WORKING PAPER | ISSUE 12/2021 | 10 NOVEMBER 2021

TOWARDS EFFICIENT INFORMATION SHARING IN NETWORK MARKETS

BERTIN MARTENS, GEOFFREY PARKER, GEORGIOS PETROPOULOS AND MARSHALL VAN ALSTYNE

Digital platforms facilitate interactions between consumers and merchants that allow collection of profiling information, which drives innovation and welfare. Private incentives, however, lead to information asymmetries resulting in market failures both on-platform, among merchants, and off-platform, among competing platforms. This paper develops two product-differentiation models to study private and social incentives to share information within and between platforms. We show that there is scope for *ex-ante* regulation of mandatory data sharing that improves social welfare better than competing interventions, such as barring entry, break-up, forced divestiture or limiting recommendation steering. These alternate proposals do not make efficient use of information. We argue that the location of data access matters and develop 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 sharing of information together with its context, without reducing the value created by network effects. It also enables regulatory oversight but limits data privacy leakages. We discuss crucial elements of its implementation in order to achieve innovation-friendly and competitive digital markets.

Keywords: Information sharing, data rights, digital platforms, market failure, in-situ, algorithms, data portability.
JEL Classi ication: D47, D82, K21, L21, L22, L40, L41, L43, L51, L86

Bertin Martens (Bertin.Martens@ec.europa.eu) is a Senior Economist at the European Commission Joint Research Centre
Geoffrey Parker (Geoffrey.G.Parker@dartmouth.edu) is a Professor of Engineering at Dartmouth College
Georgios Petropoulos (georgios.petropoulos@bruegel.org) is a Research Fellow at Bruegel and an affiliate at MIT Sloan School of Management
Marshall Van Alstyne (mva@bu.edu) is a professor at Questrom School of Business (Boston University)

Our paper has benefitted from inspiring discussions with Erik Brynjolfsson, Luis Cabral, Rebecca Christie, Maria Demertzis, Erika Douglas, Nestor Duch-Brown, Justus Haucap, Jan Krämer, Maciej Sobolewski, Sebastian Steffen, Tommaso Valletti, Reinhilde Veugelers, Guntram Wolff as well as participants at Ascola 2021, Yale University's Big Tech and Antitrust Conference 2020, OECD Competition Committee Hearing Dec. 2020, Bruegel, Digital Markets Competition Forum at Copenhagen Business School 16 June 2021, and the Boston University Platform Summit 27 May 2021. Georgios Petropoulos gratefully acknowledges financial support from the European Union's Horizon 2020 research and innovation programme under Marie Sklodowska-Curie grant No. 799093. Bertin Martens acknowledges that the views and opinions expressed in this paper do not necessarily reflect those of the Joint Research Centre or the European Commission.

bruegel

Recommended citation:
Martens, B., G. Parker, G. Petropoulos and M. Van Alstyne (2021) 'Towards efficient information sharing in network markets', *Working Paper* 12/2021, Bruegel

1 Introduction

Digital platforms enable efficient interactions between a supply side of business users, which includes external producers, content providers developers and a demand side with consumers that lead to value creation in online and offline trade (Constantinides et al., 2018; Parker et al., 2016). 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.

Successful platform business models depend on the ability of online intermediaries to facilitate valuable interactions and transactions between agents. While they invert the production structure of traditional firms (Parker et al., 2017; Parker and Van Alstyne, 2018), they also invert the information structure of traditional markets by switching from decentralised to centralised market data collection and processing.[1]

Hayek (1945) argued that prices are an efficient information signalling system in traditional decentralised markets. Still, users need to collect their own information on products and prices. This is costly and results 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. As Akerlof (1970) illustrates with his "lemons" market, asymmetric information reduces the efficacy of price as a signalling instrument and leads to market failures that can even eliminate trade between buyers and sellers.

The arrival of digital technology makes it possible to reach a higher level of information efficiency.[2] 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. 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. The completeness of the market overview increases the accuracy and reach of market information, and reduces costs, compared to individual data collection. Once collected, the marginal cost of using this information is very low. Platforms make this information available to users through two-way matching

[1]The ability to collect and store data and turn it into valuable information rapidly increased over the last two decades. The digitisation of nearly all media and the increasing migration of economic and social activities to the internet generate petabytes of data every second (Hill et al., 2015). At the same time, the cost for storing (and analyzing) data has sharply fallen (Byrne et al., 2015).

[2]On the efficiency of centralized marketplaces, see the 5 relevant criteria in Roth (2008).

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., 2003, 2010; Manyika et al., 2011).[3] Network effects, driven by the number of users and data externalities between users, contribute to these welfare gains (Brynjolfsson et al., 2019).

Despite these information efficiency gains and cost savings there may be further margins for improvement in the use and distribution of information in platforms. Centralised market information creates a very asymmetric information distribution between the platform and its users (Parker et al., forthcoming). 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 send only narrowly targeted information signals through indirect data sharing mechanisms such as the search rankings and advertising channels.[4] Limited data are too narrow to enable buyers and sellers to endogenously improve their market entry position and maximize producer profits or consumer utility.

Selective disclosure in the interest of the platform may 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. For example, a platform's adopted information structure may primarily be motivated by either minimizing the risk of dis-intermediation 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.

On the other hand, platforms need to maintain some degree of information asymmetry and drive an information wedge between them and their users in order to monetise their data advantage and avoid falling into the Arrow paradox (Bergemann and Bonatti, 2019): potential buyers do not know the value of the information until the see it but having seen it they no longer need to pay for it. Full public disclosure of all information collected by the platform would undermine data monetisation strategies and weaken incentives to collect data. Data can only be monetised when they have scarcity value, for instance through partial

[3]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.

[4]In the absence of a consumer market overview, however, the cost effectiveness of targeted advertising remains low (Blake et al., 2015).

release of samples, aggregated or coarse-grained data. User demand for privacy protection imposes restrictions on the release of personal and commercially valuable user data. Data releases outside the infrastructure controlled by the platform weakens these controls. Full data disclosure may also overwhelm users with limited data processing capacities. The platform's data analytics and filtering services save data processing costs for users.

