolist seller case, this scenario occurs when the trade-off between larger price (intensive margin) and larger customer base (extensive margin) is dominated by the latter for the efficient sellers. Indeed, because they face lower marginal costs of production, efficient sellers can profitably lower their prices for the final good, while still appropriating sufficiently large surplus. Accordingly, *constrained agency mode* appear to be redundant when competing against such sellers. Similarly, the *hybrid model* doesn't seem to pay off either, as the seller's price in those market already satisfies the low-type consumers' participation constraint.

What's left is to compare the *competition model* with the outcomes available to the platform if it operates in pure agency. The next Lemma summarizes the analysis.

Lemma 4. *Assume* $c^a(1-f) < c_i^s \leq \frac{(\underline{v}-\bar{v}\bar{n})(1-f)}{1-\bar{n}}$*, which implies* $c^a < \frac{\underline{v}-\bar{n}\bar{v}}{1-\bar{n}}$*. The platform prefers operating in competition model if the following condition holds:*

$$c_i^s > (1-f)(c^a + f\,\underline{v}) \equiv \tau_3$$

otherwise, it chooses to operate in pure agency.

For the above conditions to be simultaneously satisfied, it must be that

$$0 < c^a < \frac{\underline{v} - \bar{n}\bar{v}}{1-\bar{n}} - f\,\underline{v} \quad \Longrightarrow \quad \bar{n} < \frac{\underline{v}}{\bar{v}}\left(\frac{1-(1-\bar{n})f}{\bar{n}}\right) < \frac{\underline{v}}{\bar{v}}$$

Hence, the aggregate profit of the platform across all markets can be written as:

$$\Pi^a(f) = \int_0^{\tau_3} f\,\underline{v}\; dc^s + \int_{\tau_3}^{\frac{(\underline{v}-\bar{v}\bar{n})(1-f)}{1-\bar{n}}} \frac{c^s}{1-f} - c^a dc^s \tag{3}$$

Put simply, *competition mode* is an option if and only if the share of high-type consumers is very small. Otherwise, not even a platform that produces its final good for free would compete against an efficient seller.

2.3 Data-sharing

The platform has more than an instrument to "regulate" its own marketplace. In particular, the platform can i) use the price to exert competitive pressure on the seller, in order to increase market coverage; alternatively, it can ii) adjust the revenue sharing fee to influence the price of the sellers in all markets; finally, iii) by sharing information about consumers, the platform can increase the seller's pricing efficiency.

All these instruments alter the platform's profitability. More in detail, both competitive pressure and adjustments of the fee affect prices (and per-transaction revenues) and the number of trades. Put simply, the standard trade-off between intensive and extensive margins regulates the platform's strategic choices. Instead, by sharing information about consumers, the platform alters its entry opportunities in the markets. Intuitively, if both the seller in a specific market and the platform can produce the final good very efficiently, and the market characteristics are such that not all the demand is covered with the uniform price, it might be that the platform has the incentive to leave the seller uninformed. Take the following example where:

$$c^a < \underline{v}(1-f) \quad \& \quad \frac{(1-f)(\underline{v}-\bar{n}\bar{v})}{1-\bar{n}} < c_i^s < \min\{\tau_{1,i}, \underline{v}(1-f)\}$$

i.e., Lemma 2 applies and, in particular, the platform adopts a *hybrid model* and does not explicitly compete against the seller in this specific market. Here, the platform earns $\underline{v}-c^a$ from every low type consumer that it attracts via a personalized offer. Had the platform shared info about consumers type with the seller, it would now face the choice between *competition model* and agency. The *hybrid model* strategy is turned off in this scenario, as avoiding competition upon entry is not possible — i.e., the seller has always the incentive to set a personalized price to the low type consumers.

