Towards Efficient Information Sharing in Network Markets

Bertin Martens
Bruegel and Tilburg Law and Economics Center
bertin.martens@gmail.com

Geoffrey Parker
Affiliation Name
geoffrey.g.parker@dartmouth.edu

Georgios Petropoulos
MIT and Stanford University
gpetrop@mit.edu

Marshall Van Alstyne
Boston University
mva@bu.edu

Abstract

This paper develops a Salop differentiation model to study private and social incentives to share information within a platform ecosystem, between a platform intermediary and its business users. Information sharing can help business users to make more efficient decisions around their market and product innovation strategies improving social welfare. However, private and social incentives for information sharing do not coincide when the platform intermediary is vertically integrated and competes directly with its business users in the upstream market of the platform market. So, there is a scope for an ex-ante *regulation of mandatory data sharing. We argue that the location of data access matters and propose a regulatory framework that introduces a new data right for platform users, the* in-situ *data right, which is associated with positive welfare gains. By construction, this right enables effective information sharing, together with its context, without reducing the value created by network effects. We discuss crucial elements of its implementation in order to achieve innovation-friendly and competitive digital markets.*

Keywords: Information sharing, data rights, digital platforms, in-situ, data portability.

1. Introduction

Digital platforms enable efficient interactions between a supply side of producer users and a demand side of consumer users that create value in online and offline trade (Constantinides et al. (2018) and Munger (2015)). They adopt open digital infrastructures that allow multiple stakeholders to interact and governance rules that balance platform control with incentives for users to engage with the platform.

In this way, they do not only invert the production structure of traditional firms (Parker et al. (2017)), but also the information structure of traditional markets by switching from decentralized to centralized market data collection and processing (Martens (2021)). Moving to more centralized information structures has two effects: On the one hand, it increases market efficiencies because new (transaction-based) valuable opportunities emerge from the better management and process of information. On the other hand, it also introduces asymmetric information structures since the managers of information, the digital platforms, in order to protect their market position and the value of their intermediary services, disclose only a small subset of the information they collect and process to their users. In this paper, we show that such asymmetric information structures can have negative implications for platform users and welfare and we propose an efficient mechanism to create more efficient digital platform markets.

Our paper primarily deals with the implications of digital platform markets and their chosen centralised information structures on social welfare contributing to the interesting debate that is currently evolving in the literature (Clemons, Schreieck, et al. (2022), Clemons, Waran, et al. (2022), and Rowe and Markus (2022)). We illustrate that the increasing dominance of big tech platforms can have some adverse effects on social welfare, for example, by limiting competition and, ultimately, innovation due to the asymmetric information structures they adopt.

1.1. The benefits of centralized information structures

Prior to digital platform markets, trade occurred in traditional decentralised markets with prices being

URI: https://hdl.handle.net/10125/106899
978-0-9981331-7-1

the main efficient information signalling system (Hayek (1945)). However, users needed to collect their own information on products and prices. This was costly and resulted in substantial information duplication costs among users. High information collection costs limit the reach of individual market overviews and contribute to asymmetric information between buyers and sellers which makes markets inefficient (Akerlof (1970)).

The arrival of digital technology makes it possible to reach a higher level of information efficiency. Online markets reduce information costs for users because the intermediary platform centralizes all market information. Platforms observe all sellers and product characteristics. Buyers reveal their preferences through their browsing behaviour and transactions. Platforms then use artificial intelligence (AI) algorithms for probabilistic matching of supply and demand through search and advertising channels (Agrawal et al. (2018), Chen and Horton (2016), McAfee et al. (2012), and Varian (2014)). The quantity and quality of data inputs, and the efficiency of algorithmic processing of the data, determine the value of matching services produced by intermediary platforms (Alhassan et al. (2016), Baesens et al. (2016), Cooper et al. (2000), Gregory et al. (2021), Kitchens et al. (2018), and Wixom and Watson (2001)). The completeness of the market overview increases the accuracy and reach of market information, and reduces costs, compared to individual data collection (Byrne et al. (2015) and Steelman et al. (2014)). Once collected, the marginal cost of using this information is very low (Rifkin (2014)). Platforms make this information available to users through two-way matching services, including seller-pushed and paid advertising signals to consumers and buyer-pulled search rankings that display a selection of available products and characteristics. Overall, the reduction in information costs and increase in information accuracy and reach triggers significant real welfare gains for platform users on all sides of the market (Brynjolfsson et al. (2010), Brynjolfsson et al. (2003)). These gains follow the Gale and Shapley (1962) reasoning: in an efficient assigned matching, the seller and the buyer mutually prefer to be matched to each other rather than to other market participants. Network effects, driven by the number of users and data externalities between users, contribute to these welfare gains (Brynjolfsson et al. (2019)).