Our primary research question is how to achieve a more symmetric information distribution that narrows the gap between the private 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 rights holders can bring algorithms to their data inside the platform where it resides, as opposed to existing *ex-situ* data portability where rights holders remove their own data, transferring it to another platform. Under our proposal, data are not separated from their networked context and become directly actionable. When partner 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. *In-situ* access allows consumers to invite firms inside and outside the platform to access their interaction data in the context of other products and other consumers, create value, and propose competing offers.

An *in-situ* data access right would give firms and consumers access to a much wider range of contextual interaction data, compared to existing data portability rights that only allow access to a user's own data. Portability requires a transfer of own data to another site where it loses social interaction context with firms and consumers. That reduces the social value of ported data. Hence, this approach still leaves in place a considerable degree of information asymmetry between the platform and its users. Compared to *ex-situ* data portability rights, *in-situ* access prevents data leakage and reinforces data privacy and security which is all the more important with a second-degree expanded access right. The data holder keeps exclusive control of his data. Traditionally, data economics considers only two modes of data sharing: either direct whereby data are transferred from one party to another, like the data portability right of the EU's General Data Protection Regulation (GDPR), or indirect whereby only a data-based service is delivered but no data (e.g., online advertising, online search). *In-situ* access is an intermediate modality: no data transfer but still direct access.

The *in-situ* data access right implies an obligation for platforms to open their infrastructure to imported algorithms and share a larger part of their information with users, compared to information available in consumer-pulled search rankings and seller-pushed advertising channels. At the same time, the integrity and economic sustainability of the platform as a central data collector is secured and privacy rights for users are preserved. Solving information asymmetry market failures leads to more competitive platform markets and increased social welfare. It boosts incentives for innovation and helps small efficient firms scale when operating on a platform.

Our second and related research question explores how changes in the asymmetric information distribution in platforms affects competition and innovation inside and between platforms. Several reports (Furman et al., 2019; Scott Morton, 2019; Crémer et al., 2019) have indicated that large-scale data collection by platforms can generate market failures due to static competition problems (access and pricing) and dynamic innovation (market entry) bottlenecks in data-driven services markets. The problem is particularly acute for very large "gatekeeper" platforms. They achieve dominant positions because of traditional number-driven network effects: more users attract more users. Once the market has "tipped" and consolidated a platform's dominant position, network effects make new entries very difficult and may result in monopolistic behaviour. Data-driven network effects amplify traditional number-driven network effects and further entrench platform market positions. Economies of scale and scope in data aggregation (Martens, 2020) generate data-driven network effects or externalities: the quality of service for a particular user is affected by data collected from other users (Acemoglu et al., forthcoming; Choi et al., 2019). More and better data help to improve the quality of algorithms through learning-by-doing within and across users that further entrench market positions for incumbent platforms (Hagiu and Wright, 2020). Data can be re-used and bundled to invade adjacent markets (Condorelli and Padilla, 2020; Langus and Lipatov, 2021). Policy makers seek to address these competition issues by means of a mixture of traditional competition policy tools and new tools, including mandatory data access rights. For example, under the proposed EU Digital Markets Act (DMA), gatekeeper platforms would have to grant business users access and portability rights to the data generated by their activities on a platform; gatekeeper search engines would have to share search terms with smaller search service providers; and advertising platforms would have to grant advertisers and publishers access to market data. We demonstrate in this paper that an *in-situ* data access right for consumers and firms increases competition within and between platforms. While the *in-situ* right can in principle be applied to all platforms, irrespective of size or market power, it imposes more onerous infrastructure access conditions on relatively small platforms. We discuss specific options that can address potential adverse effects on small platforms by the introduction and application of this access right.

The remainder of this paper is structured as follows. Section 2 examines market failures in platforms, due to asymmetric information between the platform and its users and the biased and incomplete information signals that it sends to consumers and sellers. Section 3 explores the scope of the *in-situ* data access right for firms and consumers. It also addresses some further issues on the implementation of the *in-situ* data rights. Section 4 discusses how the *in-situ* access right compares with other proposed data rights. Section 5 concludes.

2 Information market failures among digital platforms

We illustrate information market failures within and between digital platforms with the help of two product differentiation models. We first examine the economic implications of information structure and sharing within a platform market following a Salop (1979) model. This standard framework allows us to demonstrate that a more efficient market results from changing information structure as opposed to banning vertical integration. We also observe that platform entry reduces prices yet is not socially optimal. Absent data access rights, a platform can nudge partners toward less profitable market segments. It can also thwart competitive entry and oversight.

A market with n horizontally differentiated products can be represented by a Salop circle with 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.[5]

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 over market conditions and locations of all players.[6] The actual distance between sellers represents consumer demand for their products. Consumers located closer to a seller have 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.

[5] 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.

[6] We assume that consumers can only interact with sellers through the platform. Due to information asymmetry, search costs to find relevant sellers outside the platform are prohibitively high.

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 good produced from the platform's recommended firm.

We consider three different information regimes on the sellers side:[7] 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 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.[8] Hence, sellers only form expectations over these locations (prior to stage 1 of the above timing) based on a distribution that is common knowledge. 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.

We now describe the case of full information when $n = 3$. Firms are able to observe the exact location of their competitors on the circle. Let X_{12}, X_{23} and X_{31} be the distances between firms 1 and 2, firms 2 and 3, and firms 3 and 1, respectively. This implies $X_{12} + X_{23} + X_{31} = 1$.

[7] Our motivation to focus on information structures and regimes with emphasis on the sellers side and not on the consumers side is the limited capacity of the latter group to process multidimensional information sets. For sellers, on the other hand, it is, in principle, easier to develop automated algorithmic systems for processing such information. Indeed, the European Commission's E-commerce Sector Inquiry (https://ec.europa.eu/competition/antitrust/sector_inquiry_final_report_en.pdf) indicates that the vast majority of online sellers use such algorithmic systems.