If the platform chooses agency, it gets $f\underline{v}$; if instead it competes, the payoff per low-type consumer is $\frac{c_i^s}{(1-f)} - c^a$. The latter payoff is dominated by the ex-ante *hybrid model*'s outcome as, by assumption $c_i^s < \underline{v}(1-f)$. Comparing agency and *hybrid model*, instead, leads to the following result

$$\underline{v} - c^a > f\underline{v} \quad \text{if} \quad c^a < \underline{v}(1-f)$$

The platform would never share the data with an efficient seller if she is very efficient too. In fact, the information asymmetry allows the platform to exploit the inefficient pricing of the seller and maximize revenues. On the other side, however, a less efficient platform that understands that the seller is able to cover all the market more efficiently may decide to share the data and operate in pure agency.

Proposition 1. *The platform, if possible, would strategically share data only with efficient sellers, namely those who are active in the markets where the platforms operates in agency model.*

When the platform shares data, third-party sellers are overall more profitable and the platform is able to raise the fee (see Figure 1). However, when the platform is less cost efficient it has the incentives to lower the fee even below the level without data-sharing, this compensates for the high cost of the platform and enables a sufficiently high number of firms to join the marketplace.
However, with data sharing consumers pay a price equal or higher than without data sharing, thus consumer surplus is never enhanced by this strategy, as shown in Figure 2.
Overall, the social welfare (defined by an utilitarian function) is generally higher with data sharing although we may observe parameter combinations for which welfare is higher in the standard set-up.

Proposition 2. *Both the platform and firms benefit from data sharing while consumer are harmed. The overall effect of data sharing on total welfare is positive (negative) when the platform has high (low) marginal costs.*

Figure 3 shows the ambiguous impact of platform's data sharing on the utilitarian total welfare. This result is driven by the level of the fee, which is much lower without data sharing when the cost of the platform is low than with data sharing, resulting in higher consumer surplus to the extent that outweighs lower profits absent data-sharing.

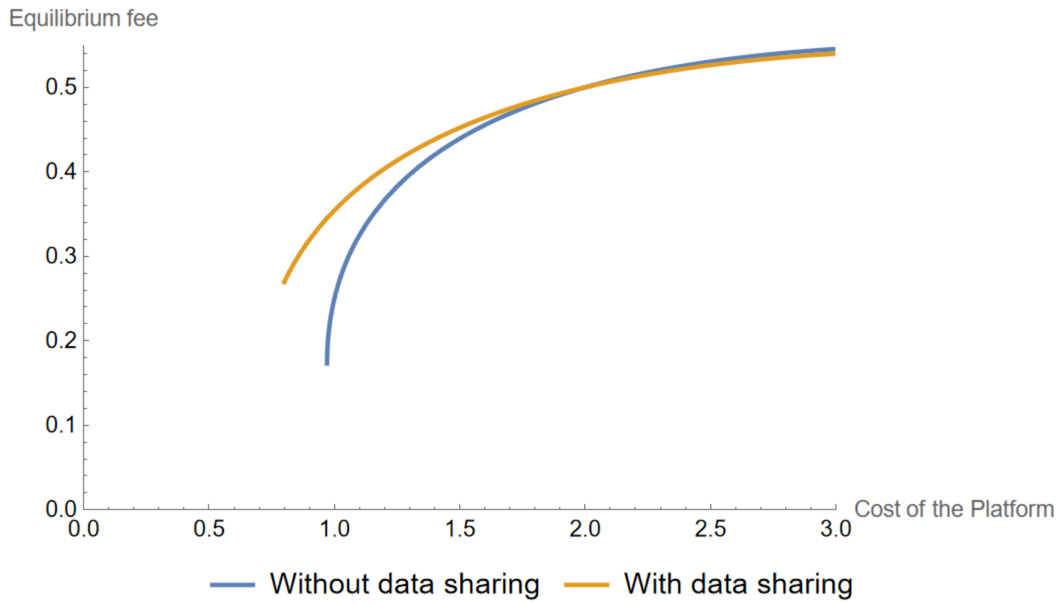

Figure 1: Equilibrium platform fee with and without data sharing as a function of c_p (with $\bar{v} = 4$, $\underline{v} = 1$ and $\bar{n} = 0.3$)