1.2. The problem of asymmetric information

Despite these information efficiency gains and cost savings, there may be further margins for improvement in the use and distribution of information in platforms. Centralized market information creates a very asymmetric information distribution between the platform and its users (Parker et al. (2022)). This may lead to information market failures. Platforms, as profit-maximizing private firms, use their privileged market overview and exclusive control over market information to their private advantage and not necessarily in a socially optimal way. This may result in biased matching between users (De los Santos and Koulayev (2017)). Platforms do not share all market information with users. They only send narrowly targeted information signals through indirect data-sharing mechanisms such as search rankings and advertising channels. In the absence of a consumer market overview, however, the cost-effectiveness of targeted advertising remains low (Blake et al. (2015) and Farahat and Bailey (2012)). Limited data are too narrow to enable buyers and sellers to endogenously improve their market entry position and maximize producer profits or consumer utility.

So, selective data disclosure in the interest of the platform can result in under-utilization of the data and information frictions that reduce platform market efficiency – to the benefit of the platform but at the expense of overall social welfare across all types of users. Besides, a platform's adopted information structure may primarily be motivated by either minimizing the risk of disintermediation in the interactions of its users or securing a competitive advantage, instead of increasing the efficiency of these transactions. A gap between private and social welfare generated by data access restrictions and inefficient information use would then provide an argument for regulatory intervention in data markets.

1.3. Research findings and objectives

Our first objective is to illustrate the asymmetric information problems identified in the previous subsection with a formal model. We develop a Salop product differentiation model (Salop (1979)) of a platform and three firms - sellers that join the platform to list their products. We compare different information structures. We show that an asymmetric information structure where the platform has all relevant market information but it does not share it with its third-party seller users leads to market inefficiencies. The platform has increased incentives not to share valuable market information especially when it is vertically integrated with one of the sellers. We show that information sharing which allows third-party sellers to get information about demand and supply conditions restores market efficiency

Then, our research question becomes how to achieve a more symmetric information distribution that narrows the gap between the platform and social welfare value of data. A classic solution to reduce information asymmetries and facilitate more efficient use of available platform information is to introduce direct or indirect data sharing between parties. We propose a new solution with an intermediary data sharing modality, the *in-situ* data access right, whereby platform sellers are the rights holders who can bring their algorithms to their data inside the platform where it resides. Under our proposal, data are not separated from their networked context and become directly actionable. When third-party firms gain access to a wider market dataset, including consumer preferences and willingness-to-pay, they can improve their market positioning inside the platform, including in competition with vertically integrated platform services.

2. Model

A market with $n = 3$ horizontally differentiated products is represented by a Salop circle with a circumference of 1. The relevant players are sellers, consumers and a platform. Each product is produced at marginal cost c by a seller. Sellers first decide their location on the circle, then set their respective prices. Consumers are uniformly distributed along the circle; they have unit demand and the market is fully covered. The platform observes both sellers' locations and offers and each consumer's location. It matches sellers with consumers.[1] For its services, the platform charges each seller a share ρ of its profit.

Consumers are not able to directly observe the location of sellers on the circle. So, they rely on the recommendation of the platform, which has superior information on market conditions and locations of all players. The actual distance between sellers represents consumer demand for their products. Consumers located closer to a seller have a higher willingness to pay for the product of that seller because it is closer to their preferences. Distance from that seller reduces willingness to pay. In the terminology of the model, each consumer incurs a transportation cost to arrive at the location of that seller. The higher the distance is, the higher will be the incurred transportation cost to buy from that seller. Let t be the transportation cost per unit of length.

The timing of the static game is as follows:
1. Firms select their locations on the circle.
2. Firms choose their prices based on their chosen locations.
3. Platform provides an individual recommendation to each consumer which firm to visit.
4. Consumers purchase one unit of the good produced from the platform's recommended firm.

We consider three different information regimes on the seller side: In the first (full information), firms are able to observe the exact location of their competitors on the circle (this can be achieved when data access rights are in place). In the second (no market-wide information), sellers enter the market without information on firms' market positions - a realistic situation in online platform markets. Sellers are "blind" in the sense that they are not able to observe their competitors' locations in order to assess their distance from them on the circle. Last but not least, we consider the case that one firm is vertically integrated with the platform that operates the market and has access to all relevant information about market conditions and market players. As a result, this firm can observe the locations of its competitors, which are still "blind". Hence, the vertically integrated firm has an information advantage over its competitors.

Our objective is to compare the consumer welfare and sellers' profitability among these regimes in order to conclude the information structure that leads to the highest welfare of the platform's participants.

3. Analysis

Full information:

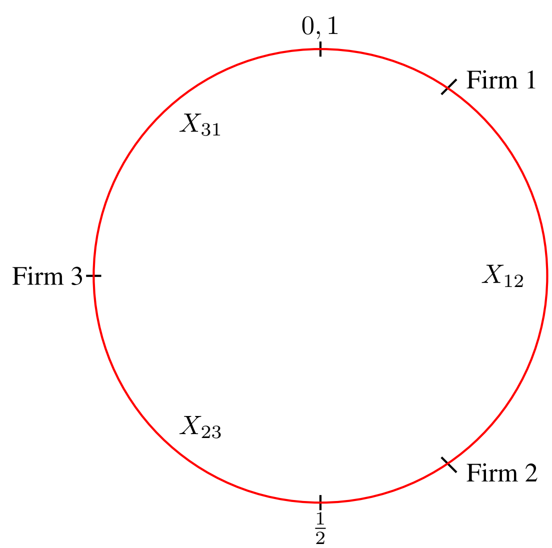

Figure 1. The Salop circle.