[8] This regime as it will be explained below represents the data portability right that is currently enforced by some privacy regulations.

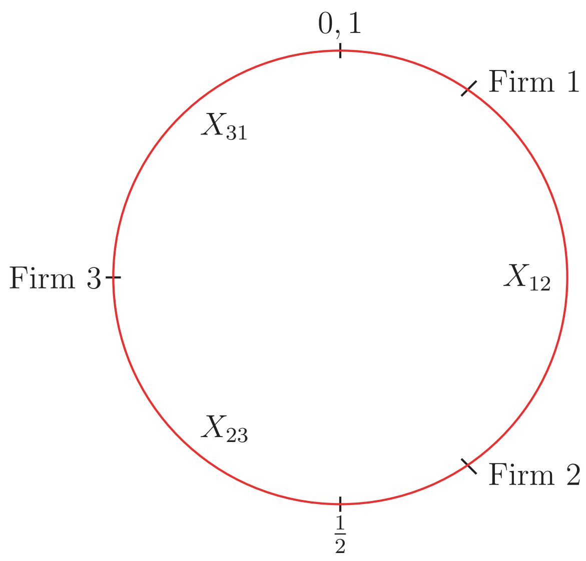

Figure 1: The Salop circle.

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:

$$\begin{aligned} p_i &= c + \frac{t(X_{jk} + 2X_{ij} + 2X_{ki})}{5} \\ &= c + \frac{t(1 - X_{ij} - X_{ki} + 2X_{ij} + 2X_{ki})}{5} \\ &= c + \frac{t(1 + X_{ij} + X_{ki})}{5}. \end{aligned}$$

By substituting p_i in D_i we conclude that profit function is a function of X_{ij} and X_{ki} as follows:

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

At the first stage, each firm i simultaneously chooses its distance from firms j and k to maximize Π_i. Due to the fact that firms are symmetric, in equilibrium, they choose locations such that

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

The respective equilibrium prices and profits are:

$$p_1^* = p_2^* = p_3^* = p^* = c + \frac{t}{3} \qquad \textit{and} \qquad \Pi_1^* = \Pi_2^* = \Pi_3^* = \frac{t}{9}.$$

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 choice of position, since we achieve both the maximization of sellers' profits and the minimization of consumers' average transportation costs.

Consider now what happens in the no information case, when $n = 3$. 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 at the second stage its price as before. 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 low willingness to pay and consumers at the short end of the market will have 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 cost.

In the n=3 case, we consider now that one firm (say, firm 3) is vertically integrated with the platform.[9] 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 so as to maximize the payoff from its information advantage.[10] 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.[11] In stage 3,

[9] Vertical integration is frequently related with abuse of dominance of platforms' market power competition policy cases, like the ones in Europe: In the Google Shopping case (https://ec.europa.eu/competition/antitrust/cases/dec_docs/39740/39740_14996_3.pdf), Google Search was fined for self-preferencing, because, by promoting its own services to consumers, it distorted competition by making it difficult for competitors to reach consumers. The practice of tying Microsoft's Internet Explorer to the Windows operating system was similarly found to be anticompetitive (https://ec.europa.eu/competition/antitrust/cases/dec_docs/39530/39530_3162_3.pdf) for providing an illegal advantage to Microsoft's own vertical integrated services. Apple is currently under scrutiny for setting disproportionally high fees to businesses that seek to participate in its app store, while its own subsidiaries and direct competitors to those businesses do not pay any fee (https://ec.europa.eu/commission/presscorner/detail/en/ip_21_2061). Amazon is currently under investigation for using non-public business data from independent sellers who sell on its marketplace to benefit Amazon's own retail business, directly competing with those third-party sellers (https://ec.europa.eu/commission/presscorner/detail/en/ip_20_2077).

[10] In the case of vertical mergers and acquisitions in which big platforms are involved, Parker et al. (2021) illustrate that 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. For example, after its acquisition, Hotmail offered to 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.

[11] Given a royalty rate ρ, platform profits with independent sellers are $\rho(\Pi_1^* + \Pi_2^* + \Pi_3^*) = \frac{3\rho t}{9}$, whereas under vertical integration, platform profits are $\rho(\Pi_1^* + \Pi_2^*) + \Pi_3^* = \frac{5\rho t}{9}$. Incremental revenue $(1 - \rho)\Pi_3^*$ motivates bias.

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.[12] The degree of bias b is chosen by the platform and there are two possibilities:

- *Bias up to nearest neighbor:* For values of b, such that $b < X_{23} - y$, bias will not affect consumers located within distance X_{12}. It will only reduce demand for firm 1 and 2 on consumers located between them and firm 3.
- *Bias beyond nearest neighbor:* For values of b, such that $b > X_{23} - y$, bias squeezes firms 1 and 2 in the region X_{12} where firm 3 previously had no market share.

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 price and profit:

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

Due to the fact that firms 1 and 2 are symmetric, at the first stage, the integrated firm select 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:[13]

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

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, 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.

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

[13] 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 $\Pi_i^{nvi} = \frac{t}{3}\left(\frac{p^{vi} - p^*}{2t} + \frac{X_{12} + X_{i3}}{2} - b\right)$.

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 an 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 position to observe whether the platform is biased or not and react to that accordingly. They can take a proper legal action or inform authorities about platform wrongdoing. Under asymmetric information, however, blind sellers do know 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. This will motivate an *in-situ* data right, allowing access not just to data but also its context, as introduced in the next section. Information sharing enables uninformed firms to observe market conditions and the locations of their competitors and update their strategies such that consumer welfare is maximized.[14]

[14] Regulators, especially in the US, have focused on the negative implications of vertical integration in platform markets. One of the proposed solutions which is currently discussed in the US Congress is the Ending Platform Monopolies Act (`https://www.congress.gov/bill/117th-congress/house-bill/3825`) which prohibits vertical integration for specific platforms. Our results here indicate that we can achieve even higher consumer surplus if instead we focus on enabling mechanisms that lead to effective information sharing across the platform market.