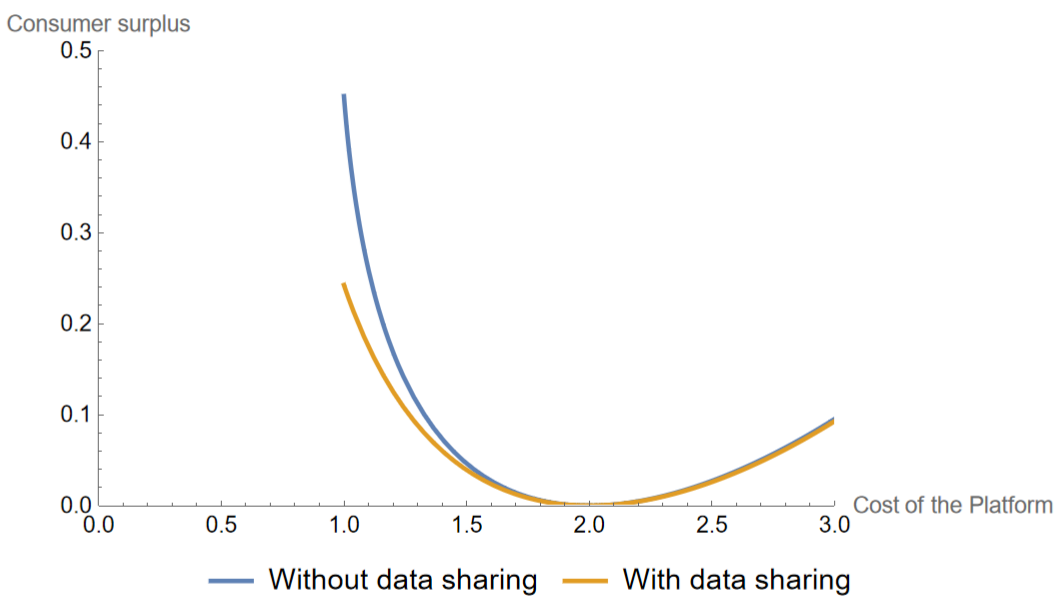

Figure 2: Consumer surplus with and without data sharing as a function of c_p (with $\bar{v} = 4$, $\underline{v} = 1$ and $\bar{n} = 0.3$)