Firms are able to observe the exact location of their competitors on the circle. Let X_{12}, X_{23} and X_{31} be the

[1] We can interpret this as follows: Consumers submit a precise product query to the platform, which can then infer (based on the profile of each consumer and the available products by sellers that are relevant to the query) what the best match between supply and demand will be. The platform then provides advice to the consumer.

distances between firms 1 and 2, firms 2 and 3, and firms 3 and 1, respectively. This implies $X_{12}+X_{23}+X_{31}=1$.

We solve the game using backward induction to derive the equilibrium prices and locations for each of the three firms. Since firms observe their in-between distance, the platform provides consumer recommendations that are consistent with sellers' demand expectations. Hence, each firm i faces demand D_i which is given by:

$$D_i(p_i, p_j, p_k; X_{ij}, X_{ki}) = \frac{p_j + p_k - 2p_i}{2t} + \frac{X_{ij} + X_{ki}}{2},$$

where, p_i is the price set by each firm i, where $i, j, k = 1, 2, 3$, with $j \neq k \neq i$. In the second stage of the game, each firm i maximizes its profit $\Pi_i = (p_i - c)D_i$, with respect to its price p_i. This leads to the equilibrium profit as function of X_{ij} and X_{ki}:

$$\Pi_i = \frac{t(1 + X_{ij} + X_{ki})^2}{25}.$$

In the first stage, each firm i simultaneously chooses its location based on its expectation of the locations of firms j and k. So, essentially, each firm i chooses its distance from its two competitors in order to maximize Π_i. We focus on the symmetric Bayesian Nash equilibrium of this location game. In equilibrium, firms choose locations such that

$$X_{12}^* = X_{23}^* = X_{31}^* = \frac{1}{3}.$$

The respective equilibrium prices and profits are:

$$p_1^* = p_2^* = p_3^* = p^* = c + \frac{t}{3}$$

and

$$\Pi_1^* = \Pi_2^* = \Pi_3^* = \frac{t}{9}.$$

In fact, this equilibrium involves locations of firms that minimize the average transportation cost for consumers. To see this, let $\alpha, \beta, \gamma \in \mathbb{R}$ such that the distance between the three firms is respectively, $X_{12}^*+\alpha$, $X_{23}^*+\beta$ and $X_{31}^*+\gamma$. We want to find the real numbers α, β, γ which minimize the average transportation cost in the equilibrium:

$$\min_{\alpha,\beta,\gamma} \left\{ 2t \left(\int_0^{\frac{1}{6}+\frac{\alpha}{2}} y dy + \int_0^{\frac{1}{6}+\frac{\beta}{2}} y dy + \int_0^{\frac{1}{6}+\frac{\gamma}{2}} y dy \right) \right\}.$$

The minimum of this objective function corresponds to $\alpha = \beta = \gamma = 0$.

Lemma 1. *When $n = 3$ firms compete in a Salop circle under full information, in equilibrium, they choose equidistant locations and set price $c + \frac{t}{3}$. This equilibrium minimizes the average transportation cost incurred by consumers, maximizes seller profits, and results in highest welfare.*

So, under full information, there is no market failure in the choice of position, since we achieve both the maximization of sellers' profits and the minimization of consumers' average transportation costs.

No market-wide information: Before stage 1, in the beginning of the game, the location of each firm is drawn identically and independently from a uniform distribution over the support $[0, 1]$ which is common knowledge. Following their expectations, at the first stage, each firm decides its location on the circle and in the second stage its price as before. At the symmetric equilibrium, firms choose their locations such that $\mathbb{E}[X_{12}] = \mathbb{E}[X_{23}] = \mathbb{E}[X_{31}] = \frac{1}{3}$, and the equilibrium price of the second stage is again $p_1 = p_2 = p_3 = c + \frac{t}{3}$. This is expected to be a suboptimal equilibrium, however, because in general, it is highly unlikely that firms' expectations over their respective distances will exactly meet market conditions. It is $|X_{ij} - E[X_{ij}]| \geq 0$, for every $i, j = 1, 2, 3$, with $i \neq j$, and with strict inequality holding most often. Some sellers may actually end up in a location further than $\frac{1}{3}$ away from their nearest competitor but get squeezed on the other side by a very close competitor at less than $\frac{1}{3}$ distance away. While consumers at the long end of the market will have a low willingness to pay and consumers at the short end of the market will have a higher willingness to pay, sellers are not able to set their prices optimally. They need wider market information in order to select a more optimal location at $\frac{1}{3}$ distance from their competitors. Without that market information, they cannot form reliable expectations about the consequences of moving along the Salop circle.