So far we have explored market failures given the information structure within a platform market. Other market failures related to information structures are also present between platforms that compete with each other. Platform users generate information from interacting with other users. Interactions include clickstream behavior, posts, likes and comments they make. This allows platforms to collect information on their users and improve the quality of their services towards them. In the presence of strong network effects, the emergence of very big platforms with accurate user profiling limits opportunities for consumers to interact with the supply side by using an alternative channel to the big platform's. Smaller competing platforms are not able to provide efficient transactions because they lack access to vital information collected by big platforms for generating efficient transactions.

As a result, consumers are forced to single-home, giving rise to a competitive bottleneck market failure (Armstrong, 2006; Armstrong and Wright, 2007): By attracting consumers, the platform extracts excessive rents on the supply side by charging a monopoly price. Information sharing, in this case, can help competing platforms to better understand the preferences of their users and improve their services. This can give rise to multi-homing on the consumer side which implies that the platform will have to reduce its price on the supply side resolving the market failure (Belleflamme and Peitz, 2019; Bakos and Halaburda, 2020). Sellers also have more incentives to multi-home as they can interact with consumers through alternative channels. As a result, information sharing through enabling multi-homing improves both consumer and seller surplus.

To see this, consider a big platform A and a competing smaller platform B which are located at the extremes of a Hotelling (1929) line, at 0 and 1. The two platforms offer differentiated services to consumers. Consumers and sellers have mass 1 and are distributed uniformly across the line. Consumers that are closer to platform A have preferences that better match platform A's service. Again we use the transportation cost incurred by the user to arrive at one of the platforms, as the parameter that indicates the distance between platform's service and consumers' preferences (without loss of generality, the transportation costs per unit of length for consumers are normalized to 1). Sellers do not incur any transportation cost. Platforms' marginal costs equal 0. Let $\mu(\nu)$ be the indirect network effect of sellers (consumers) on consumers (sellers). We assume that $\nu > \max\{0, \mu\}$, where $\mu\nu < 1$ and μ can be either positive or negative.[15]

[15] We can expect μ to be negative if consumers are annoyed by sellers' advertising efforts while they are consuming platform's service and positive when sellers' presence can increase the variety of products and services consumers are looking for in the platform market. For simplicity, we consider the case that consumers exert a stronger network effect on sellers than the sellers on consumers. But, note that this assumption is not necessary for illustrating the competitive bottleneck market failure and how it can be resolved through information sharing and multihoming. Last but not least, the assumption that $\mu\nu < 1$ is a standard assumption of the Hotelling product differentiation model.

Each platform j, based on its user information, provides a service of quality V_j to consumers (where, $j = A, B$). Platform j sets an entry price p_j for each seller who joins the platform. Consumer i's utility from participation in the platform j is $V_j + \mu s_j - x_{ij}$, where s_j is the share of sellers who join platform j and x_{ij} is the distance of the consumer i from the platform j. Seller's payoff is defined in an analogous way, as $\nu x_j - p_j$, where x_j is the share of consumers who join platform j. Consumers differ with respect to their transportation cost (distance) x_{ij} to arrive to platform j.

Under information asymmetry and without any mechanism in place to reduce this asymmetry between the two platforms, we have: $V_A > V_B = V_B^-$. Information sharing allows smaller platform B to provide more efficient services increasing the value to $V_B^+ > V_B^-$ and compete on a more equal footing with larger platform A for consumers. That makes multihoming on the consumer side more possible, in equilibrium.

Let $\overline{x}_A$ and $\overline{x}_B$ be defined as the marginal consumers who would join platform A and B, respectively, if they were alone in the market (and therefore, all sellers would have joined that platform). We have:

$$\begin{aligned} V_A + \mu - \overline{x}_A = 0 \quad &\Rightarrow \quad \overline{x}_A(V_A) = \overline{x}_A = V_A + \mu \\ V_B + \mu - (1 - \overline{x}_B) = 0 \quad &\Rightarrow \quad \overline{x}_B(V_B) = \overline{x}_B = 1 - \mu - V_B. \end{aligned}$$

We assume that $\overline{x}_B^+ = \overline{x}_B(V_B^+) \geq 0$. The question then becomes what the welfare implications of information sharing are in the equilibrium of the Hotelling platform game. Based on the quality offered by each platform, we have two different cases each of which give different equilibria when there is no information sharing:

- Case 1: $V_A + V_B^- \leq 1 - 2\mu$, which implies that $\overline{x}_A \leq \overline{x}_B^- = \overline{x}_B(V_B^-)$. Consumers either single-home or do not visit any platform.
- Case 2: $V_A + V_B^- > 1 - 2\mu$. There is the potential for multihoming for consumers in $(\overline{x}_A, \overline{x}_B^-)$. The rest of consumers either single-home or do not visit any platform.

In both cases sellers can multihome if they wish.

We show that in each of these cases, information sharing is associated with a greater surplus for consumers and sellers in each of these two cases.

Starting with Case 1, note that the share of consumers who joins platform A increases with the difference in the value of the services of platform A and B, $V_A - V_B^-$. When this difference is large, the majority share of consumers is captured by platform A. In equilibrium, platform A finds optimal to set a price $p_A = \nu \overline{x}_A$, while platform B sets price $p_B = \nu(1 - \overline{x}_B^-)$. The share of consumers $[0, \overline{x}_A]$ visits platform A, the share $[\overline{x}_B^-, 1]$ visits platform B.

Consumers in $(\overline{x}_A, \overline{x}_B^-)$ do not visit any platform. Sellers visit both platforms but derive zero surplus.[16]

By being able to capture the largest share of the single-homing side of consumers, platform A can extract all surplus from sellers who wish to join A to interact with consumers. In this competitive bottleneck equilibrium, sellers have only weak incentives to multihome when $V_A - V_B^-$ is large because, by selling through platform A only, they can already interact with the largest share of consumers $\overline{x}_A$.

Network effects make it easier for the competitive bottleneck equilibrium to arise and maximize the excessive rents extracted by platform A for two reasons. First, the higher this network value for sellers ν is, the higher are the rents, the platform A can extract from the sellers. Second, the share of consumers that prefer platform A increases with network effect μ and so does the price platform A charges the sellers.