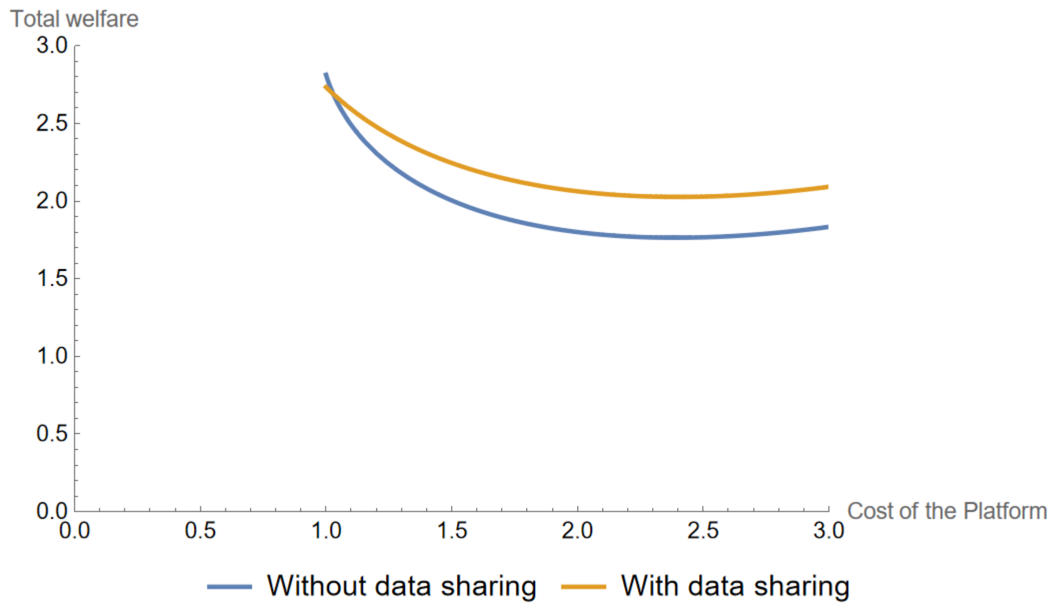

Figure 3: Total welfare with and without data sharing as a function of c_p (with $\bar{v} = 4$, $\underline{v} = 1$ and $\bar{n} = 0.3$)

2.3.1 Discussion

The analysis above suggests that data-sharing can be used by the platform to govern the interaction between sellers and consumers and generate the highest surplus. However, the platform would use it strategically, sharing information only in those markets where it can profitably do so. The *private* equilibrium would look like a marketplace where some sellers are informed while some others are not.

A mandated data-sharing policy has the virtue to force the platform to share data with all sellers in all markets the platform would not share them. By doing so, mandated data sharing can induce a positive effect on consumers, as the informed seller might decide to compete, if possible, against the platform to attract all consumers of low-type.

Yet, data-sharing *per se* is not necessarily a better outcome than the one in which there is data asymmetry or no data at all. Indeed, in this model, the platform benefits from sellers' large prices via the ad-valorem transaction fee. Hence, from the platform's perspective, sharing info makes the seller more efficient in pricing and increases the average per-transaction revenue rate. This is particularly true if the seller is, absent any information, willing to cover the entire market at a low price. In a situation like this, the regulator might paradoxically prefer forbidding data-sharing, rather than promoting and mandating it.

3 Extension

Consider now a modified version of the model, where data have not a price-discrimination purpose, but it merely contains contact information of the consumers on the platform. Sharing those data allows the sellers to communicate directly with the potential consumers and offer them the good outside of the platform's marketplace, posing the risk of disintermediation.

In particular, consider consumers the same set of agents with the same characteristics as in the baseline model, but now, with data sharing, the sellers can tempt a (minority) share $\rho \in (0, 1/2)$ of them in each group (high- and low-willingness to pay type) to buy the good directly from the merchants' website at a price which is discounted by a positive factor $\eta \in [0, f]$, as there is no transaction fee involved. The assumption that only an exogenous and small share of consumers would turn to the direct channel of the seller is imposed to ensure that the platform maintains a sizable market power, and proxies the presence of barriers such as brand reputation, consumer inertia, and other behavioural bias that makes digital market weakly contestable.[10] The strategy of the sellers changes as there is an additional channel to promote their goods. In particular, the sellers do not compare, as in the baseline specification, $\underline{v}(1-f) - c_{s_i} > \bar{n}(\bar{v}(1-f) - c_{s_i})$, but rather

$$\underline{v}(1-\rho)(1-f) + \rho\, \underline{v}(1-\eta) - c_{s_i} > \bar{n}((1-\rho)\bar{v}(1-f) + \rho\, \bar{v}(1-\eta) - c_{s_i})$$

Notice that we do not allow for the sellers to price discriminate consumers. The only assumption is that Price-Parity clauses are not enforced, so that the sellers can charge a lower price for the same good on different distribution channels.[11] To keep the analysis simple and conservative, we assume that $\underline{v} < \bar{v}(1-\eta)$ — i.e., that the discounted price for high-willingness to pay consumers offline cannot activate the market for low-willingness to pay consumers.