This also implies that the real numbers $\alpha, \beta, \gamma \neq 0$, and therefore the average transportation cost incurred by consumers increases. This equilibrium is not efficient from a welfare point of view because the consumer welfare and seller profits are both smaller than the full information case.

If firms were to receive the necessary updated information through an information-sharing mechanism, they would be able to observe their market locations and their competitors' distances and then adjust their positions and prices to both maximize their profits and minimize consumers' expected transportation costs.

Vertical integration: One firm (say, firm 3) is vertically integrated with the platform. It is fully informed about market conditions, while the other two firms cannot observe market conditions and they only expect to be at the optimal distance of $\frac{1}{3}$ from their competitors and set price $p^* = p_1^{ni} = p_2^{ni} = c + \frac{t}{3}$. Firm 3 now chooses its optimal location and price so as to maximize the payoff from its information advantage.[2] Moreover, the platform is now biased towards its own subsidiary in the sense that it can profit by directing more consumers to firm 3, even if consumers benefit by visiting one of the other two firms. Absent bias, let y be the indifferent consumer between firm 3 and firm i, where $i = 1, 2$. Given bias, let $b > 0$ be the measure of biased recommendation in market share. In stage 3, an unbiased intermediary would have recommended to all consumers located between firm 3 and y to visit firm 3. Nevertheless, under bias, the platform recommends to $y + b$ consumers that they visit firm 3 at the expense of its two competitors.[3] Firm 3's demand function is

$$D_3(p, p^*; X_{12}, b) = \frac{c + \frac{t}{3} - p}{t} + \frac{1 - X_{12}}{2} + 2b.$$

The integrated firm selects price p so that it maximizes its profit $(p - c)D_3(p, p^*; X_{12}, b)$. Solving this maximization problem, we find firm 3's equilibrium profit:

$$\Pi^{vi} = t\left(\frac{1}{6} + \frac{1 - X_{12}}{4} + b\right)^2.$$

Due to the fact that firms 1 and 2 are symmetric, at the first stage, the integrated firm selects its optimal location such that $X_{23} = X_{31} = \frac{1-X_{12}}{2}$. Hence, the respective profits of the two uninformed firms with price p^* become:[4]

$$\Pi_1^{nvi} = \Pi_2^{nvi} = \frac{t}{3}\left(\frac{7}{24} + \frac{X_{12}}{8} - \frac{b}{2}\right).$$

The profit of firm 3 is increasing in platform's bias b. The vertically integrated platform has incentives to adopt a high b in order to reduce the share of consumers that visit firms 1 and 2 below X_{12} and below $\frac{1}{3}$, which is the social optimum, in order to increase its profitability.

If bias b were observable, then regulation could target the platform's tendency to steer business toward itself. Importantly, however, a regulation that targets steering alone would not restore market efficiency. It would reduce bias but it would not provide information sufficient for firms to re-position themselves or, in effect, to redesign their products. More information, including competitors' locations, is also required.

Social and private incentives deviate from each other, harming consumer welfare and independent producers. Compared to $n = 2$, platform entry as firm 3 does lower prices and transportation costs yet is not socially optimal. Indeed, uninformed sellers are worse off under vertical integration with high bias b ($> \frac{X_{12}}{4} - \frac{1}{12}$), in comparison to the full information case, since $\Pi_i^{nvi} < \Pi_i^*$, for all $i = 1, 2$. Consumers are worse off as well since $p^{vi} > p^*$ and their average transportation cost is higher due to bias.

Proposition 1. *When $n = 3$ firms compete in a Salop circle and one is vertically integrated with the platform that operates the market, the platform has the incentive to squeeze independent sellers by reducing the share of consumers directed to them. The solution differs from the socially optimal solution, which is achieved under full information. Independent sellers and consumers are worse off under vertical integration.*

Since full information means that independent sellers can observe their distance from competitors, they are also in a position to observe whether the platform is biased or not and react to that accordingly. They can take proper legal action or inform authorities about platform wrongdoing. Under asymmetric information, however, blind sellers do know if low sales result from bias or from proximity. Vertical integration is more concerning when independent sellers are uninformed and cannot observe platform bias. Improving information sharing in order to achieve a symmetric information structure is one means of resolving this information market failure.

Corollary 1. *Information sharing increases consumer welfare. When sellers are symmetrically informed about market conditions and their positions in it, consumer surplus is maximized.*

If improving information symmetry is a platform choice, then sharing information with non-integrated sellers is not incentive-compatible. This incompatibility creates scope for a regulatory intervention that facilitates effective and symmetric information sharing, in order to arrive at a full information regime.

[2] In the case of vertical mergers and acquisitions in which big platforms are involved, this shift to a new location on the circle can be achieved through the additional functionalities of the merged entity and allows it to provide valuable services to users of different preference profiles (Parker et al. (2021)). For example, after its acquisition, Hotmail offered its users new complementary services based on Microsoft websites. Whole Foods Market was integrated into the Amazon distribution network offering new options to shoppers whose preference was to purchase groceries online.