When consumers share their information with platform B, the service quality offered to them, V_B^+, increases to a level comparable to platform A's quality. Let this increase be sufficient, such that $V_A + V_B^- > 1 - 2\mu$, which is sufficient so that $\overline{x}_A > \overline{x}_B^+$.

The two platforms "play" a Bertrand pricing game to attract sellers to their market. Their monopoly power is reduced to the portion of consumers that single-home. In equilibrium, platform A sets $p_A = \nu \overline{x}_B^+$ which is declining in V_B^+ and platform B sets $p_B = \nu(1 - \overline{x}_A)$. So, the share of consumers $(\overline{x}_B^+, \overline{x}_A)$ multi-homes, while the shares $(0, \overline{x}_B^+)$ and $(\overline{x}_A, 1)$ single-home (see Figure 2).

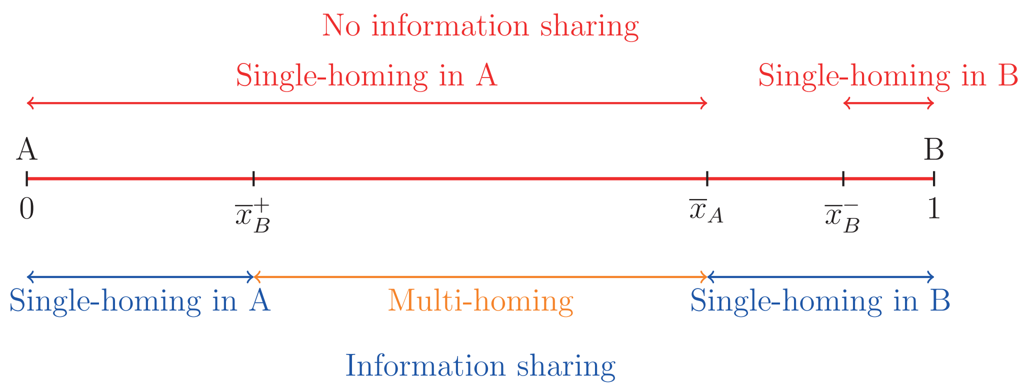

Figure 2: Platform competition across the Hotelling line with and without information sharing, when no information sharing implies that some consumers only single-home. The shares of consumers that single-home and multi-home with and without information sharing are depicted.

[16] The underlying assumption is that if sellers are indifferent between singlehoming and multihoming, they choose to multihome.

Note that under information sharing no platform has incentives to increase its price above the equilibrium level, because then sellers will only join their competitor (since accessing one platform does not only give them access to consumers that single-home on that platform, but also to the consumers that multi-home). Due to multi-homing, sellers now extract positive surplus and in equilibrium they all multi-home (given their zero transportation cost). Platform A now sees part of their rents being extracted by the sellers. This extracted surplus is proportional to the share of consumers that multi-home or in other words to the value V_B^+. Platform B realizes higher rents under information sharing and consumer single-homing since $1 - \overline{x}_A > 1 - \overline{x}_B^-$.

While consumers in $(0, \overline{x}_B^+)$ and $(\overline{x}_A, 1)$ have the same welfare as in the single-homing case, consumers within $(\overline{x}_B^+, \overline{x}_A)$ generate higher surplus under multi-homing, because they get positive surplus from both platforms instead of being unserved with zero benefit. So, overall, information sharing and the resulting rise in V_B increases both consumer and seller surplus, in comparison to the Case 1 where there is no information sharing and consumers either single-home or remain unserved. The higher the value V_B^+, the greater will be the consumer and seller welfare gains and lower the rents extracted by the big platform A.

Now we compare the welfare implications of information sharing with Case 2. Information sharing again increases the value of platform B for consumers to level V_B^+ for which we now have $\overline{x}_A > \overline{x}_B^- > \overline{x}_B^+$. Information sharing expands the market share of consumers that multi-home (see Figure 3). The expansion of multi-homing over the Hotelling line is proportional to the increase of the value of platform B from V_B^- to V_B^+.

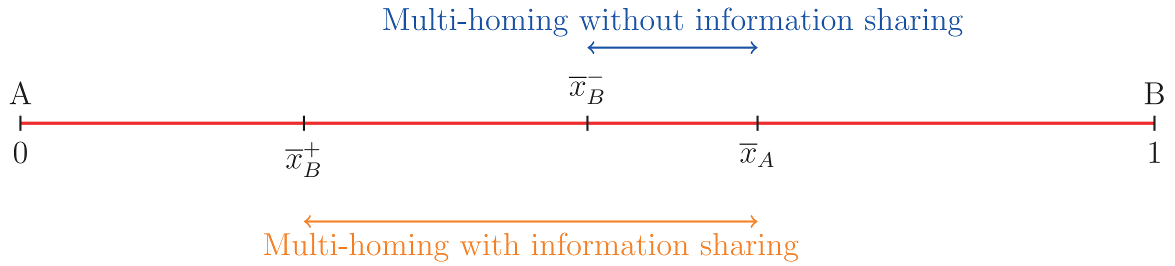

Figure 3: Platform competition across the Hotelling line with and without information sharing, when no information sharing implies that at least some consumers multi-home. Information sharing increases the share of consumers that multi-home.

The expansion of multi-homing under information sharing increases consumer surplus as there is a larger share of consumers that visit both platforms and derive extra surplus. Sellers are again better off since equilibrium platform prices are lower under information sharing. Since in equilibrium, $p_A = \nu \overline{x}_B^+ < \nu \overline{x}_B^-$ and $p_B = \nu(1 - \overline{x}_A) < \nu(1 - \overline{x}_B^-)$ information sharing constrains the ability of both platforms in this case to extract rents from sellers. Consumer welfare again increases because the share $(\overline{x}_B^+, \overline{x}_B^-)$ is strictly better off under information sharing.