We update Lemma 1 accordingly:

Lemma 5. *A monopolist seller in market i sets $p^{on}_{s_i} = \underline{v}$ online and $p^{off}_{s_i} = \underline{v}(1-\eta)$ offline if $c_{s_i} < \frac{(\underline{v}-\bar{v}\bar{n})((1-f)(1-\rho)+(1-\eta)\,\rho)}{1-\bar{n}}$ while it sets $p^{on}_{s_i} = \bar{v}$ online and $p^{off}_{s_i} = \bar{v}(1-\eta)$ offline when $c_{s_i} > \frac{(\underline{v}-\bar{v}\bar{n})((1-f)(1-\rho)+(1-\eta)\,\rho)}{1-\bar{n}}$*

The first observation is that, with respect to the baseline case with no data sharing, the monopolist seller is able and willing to serve all the market for lower level of efficiency, given the ability to profitably reach some of the consumers offline. This is somehow intuitive, as the possibility to sell the good for a higher return on the direct channel induces the sellers to cover the market.

Let us now consider the problem of the platform. As before, the platform has three strategies: i) It can operates as an intermediary and leave the market to the sellers; ii) It can produce a copycat product and enter the market, thus operating in dual mode; iii) It can produce a copycat, enter the market, but strategically lose the competition, if the revenues from intermediation are larger than the ones from direct sales. It should be noted that under the assumption that data sharing allows the sellers to reach some consumers offline, the incentives of the platform to foreclose the sellers increase. In fact, by letting the sellers winning the price competition, the platform implicitly allows some price heterogeneity among channels (the marketplace and the merchant's website). To avoid this situation, the platform could enter the market with a copycat product and secure all the demand at a price that matches the offline one. By doing so, the platform prevents, in principle, disintermediation from happening.

[10] Arguably, this environment fits the case of *Spotify vs. Apple*, where Spotify complained that Apple did not allow app-developers on the AppStore to communicate with their consumers directly to promote alternative channels to subscribe.

[11] Following lawsuites by antitrust authorities in the EU, Amazon decided not to enforce PPCs in Europe from 2017. In the US, PPCs by Amazon were dropped from 2019.

When the firm operates in agency mode, its profits can be written as:

$$\Pi_i^{ag} = \begin{cases} \underline{v}\,(1-\rho)\,f & \text{if} \quad c_{s_i} < \frac{(\underline{v}-\bar{v}\bar{n})((1-f)(1-\rho)+(1-\eta)\,\rho)}{1-\bar{n}} \\ \bar{n}\,\bar{v}(1-\rho)\,f & \text{if} \quad c_{s_i} > \frac{(\underline{v}-\bar{v}\bar{n})((1-f)(1-\rho)+(1-\eta)\,\rho)}{1-\bar{n}} \end{cases}$$

Let's now consider the case in which the platform decides to enter the market and compete à la Bertrand against the active seller. In this case, price competition ensues and the platform wins it whenever it can profitably undercut the seller's price.

The platform is not only competing against the seller on the marketplace, but it is also trying to retain the share of consumers ρ who get the personalized offer directly by the seller. The platform cannot exactly identify who are the consumers who are considering buying from the seller directly (it know ρ, but not who is part of it). Therefore, to prevent the sellers and consumers from profitably disintermediating, the platfrom needs to reduce its price even further than $c_s/(1-f)$. In fact, the seller can cross-subsidize its indirect channel on the marketplace by offering a smaller discount on its direct channel. Eventually, the average marginal cost faced by the seller on the two channels is $\bar{c}_{s_i} \equiv \rho c_{s_i} + (1-\rho)\frac{c_{s_i}}{1-f}$. In order to undercut the seller, the platform needs to set a price which is $p_a < \bar{c}_{s_i}$. This is however not enough to prevent some consumers from leaving the platform. In fact, if the seller sets a price on the direct channel which is $p^d_{s_i} \in (c_{s_i}, \bar{c}_{s_i})$, it can still attract a share ρ of consumer for a margin. In order to regain those consumers, the platform must further reduce its price to $p_a = c_{s_i}$.