[3] As consumers are unable to observe seller locations, they also do not observe the platform's bias.

[4] Note that demand for each firm $i = 1, 2$ in equilibrium will be $D_i(p^*, p^{vi}; X_{12}, b) = \frac{p^{vi} - p^*}{2t} + \frac{X_{12} + X_{i3}}{2} - b$. Therefore, its profit is given by $\Pi_i^{nvi} = \frac{t}{3}\left(\frac{p^{vi} - p^*}{2t} + \frac{X_{12} + X_{i3}}{2} - b\right)$.

4. From theory to practice and the *in-situ* right

The above analysis illustrates that firms that use digital platforms should be able to access market-wide information that allows them to make informed decisions about their market strategies in order to ensure efficiency. Existing regulatory intervention and mandatory data access rights are limited in scope and do not really address the information asymmetry between platforms and their users.

In the EU, which has the most developed data access regulation, data access and portability rights for firms operating through intermediary platforms have recently come into focus. The EU Digital Markets Act (DMA) is the first regulation to introduce a data portability right for firms provided they operate through very large gatekeeper platforms (European Union, 2022).

Article 6 and paragraph 10 of the DMA includes the following obligation for the big platforms: ”The gatekeeper shall provide business users and third parties authorised 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.”

This data portability right is an *ex-situ* access right: data are transferred directly from a platform to the server of the rights holder (business user). These are zero-degree interaction data in the sense that firms only get access to their own interaction data generated through the platform with other users (e.g., own sales, customer visits) but do not provide market-wide information that according to our analysis above is needed in order to achieve market efficiency (Krämer et al., 2020). Data portability does not provide sufficient context to fully optimize decisions. In effect, by limiting access to a firm's own data, it strips away information that incorporates a high network value. Data loss reduces the information available for making decisions, reducing market efficiency. Hence, the data portability right is, in other words, equivalent to the no market-wide information case presented above.

The full information case can be achieved if, in addition to own sales and customer visits (zero-degree data), each firm can access (Figure 2)

- first-degree interaction data: online behavior (such as online searches, views and transactions) of consumers who interacted with the firm's products as well as with products from competing firms,
- second-degree interaction data: online behavior of consumers who interacted with products from competing firms which were also looked at by consumers who interacted with the firm

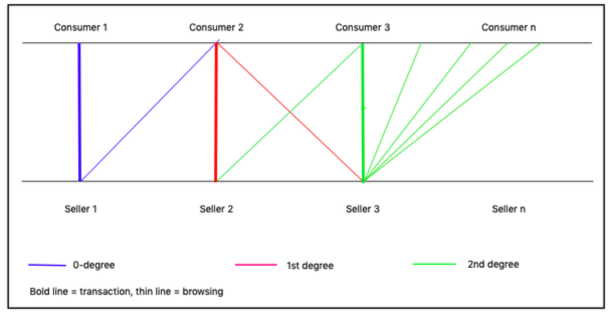

Figure 2. Degrees of platform interaction data. In Situ data rights enable context visibility beyond network degree-0 (oneself) for which negotiation or legislation can provide higher degree access.

In other words, access to first-degree interaction data provides information that allows the firm to observe competitor firms as revealed by consumer behavior, while access to second-degree interaction data provides information about competitors' sales and market shares. Under these additional information sets, sellers can observe their distance from competitors in the Salop circle, and (re-)position themselves in the platform market (adjusting their price accordingly) making the platform market more efficient.

To achieve the full information state of the Salop model, we propose the introduction of a new access right to a wider set of data, from 0-degree to 2nd degree networked data. This is a right to *in-situ* access to this wider dataset. The *in-situ* access right as an intermediate modality, in between direct and indirect data transfers. Data remains at the location where it is collected. Instead of porting data to another location as an input in an algorithm, the algorithms are transferred to the platform's infrastructure where the data is located, and data analysis is performed *in-situ*.

The switch from *ex-situ* to *in-situ* access rights is necessary in order to overcome information asymmetry between firms and the platform and enable a firm to access the data required to estimate consumer demand and willingness to pay for its own and for competing products. To achieve this, the scope of the *in-situ* access right for a firm has to reach further than the standard *ex-situ* data portability right of the DMA that would normally only include information on direct transactions between a firm and its consumers.

The *in-situ* right enables firms to assess consumer preferences and willingness to pay for a range of close-substitute competing products. This would not only enable each firm to estimate the distance to competitors. but it may also enable firms to identify gaps and niche markets where innovative entry could be possible. If business users operate on several competing platforms, they will be able to combine insights gained across platforms. Even if platforms differ in the demand and supply conditions, business users can develop an oversight across all platform markets they participate in, in order to better assess aggregate market trends and adopt the appropriate business strategies.