So, we conclude that as in the Salop (within the platform) model, also here, between platforms, private incentives (by big platform A this time) for information sharing do not coincide with public ones. This is because information sharing enables greater degree of multi-homing which transfers market power from platform A to their upstream suppliers and to competing platform B (when initial information asymmetry is high). In this way, multi-homing can contribute to the correction of platform market power failures by redistributing value from big platforms to their sellers and consumers (and also smaller platforms in highly asymmetric platform ecosystems).

Proposition 2. *Let two platforms A and B compete in a Hotelling product differentiation line, while being asymmetric in terms of the quality of service they offer to their users, $V_A > V_B$. When information sharing sufficiently increases the quality V_B at levels $V_B > 1 - 2\mu - V_A$, then, it also increases consumer multi-homing, consumer welfare and business users' profitability.*

When consumers multi-home between A and B, data collection in platform A will be reduced because some interactions between consumers and sellers take place in platform B now. That could potentially decrease the value V_A while increasing the value of platform B at level V_B^+. The welfare implications of information sharing illustrated above remain the same as long as the sum of values of the two platforms $V_A + V_B$ increases. In the *in-situ* data rights proposal we develop below, the value for data and network effects is distributed in a way that the primary focus is on increasing V_B without a severe decline in V_A.

Corollary 2. *An information sharing regime between a platform A and a platform B that increases V_B without decreasing $V_A + V_B$ leads to a higher consumer and business user welfare.*

Information sharing also increases platform B's profit if its quality V_B^- is low.

Corollary 3. *If prior to any information exchange platform B's offered quality is low enough, such that $V_B < 1 - 2\mu - V_A$, then, an information sharing mechanism that increases consumer and business user welfare, also increases platform B's profit.*

Note that the network value μ makes it more likely for information sharing to increase consumer welfare. Once quality V_B increases, consumers are more inclined to multi-home towards platform B due to network effects. While the network value provides extra market power to platform A when a significant share of consumers single-homes, it can also contribute to the competitive pressure exerted by platform B when it increases its quality through an information sharing regime. Hence, it is therefore important for this regime to be designed in a way that it also transfers network value from the one platform to another. This is a main motivation for the *in-situ* mechanism for consumers we introduce in the next section.

3 *In-situ* data access

In the previous section we presented a Salop model to demonstrate how firm access to platform data could level the playing field between vertically integrated platforms and their business users. We also used a Hotelling model to show how consumer access to platform data could boost their welfare. In this section we elaborate on the type of data that should be accessed and the modalities of that access.

3.1 Existing EU data access rights and the *in-situ* right

Platforms can monetise their data by direct data sales to a service producer, or indirectly by producing and selling its own data-driven service. Examples of indirect data transfers include search or advertising services where users do not get access to the underlying raw data, only to a service produced by the algorithm. Casual observation suggests that indirect data access via services trade is the more ubiquitous and dominant data-driven business model in online platforms. Direct data trade is less prominent because it erodes data control. That maintains information asymmetry between platforms and their users.

Existing regulatory intervention and mandatory data access rights are limited in scope and do not really address information asymmetry between platforms and their users. In the EU for example, which has the most developed data access regulation, consumers benefit from data portability rights under the General Data Protection Regulation (GDPR, 2016). It allows natural persons to transfer their "own" data, including data actively provided by a data subject and data observed from the use and interaction with a service by a data subject.[17] How far the interpretation of interaction data reaches is unclear however.

Business users have no comparable data portability rights. Outside platforms, specific data access and portability rights exist only in a few sectors in the EU, including for automotive, energy and financial services data (Graef and van den Boom, 2020; Krämer et al., 2020). In the automotive sector, maintenance service provider firms have a right to access vehicle

[17]See relevant guidelines at `Article29WorkingParty(2016)Guidelinesontherighttodataportability, WP242,rev1,April2017`.

maintenance data. In energy and financial services, consumers have a data portability right to transfer their data to competing service providers.

Data access and portability rights for firms operating through intermediary platforms have recently come into focus however. Under the existing EU Platform-to-Business Regulation,[18] business users can ask a platform to access their own contributed data. The platform may refuse but should provide substantive reasons for doing so. The proposed EU Digital Markets Act (DMA)[19] is the first to-be-regulation to introduce a set of data portability rights for firms provided they operate through very large gatekeeper platforms Cabral et al. (2021).[20]

All these existing data access rights are *ex-situ* access rights: data are transferred directly from a platform to the server of the rights holder. We propose the introduction of an *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*.

3.2 The scope of an *in-situ* access right for firms

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 a standard *ex-situ* data access and portability right that would normally only include information on direct transactions between a firm and its consumers. These are zero-degree interaction data because they do not involve an intermediary agent (see Figure 4). The *in-situ* access right should reach further and include first- and second-degree interaction data. First-degree interactions include information on the behaviour of consumers (e.g., online searches and views) who interacted with the firm's products as well as with products from competing firms. In this case, consumers act as information intermediaries between two firms.

Second-degree interactions involve two intermediaries, a consumer and a firm. This includes data from consumers who interacted with products from competing firms (second

[18] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32019R1150.

[19] https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-markets-act-ensuring-fair-and-open-digital-markets_en.

[20] However, the scope and modalities of these proposed DMA data access and portability rights for firms have not been defined yet.