Because the seller cannot quote a marketplace price below $p_{s_i} = \bar{c}_{s_i}$, the platform will only consider entry if $c_p < \bar{c}_{s_i}$. Hence, our first result in this scenario is that the platform will never consider entry in Bertrand competition for $c_p > \bar{c}_{s_i} \leq \frac{c_{s_i}}{1-f}$ for all $\rho \in [0,1]$.

Notice that the ability of the platform to raise the rival cost is hampered by the disintermediation effect, as the more people are reached by the seller on its direct channel, the less sensible the average cost to a variation of the fee. Formally,

$$\frac{\partial \bar{c}_{s_i}}{\partial f} = \frac{(1-\rho)c_{s_i}}{(1-f)^2} \qquad \text{and} \qquad \lim_{\rho \to 1} \frac{(1-\rho)c_{s_i}}{(1-f)^2} = 0$$

Small share of high-type consumers. In this scenario, the monopolist seller would independently serve the entire market. We already saw that the platform would not consider entry if $c_a > \bar{c}_{s_i}$. We can rewrite this condition in terms of c_{s_i} and write $c_{s_i} < \frac{c_a(1-f)}{1-f\,\rho}$.

Let's see what happens if $c_a < \bar{c}_{s_i}$. In this case, the platform can, in principle, enter the marketplace and out-compete the rival seller. Bertrand competition pushes prices toward the average marginal cost of the seller $p^m_{s_i} = \bar{c}_{s_i}$, $p_a = \bar{c}_{s_i}$, where the superscript m stands for the price on the marketplace. As the seller cannot price below the average marginal cost on the marketplace, the platform can secure the entire market, except those ρ consumers who got an offer from the direct channel of the seller. To them, the seller can offer a price $p^d_{s_i} = \bar{c}_{s_i}(1-\eta)$

The profit of the platform in case of entry is:

$$\Pi_i^{comp} = (1-\rho)(\bar{c}_{s_i} - c_a)$$

Notice that in this scenario, as the seller would independently cover the entire market, the strategy of entry only to nudge the seller to lower its price is clearly ruled out. The platform decides what is better between operating as a pure intermediary (agency) and taking over the market whit a first party product. In particular, the platform decides to enter if:

$$(1-\rho)(\bar{c}_{s_i}-c_a) > (1-\rho)f\,\underline{v} \;\Rightarrow\; c_{s_i} > (c_a + f\,\underline{v})\frac{(1-f)}{(1-f\rho)} \equiv \tau_4'$$

Notice that $\tau_4' = \tau_4/(1-f\rho)$. Hence, it clearly increases in ρ. Provided that $c_{s_i} < c_a < \bar{c}_{s_i}$, the platform cannot try to conquer back the consumers who received an offer on the direct channel of the seller. However, if $c_a < c_{s_i}$ the platform may be tempted to lower its price all the way down to c_{s_i}, out-competing not only the seller on the marketplace, but also the seller on the direct channel. In this case, the platform would earn:

$$\Pi_i^{comp2} = c_{s_i} - c_a$$

Notice that this strategy is profitable when:

$$c_{s_i} - c_a > \max\left\{(1-\rho)(\bar{c}_{s_i}-c_a), (1-\rho)f\,\underline{v}\right\}$$

Formally:

$$c_{s_i} - c_a > (1-\rho)(\bar{c}_{s_i}-c_a) \quad \text{if} \quad c_{s_i} > \frac{c_a(1-f)}{\rho(1-f\,\rho)-f(1-\rho)} \equiv \tau_4(i)$$

$$c_{s_i} - c_a > (1-\rho)f\,\underline{v} \quad \text{if} \quad c_{s_i} > c_a + (1-\rho)f\,\underline{v} \equiv \tau_4(ii)$$

The two thresholds $\tau_4(i)$ and $\tau_4(ii)$ (sorry for the poor creativity) are such that: if $\tau_4(ii) < \tau_4'$, then $\tau_4(i) < \tau_4(ii) < \tau_4'$ and the platform will change from agency directly to intense competition. Otherwise, if $\tau_4(ii) > \tau_4'$, then $\tau_4(i) > \tau_4(ii) > \tau_4'$ and the platform will gradually move from agency if $c_{s_i} < \tau_4'$, to competition if $\tau_4' < c_{s_i} < \tau(i)$, to intense competition if $c_{s_i} > \tau(i)$. The role of ρ is twofold: on the one hand, it reduces the number of market where soft competition is preferred to pure agency; on the other hand, instead, it increases the number of market in which intense competition occurs.