Second-degree interaction data is unlikely to give a complete market overview, especially not for small businesses in the long tail of the sales distribution that have fewer consumers. However, if any of their consumers would directly interact with competing superstar sellers (seller 3 in Figure 2), second-degree data access allows these small sellers to "hop" to a much wider market overview and enables them to compete better with superstar sellers. On the other hand, second-degree access would have to be limited to a market definition that does not extend beyond well-defined product categories or close-substitute products. Business users who sell smartphones cannot claim access to TV market data, even though their consumers may have interacted with TV sellers. Limiting the scope of *in-situ* access rights to first and second-degree interactions implies that not all platform data are shared with its business users. The platform still has an information advantage, though more limited than with *ex-situ* zero-degree portability rights.

It is important to note that the *in-situ* access right for sellers does not involve access to consumers' personal data. Access to first- and second-degree interaction data should be viewed as sufficient statistics that allow sellers to get information on consumers' preferences over alternative options within a given product category in order to assess their products' degree of substitutability and their specific location on the Salop circle. But, the identity of each of these consumers is never revealed. Sellers, by exercising their *in-situ* access right, get access to the general characteristics of those consumers who bought from each firm (up to second-degree), as well as firm/product profile data (product characteristics and prices, sales volumes, characteristics and browsing behavior of buyers, etc.) in a way that consumers remain anonymous.

Such information sharing can only occur *in-situ*. If such multidimensional information were shared *ex-situ*, the risk to de-anonymize consumers' data would have been much higher (De Montjoye et al. (2015)), violating in this way data privacy rules. When data is released *ex-situ*, the data receivers (e.g., sellers) can do whatever computations they want with the data. *In-situ*, it is much easier to control computations and restrict information output to aggregate and derived data while protecting the detailed input data. Besides, there are new methods that can be applied so that with data aggregation the value from the input data will not be reduced. For example, Google has proposed to aggregate personal data in consumer cohorts for advertising profiling purposes (Federated Learning of Cohorts, FLoC), without significant loss of advertising value. Other methods of differential privacy as well as the secure multiparty computation approach can be helpful as well.

As a result, an important restriction on the scope of *in-situ* data access rights for sellers is that platforms retain exclusive direct access to consumers' personal information and profiles, including direct write access to consumers via search and paid advertising channels. If firms would have directed write access to all consumers they could circumvent the platform and cause a substantial degree of "leakage" that may erode the economic viability of the platform and undermine the positive network externalities that it generates.

In-situ access for firms will also mitigate the problems from vertical integration of the platform with upstream providers, identified in the analysis of the model, in terms of upstream competition and level playing field. This is because platforms will be less able to take advantage of information asymmetry for the benefit of their own upstream subsidiary. Third-party firms can use their increased market insights to better fend off competition inside the platform from vertically integrated platform services. In this way, the benefits from information sharing expand not only horizontally, between competing firms, but also vertically, across the value chain.

However, *in-situ* access will not result in a complete information symmetry between a platform and its business users. Platforms retain important information advantages, notably in terms of access to personal data and an overall market overview. Business users cannot replicate the entire platform dataset. Preserving part of the information asymmetry is important in order not to undermine the economic viability of the platform as a central intermediary. Putting all platform information into the public domain would undermine the commercial viability of the platform and put at risk the positive social network effects that it generates for its users. Exclusive control over at least part of its data pool facilitates monetization of the data through the twin advertising and search communication channels.

So far, we implicitly assumed that firms that have an *in-situ* data access right in a platform also have the necessary information processing capacity and skills to develop and run algorithms that can harvest data-driven market insights in these platforms. This is not necessarily true for some small business users on a platform. They may require help from a third-party data analytics provider, or the platform itself may provide this service. In the latter case, the platform would have a privileged insight into the strategies of its business users. Third-party analytics would avoid that problem, especially when encrypted *in-situ* processing technologies are used. Platforms will no longer be in a position to charge a monopolistic price for such services since they can be performed by many competing providers. They may still be able to charge a monopolistic price for *in-situ* use of the platform server infrastructure. This should be regulated as part of the *in-situ* access right.

Since the *in-situ* data right corresponds to the full information and the data portability right corresponds to the no market-wide information cases analyzed above, we conclude that: i) As measured by social welfare, a policy banning vertical integration is dominated by a policy of *in-situ* data rights; ii) A policy of data portability is strictly dominated by a policy of *in-situ* data rights.

5. Potential risks and ways to address them

The *in-situ* proposal can incorporate significant benefits for the users of platform markets but there are some potential risks that should not be overlooked. The first one concerns data privacy violations. Even if *in-situ* outperforms *ex-situ* data access with respect to privacy protection and firms get access to anonymized, aggregated data, it may be the case that the risk of de-anonymization is not negligible if relevant datasets are not aggregated at a sufficient level. So, it is important to minimize this risk when we decide about the type of networked data that will be shared and the degree of its aggregation when we are in the phase of the implementation of this mechanism.

At the same time, not all business users of platforms should be considered a priory as benevolent. There may be some firms that want to use the increased market information they get through the *in-situ* mechanism for malicious purposes and illegal activities. It is therefore important to safeguard this right with a certification mechanism for the right holders. Firms that want to participate in the mechanism and access additional information in platform markets should satisfy some security criteria of good market conduct and intentions to minimize the risk of observing adverse market effects.