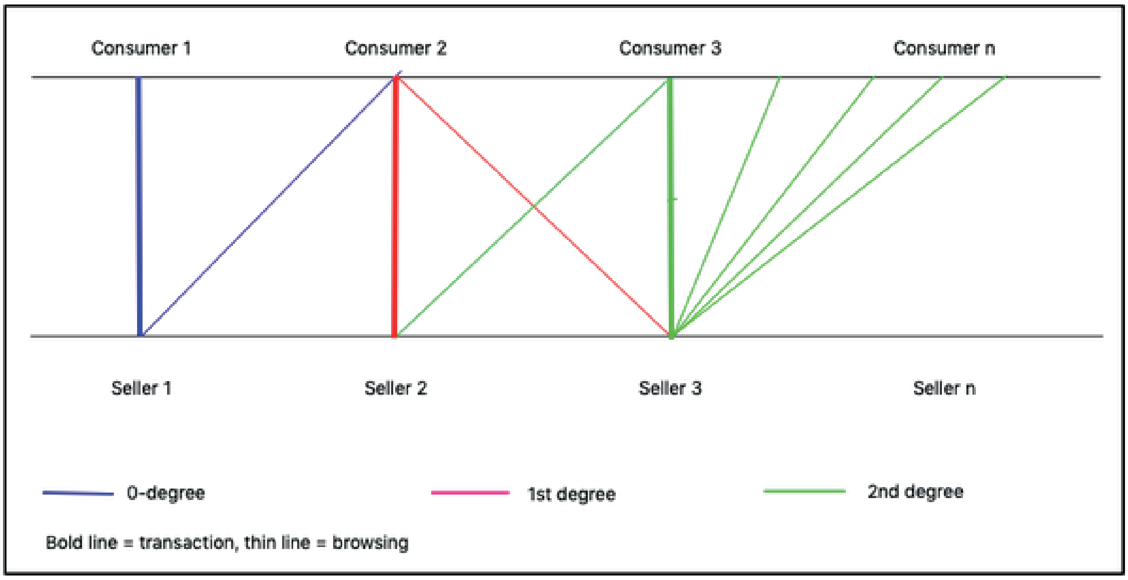

Figure 4: Degrees of platform interaction data. *In Situ* data rights enable context visibility beyond network degree-0 for which negotiation or legislation can provide higher degree access.

intermediary) that were also looked at by consumers who interacted with the firm (first intermediary). Such wider data access rights enable firms to assess consumer preferences and willingness-to-pay for a range of close-substitute competing products. It would provide sufficient information to a firm to estimate distance to competitors and (re-)position itself more optimally in the platform market, as explained in the Salop model in Section 2. 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 less consumers. However, if any of their consumers would directly interact with competing superstar sellers (seller 3 in Figure 4), second-degree data access allows these small sellers to "hop" to a much wider market overview and enable 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.[21] 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 behaviour of buyers, etc.) in a way that consumers remain anonymous.

Such information sharing can only occur *in-situ*. If such multidimensional information would be 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.[22]

As a result, another 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 the search and paid advertising channels. If firms would have direct 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 reduces asymmetric and incomplete information market failures inside platforms. Firms can use these 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. This will mitigate the problems from vertical integration of the

[21]See the next subsection for how personal data can be accessed *in-situ*.

[22]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. See also relevant discussion below.

platform with upstream providers, in terms of upstream competition and level playing field. Platforms will be less able to take advantage of information asymmetry for the benefit of their own upstream subsidiary.

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 monetisation 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 in 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.

In-situ data rights can lead us to the full information case of the Salop model of Section 2.[23] Data portability is equivalent to the no information case because it does not help firms to observe their distance from competitors. Hence, from Section 2, we conclude that:

Proposition 3. *As measured by social welfare, a policy banning vertical integration is dominated by a policy of* in-situ *data rights.*

Proposition 4. *A policy of data portability is strictly dominated by a policy of* in-situ *data rights.*

Data portability allows a firm to observe its own sales, i.e. data of network degree zero, but does not provide sufficient context to fully optimize decisions. In effect, by limiting access to a firm's or an individual's own data, it strips away information of network higher

[23] Full information should be understood as the observation of market location of competitors, which, as we showed in Section 2, is crucial to arrive at the socially optimal outcome.

degree. Data loss reduces the information available for making decisions, reducing market efficiency.

In effect, changes in information structure are both more welfare-enhancing and lighter touch regulation than changes in ownership structure. Changes in ownership structure eliminate unfair competition based on the platform's information advantage. They solve a competition problem but do not solve the asymmetric information problem. Nor do structural solutions solve the problem of shrinking network effects value based on shrinking network size. Instead, they make inefficient use of available information and make it harder to combine data pools.

3.3 *In-situ* access right for consumers

While *in-situ* data access on the supply side of a platform helps business users to position themselves more competitively inside the platform market, an *in-situ* right for consumers follows a somewhat different logic. Consumers come to a platform because it offers them access to much better market information than they can ever collect on their own. Platforms share pieces of their huge information pool with consumers through two information channels, consumer-pulled search rankings and seller-pushed advertising. These channels filter the total amount of information available in the platform and select only the supposedly most relevant information in response to a consumer query. Consumers need that filtering because bounded rationality and human cognitive capacity constraints do not allow them to process all available information. Filtering is subject to bias because platforms and advertisers are profit-maximizing firms that act in their own interest, not necessarily in the interest of the consumer (Ursu, 2018). As a result, consumers cannot be sure that the information signals they receive from a platform steer them to their welfare-maximizing choice. They can invest time in more search to further explore the market but search is costly. The design of search engine output signals affects these costs (Gu, 2016). When marginal search costs exceed expected marginal welfare gains, consumers will stop searching and settle for the available options or quit the market.

What can consumers do to make better use of the available information in platforms and reduce their search costs? Contrary to firms, consumers do not have autonomous data processing capacities and have to rely on the processing capacity of third-parties. Consumers can only rely on competition between search algorithms and platforms to reduce their information costs, as suggested by the Hotelling model of Section 2.

Without data sharing, consumers can still multi-home to another platform and start searching over again. That implies a loss of previous search and browsing histories and a costly re-introduction of these data in another platform. To avoid that cost, consumers could use their personal data portability right under the GDPR. That enables them to instruct the operator of the original platform to directly and digitally transfer their personal data to another destination platform. Effective real-time data portability[24] could result in more efficient information outcomes, facilitate consumer multi-homing and increase consumer welfare through competition between platforms. Still, since portability is limited to zero-degree interaction data, it only works in the absence of network effects.