The aggregate profits of the platforms are:

$$\Pi_a = \begin{cases} (1-\rho)\left(\int_0^{\tau_4'} f\underline{v}dc_{s_i} + \int_{\tau_4'}^{\tau_4(i)}\left(c_{s_i}\frac{(1-f\,\rho)}{1-f} - c_a\right)dc_{s_i}\right) + \int_{\tau_4(i)}^{\tau_1'}(c_{s_i}-c_a)\,dc_{s_i} & \text{if}\, \tau_4(ii) > \tau_4' \\ (1-\rho)\int_0^{\tau_4(ii)} f\underline{v}dc_{s_i} + \int_{\tau_4(ii)}^{\tau_1'}(c_{s_i}-c_a)\,dc_{s_i} & \text{otherwise} \end{cases}$$

Discussion. The importance of consumers' contact information is well-known. Sellers that are able to directly communicate (via notifications or emails) with their consumers can disintermediate from the platform ecosystem and bypass the transaction fee. If used strategically, this strategy can prove beneficial for both sellers and consumers, who benefit from a discount.

Interestingly, the platform reacts in a very different way compared to the case in which data do not allow firms to directly trade with consumers. Indeed, following data-sharing, a proportion of users may find it profitable to buy from the direct channel of the seller, to the detriment of the platform. The platform's best reaction is therefore to enter in more markets than it would be optimal otherwise, and outcompete the active sellers both online and offline. By offering a very competitive price on the platform, consumers remain captive of the marketplace and do not disintermediate.

The welfare effect is clearly positive for consumers, who benefit from large discounts. Instead, sellers are collectively worse off if relatively inefficient, as they are systematically outperformed by the more efficient platform. Efficient sellers, instead, thrive, as the platform cannot outcompete them.

A Proofs

TBW

References

Abrardi, L., Cambini, C., Congiu, R., and Pino, F. (2024). User data and endogenous entry in online markets*. *The Journal of Industrial Economics.*

Anderson, S. P. and Bedre-Defolie, Ö. (2021). Hybrid platform model. *CEPR Discussion Paper No. DP16243.*

Aparicio, D., Metzman, Z., and Rigobon, R. (2021). The pricing strategies of online grocery retailers. NBER working paper n. 28639.

Beard, T. R., Kaserman, D. L., and Mayo, J. W. (2001). Regulation, vertical integration and sabotage. *The Journal of Industrial Economics*, 49(3):319–333.

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

Bhargava, H. K. and Choudhary, V. (2008). Research note—when is versioning optimal for information goods? *Management Science*, 54(5):1029–1035.

Bounie, D., Dubus, A., and Waelbroeck, P. (2021). Selling strategic information in digital competitive markets. *The RAND Journal of Economics*, 52(2):283–313.

Bourreau, M. and Gaudin, G. (2022). Streaming platform and strategic recommendation bias. *Journal of Economics & Management Strategy*, 31(1):25–47.

Campbell, J., Goldfarb, A., and Tucker, C. (2015). Privacy regulation and market structure. *Journal of Economics & Management Strategy*, 24(1):47–73.

Casner, B. and Teh, T.-H. (2023). Content-hosting platforms: discovery, membership, or both? *Membership, or Both.*

De Corniere, A. and Taylor, G. (2020). Data and competition: a general framework with applications to mergers, market structure, and privacy policy.