Since, as we discussed, private incentives for information sharing do not coincide with social ones, platforms may try to find ways to limit the information they provide to their users. It is therefore important for the regulation that will enforce the *in-situ* mechanism to provide concrete and precise obligations over the responsibilities of the platforms when they share information with their users. Imposing some compatibility standards on the way information is transmitted and how data is accessed by users' algorithms is a good start in this direction.

The fact that business users can use the *in-situ* right to get access to more information about competitors' market performance may increase the risk of some form of price collusion with detrimental effects on consumers. It is therefore important to minimize this risk by excluding from the information shared key parameters as the price competitors charge for their products. Demand characteristics and price elasticity are also important factors that might mitigate this risk. Competition authorities can also be more effective in fighting price-fixing explicit agreements because under *in-situ* market transparency achieves higher standards and makes it less costly to intervene.

6. Conclusions

Online platforms centralize large data sets regarding the behaviors of users on all sides of their markets. This allows them to efficiently match users and generate stronger welfare-enhancing network externalities as compared to traditional decentralized offline markets where users must collect their own market information. Policymakers have focused on the competition implications of network effects and 'market tipping' in online platforms mostly by considering ownership structure. In this paper, we turn our attention to market failure driven by information asymmetry between platforms and their users. Without sufficient access to market data, merchants only receive small excerpts from the pool of comprehensive market information collected by the platform.

The lack of comprehensive market information creates obstacles for merchants in selecting an optimal market entry point into the platform. This reduces business revenue and consumer welfare and reduces market entry for innovative products.

To overcome these problems, we propose that regulators introduce a new *in-situ* data access right, both for firms and consumers, with a wider scope that goes beyond “own” data and includes a sufficient degree of networked context to enable the reconstructions of a market overview. Instead of porting data off-platform, the *in-situ* mechanism requires digital platforms to open their infrastructures and allow consumers and merchants to bring algorithms to their data. Data is accessed at the location where it is collected, preserving the option value of context.

The European Parliament, in its proposed DMA amendments (see specifically Amendment 17) considers the *in-situ* data rights as a viable and effective policy option for the future of big platform regulation (European Parliament (2020)): ”Gatekeepers should also facilitate access to these data in real time by means of appropriate technical measures, such as for example putting in place high quality application programming interfaces **or enabling access to data by the business user “in situ”, without a transfer by the gatekeeper**.”

While empirical evidence about the implications of such an information-sharing model are scarce, they are promising for the efficiency gains it incorporates. The open algorithms project (OPAL (2019)) is a case study that follows closely the basic principles of the *in-situ* mechanism. The real-world deployment of the project started in mid-2017 in Colombia and Senegal. It managed to unlock the potential of data collected by private organizations in these places which led the participating organizations to make more informed and better decisions ”in support of the sustainable development of their goals”. Furthermore, a second relevant case study is the open banking initiatives in the EU and the UK. These initiatives enable small fintech corporations to access data from big bank institutions *in-situ*. Even if it is too early to thoroughly evaluate the implications of these regulations, we have already seen some benefits with respect to market efficiency and innovation (Ozcan and Zachariadis (2021)).

References

Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction machines: The simple economics of artificial intelligence*. Harvard Business Press.

Akerlof, G. A. (1970). The market for “lemons”: Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, *84*(3), 488–500.

Alhassan, I., Sammon, D., & Daly, M. (2016). Data governance activities: An analysis of the literature. *Journal of Decision Systems*, *25*(sup1), 64–75.

Baesens, B., Bapna, R., Marsden, J. R., Vanthienen, J., & Zhao, J. L. (2016). Transformational issues of big data and analytics in networked business. *MIS quarterly*, *40*(4), 807–818.

Blake, T., Nosko, C., & Tadelis, S. (2015). Consumer heterogeneity and paid search effectiveness: A large-scale field experiment. *Econometrica*, *83*(1), 155–174.

Brynjolfsson, E., Collis, A., & Eggers, F. (2019). Using massive online choice experiments to measure changes in well-being. *Proceedings of the National Academy of Sciences*, *116*(15), 7250–7255.

Brynjolfsson, E., Dick, A. A., & Smith, M. D. (2010). A nearly perfect market? *QME*, *8*(1), 1–33.

Brynjolfsson, E., Hu, Y., & Smith, M. D. (2003). Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. *Management science*, *49*(11), 1580–1596.

Byrne, D. M., et al. (2015). *Prices for data storage equipment and the state of it innovation* (tech. rep.). Board of Governors of the Federal Reserve System (US).

Chen, D. L., & Horton, J. J. (2016). Research note—are online labor markets spot markets for tasks? a field experiment on the behavioral response to wage cuts. *Information Systems Research*, *27*(2), 403–423.

Clemons, E. K., Schreieck, M., Krcmar, H., & Bui, T. (2022). Social welfare computing and the management and regulation of new online business models. *Electronic Markets*, *32*(2), 411–414.

Clemons, E. K., Waran, R. V., Hermes, S., Schreieck, M., & Krcmar, H. (2022). Computing and social welfare: Minimizing societal harm from digital transformation while preserving the benefits of innovation in online businesses. *Electronic Markets*, *32*(2), 417–436.

Constantinides, P., Henfridsson, O., & Parker, G. (2018). Introduction—platforms and infrastructures in the digital age.

Cooper, B. L., Watson, H. J., Wixom, B. H., & Goodhue, D. L. (2000). Data warehousing supports corporate strategy at first american corporation. *MIS quarterly*, 547–567.

De los Santos, B., & Koulayev, S. (2017). Optimizing click-through in online rankings with endogenous search refinement. *Marketing Science*, *36*(4), 542–564.

De Montjoye, Y.-A., Radaelli, L., Singh, V. K., & Pentland, A. “ (2015). Unique in the shopping

mall: On the reidentifiability of credit card metadata. *Science*, *347*(6221), 536–539.
European Parliament. (2020). *Draft report on the proposal for a regulation of the european parliament and of the council contestable and fair markets in the digital sector (digital markets act)* (tech. rep.) [COM(2020)0842 – C9-0419/2020 – 2020/0374(COD), p. 1-82]. European Parliament. https://www.europarl.europa.eu/doceo/document/IMCO-PR-692792_EN.pdf
European Union. (2022). *Regulation (eu) 2022/1925 of the european parliament and of the council of 14 september 2022 on contestable and fair markets in the digital sector and amending directives (eu) 2019/1937 and (eu) 2020/1828 (digital markets act)* (tech. rep.) [OJ L 265, 12.10.2022, p. 1–66]. European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R1925
Farahat, A., & Bailey, M. C. (2012). How effective is targeted advertising? *Proceedings of the 21st international conference on World Wide Web*, 111–120.
Gale, D., & Shapley, L. S. (1962). College admissions and the stability of marriage. *The American Mathematical Monthly*, *69*(1), 9–15.
Gregory, R. W., Henfridsson, O., Kaganer, E., & Kyriakou, H. (2021). The role of artificial intelligence and data network effects for creating user value. *Academy of management review*, *46*(3), 534–551.
Hayek, F. A. (1945). The use of knowledge in society. *The American economic review*, *35*(4), 519–530.
Kitchens, B., Dobolyi, D., Li, J., & Abbasi, A. (2018). Advanced customer analytics: Strategic value through integration of relationship-oriented big data. *Journal of Management Information Systems*, *35*(2), 540–574.
Krämer, J., Senellart, P., & de Streel, A. (2020). *Making data portability more effective for the digital economy: Economic implications and regulatory challenges*. Centre on Regulation in Europe asbl (CERRE).
Martens, B. (2021). Data access, consumer interests and social welfare–an economic perspective on data. *Data Access, Consumer Interests and Public Welfare*, 69–102.
McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D., & Barton, D. (2012). Big data: The management revolution. *Harvard business review*, *90*(10), 60–68.
Munger, M. (2015). Coase and the ‘sharing economy’. In C. Veljanofski (Ed.), *Forever contemporary: The economics of ronald coase* (pp. 187–208). Institute of Economic Affairs.
OPAL. (2019). *Open algorithms project case study* (tech. rep.). https://static1.squarespace.com/static/599ef170197aeac586fed53f/t/5c182ccd8a922daa605547b1/1545088208773/OPAL+Unlocking+Private+Sector+Data_Case+Study.pdf
Ozcan, P., & Zachariadis, M. (2021). Open banking as a catalyst for industry transformation: Lessons learned from implementing psd2 in europe.
Parker, G., Petropoulos, G., & Van Alstyne, M. W. (2021). Platform mergers and antitrust. *Industrial and Corporate Change*, *00*, 1–30.
Parker, G., Petropoulos, G., & Van Alstyne, M. W. (2022). Digital platforms and antitrust. *Oxford Handbook of Institutions of Economic Governance and Market Regulation*.
Parker, G., Van Alstyne, M. W., & Jiang, X. (2017). Platform ecosystems: How developers invert the firm. *MIS Quarterly*, *41*(1), 255–266.
Rifkin, J. (2014). *The zero marginal cost society: The internet of things, the collaborative commons, and the eclipse of capitalism*. St. Martin’s Press.
Rowe, F., & Markus, M. L. (2022). Taking the measure of digital giants: Amazon and the social welfare computing research agenda. *Electronic Markets*, *32*(2), 437–446.
Salop, S. C. (1979). Monopolistic competition with outside goods. *The Bell Journal of Economics*, 141–156.
Steelman, Z. R., Hammer, B. I., & Limayem, M. (2014). Data collection in the digital age. *MIS quarterly*, *38*(2), 355–378.
Varian, H. R. (2014). Beyond big data. *Business Economics*, *49*(1), 27–31.
Wixom, B. H., & Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success. *MIS quarterly*, 17–41.