Delbono, F., Reggiani, C., and Sandrini, L. (2021). Strategic data sales to competing firms. Technical Report JRC126568, JRC Digital Economy Working Paper, Seville, Spain.

Dryden, N., Khodjamirian, S., and Padilla, J. (2020). The simple economics of hybrid marketplaces.

Economides, N. (1998). The incentive for non-price discrimination by an input monopolist. *International Journal of Industrial Organization*, 16(3):271–284.

Etro, F. (2021a). Hybrid marketplaces with free entry of sellers. *Available at SSRN.*

Etro, F. (2021b). Product selection in online marketplaces. *Journal of Economics & Management strategy*, 30(3):614–637.

Foerderer, J., Kude, T., Mithas, S., and Heinzl, A. (2018). Does platform owner's entry crowd out innovation? evidence from google photos. *Information Systems Research*, 29(2):444–460.

Gambato, J. and Sandrini, L. (2023). Not as good as it used to be: Do streaming platforms penalize quality? *ZEW-Centre for European Economic Research Discussion Paper 23-032*.

Goldfarb, A. and Tucker, C. (2019). Digital economics. *Journal of Economic Literature*, 57(1):3–43.

Hagiu, A., Teh, T.-H., and Wright, J. (2022). Should platforms be allowed to sell on their own marketplaces? *The RAND Journal of Economics*, 53(2):297–327.

Krämer, J. and Schnurr, D. (2022). Big data and digital markets contestability: Theory of harm and data access remedies. *Journal of Competition Law & Economics*, 18(2):255–322.

Krämer, J. and Shekhar, S. (2022). Regulating algorithmic learning in digital platform ecosystems through data sharing and data siloing: Consequences for innovation and welfare. *Available at SSRN*.

Liu, Z., Zhang, D., and Zhang, F. (2021). Information sharing on retail platforms. *Manufacturing and Service Operations Management*, 23(3):547–730.

Madsen, E. and Vellodi, N. (2022). Insider imitation. *working paper*.

Magnani, M. and Navarra, F. (2023). Information provision in hybrid platforms. *Available at SSRN*.

Mikians, J., Gyarmati, L., Erramilli, V., and Laoutaris, N. (2012). Detecting price and search discrimination on the internet. In *Proceedings of the 11th ACM Workshop on Hot Topics in Networks*, HotNets-XI, pages 79–84, New York, NY, USA. Association for Computing Machinery.

Montes, R., Sand-Zantman, W., and Valletti, T. (2019). The Value of Personal Information in Online Markets with Endogenous Privacy. *Management Science*, 65(3):955–1453.

Navarra, F., Pino, F., and Sandrini, L. (2023). Mandated data sharing in hybrid marketplaces. *Available at SSRN*.

Padilla, J., Perkins, J., and Piccolo, S. (2022). Self-preferencing in markets with vertically integrated gatekeeper platforms. *The Journal of Industrial Economics*, 70(2):371–395.

Pino, F. (2022). The microeconomics of data–a survey. *Journal of Industrial and Business Economics*, 49(3):635–665.

Prüfer, J. and Schottmüller, C. (2021). Competing with big data. *The Journal of Industrial Economics*, 69(4):967–1008.

Raith, M. (1996). A general model of information sharing in oligopoly. *Journal of economic theory*, 71(1):260–288.

Teh, T.-H. (2022). Platform governance. *American Economic Journal: Microeconomics*, 14(3):213–254.

Thisse, J.-F. and Vives, X. (1988). On The Strategic Choice of Spatial Price Policy. *The American Economic Review*, 78(1):122–137.

Wen, W. and Zhu, F. (2019). Threat of platform-owner entry and complementor responses: Evidence from the mobile app market. *Strategic Management Journal*, 40(1).

Zhu, F. and Liu, Q. (2018). Competing with complementors: An empirical look at amazon.com. *Strategic Management Journal*, 39(10):2618–2642.