Network effects or externalities are a key component of the social welfare benefits for platform users, both business users and consumers. We can distinguish between number-driven and data-driven network effects. Number-driven network effects imply that more users attract more users, more buyers attract more sellers to the platform, and vice versa. Data driven network effects imply that data collected from one set of users improves the service quality for another set of users, irrespective of the number of users on the platform. Continuous feedback between algorithmic services outputs and newly observed behavioural data inputs creates a learning-by-doing loop in platforms (Hagiu and Wright, 2020), both within and between individual consumers. The larger and more comprehensive the dataset collected by the platform, the higher the quality of the service, up to the point where diminishing returns to economies of scale and scope in data aggregation set in. Data-driven network effects imply that the social value of aggregated data is higher than the sum of values of segmented private data.

When data-driven network effects are important, individual datasets may exclude valuable market information from other users that is pertinent for a specific user. Porting one's own data removes that interaction context. In many applications, network effects and interactions are important in the value of the data. For example, downloads of personal Facebook data do not include posts by friends and colleagues – that is their data. Lack of context renders the data less useful,[25] unless context data-driven network effects are unimportant. For example, if wearables and home appliances are solely used to track one's own behaviour and consumption, network effects play no role. However, if behaviour and consumption patterns of others add value to the data, as in the case of social media for example, then networked context becomes important (Acemoglu et al, forthcoming). Platforms create value by the large-scale aggregation of small – scale network effects – those too small for individuals to create for themselves. Porting data off-platform adds friction to the creation of data driven network effects thereby reducing value. It is precisely these user interaction data that are so

[24]The GDPR currently does not mandate real-time data portability. Delays of up to three months are allowed. That makes portability useless for time-critical applications.

[25]See the experiment by David Berlind's team that re-affirms that: https://www.programmableweb.com/api-university/how-facebook-makes-it-nearly-impossible-you-to-quit.

valuable and that leaves recipients of ported individual data at a disadvantage compared to the platform.

Moreover, data is often not actionable when removed from the infrastructure where it was formerly resident. Off-platform data cannot be used to make a purchase i.e. to push a transaction, or to receive a reply or benefit, i.e. to pull a transaction, unless it is paired again with users on that platform. Unless the third-party data processor has ways to circumvent the platform and contact users directly, the data has less value outside the platform.

In addition, portability takes data outside the secured perimeter and introduces a moral hazard or hidden action problem where the receiving party can for reasons of intent e.g. economic gain or for reasons of negligence e.g. lax security allow data access to parties not authorized by the consumer. Once multiple data sources exist, tracing liability for improper disclosure becomes increasingly difficult with each additional source.

To overcome these limitations to *ex-situ* data portability, we could assign *in-situ* data rights to consumers over their interaction data. Consumers' behavioural interaction data would include volunteered or contributed data, such as search queries, and observed data, such as browsing behaviour and interactions with other products on the platform, and information contributed by business users on product or services characteristics, prices and other sales conditions. Observed data will include user interactions which to some extent are induced by platforms' algorithmic outputs, such as rankings and ads (that are based on user data inputs).

The *in-situ* mechanism works as follows: Consider again, as in the Hotelling model above, an incumbent platform A and an entrant competitor platform B.[26] Platform B requests consumer i for permission to access her interaction (personal) data located in platform A or, alternatively, consumer i invites platform B to access her data on platform A. Once consumer i gives her consent, platform A grants platform B access to i's interaction data at its location on platform A and use that data as an input for running its algorithmic applications on that site. In other words, instead of bringing the data to the entrant, the entrant's algorithm can be brought to the data located at the infrastructure of platform A. Consumer i's data is not transferred outside the infrastructure of platform A at any point in this process. However, platform B, through algorithmic analysis on site, can gain unique insights over consumer

[26] Instead of a platform B we may have a single firm or even a seller that already operates in platform A's market. *In-situ* access can help these firms to offer better quality and more personalised services to the consumer interacting directly with her, without having to rely anymore on platform A. This can increase the threat of disintermediation by putting competitive pressure on the platform to redistribute a larger share of the created value to platforms' users, in a similar way, as we saw in the Hotelling model of Section 2. Competition does not only increase by the entry of other intermediaries but also with the increased ability of users to interact with each other directly without platform intermediation. This can be an extra discipline device that will prevent platforms from extracting excessive rents from their business users.

i's preferences and characteristics. Hence, *in-situ* access allows consumers to share their search and interaction history directly with other platforms and get their (better quality) responses. This enables competition on top of a platform's infrastructure and exclusive data pool, which has the desired effect of forcing a platform to share more of the data benefits with users.

In practice the information sharing under *in-situ* can take place through the employment of application programming interfaces (APIs). In fact, a system of federated APIs (e.g., Google's FLoC) can be implemented. This allows a digital firm to get access to the data of a given user which is located in multiple big platforms at the same time, provided that user has given her consent. Extending the example of the previous paragraph, Platform B can get *in-situ* access to consumer i's data located not only in platform A but also in platforms C and D. For example, a user can give her consent to the Zalando platform for *in-situ* access of her data located at big platforms like Amazon and Google at the same time. Such a federated system can maximize the benefits from network effects through information sharing. This is because it makes it possible to build a large network where users can transfer the information they generate in one platform to another with multiple derived benefits from the better personalization of services as well as the value per interaction with other users. Platform B can offer a higher quality of service if it accesses data from platform A and C than if it only accesses data from platform A.

Note that the *in-situ* right for consumers does not reduce information asymmetry between consumers and platforms directly. In fact, there is not much point in trying to reduce that asymmetry because consumers have little or no autonomous data processing capacity to deal with a huge increase in information inflows. *In-situ* data access for consumer data does however reduce information asymmetry between larger and smaller platforms and thereby increases competition between platforms (see Hotelling model in Section 2). That, in turn, may help consumers to get better service quality from platforms and reduce biases in the signals that consumers receive from platforms. Consumers will receive a wider variety of service offers, at lower prices, compared to what a single platform offers.

The *in-situ* data right resolves the problems of *ex-situ* data portability identified above, when users share information. Since data always remains on site behind platform's firewall protection, it is possible to apply techniques such as encryption of user data, secure multi-party computation and federated learning to both keep the data in context and achieve some degree of anonymisation and aggregation to preserve the privacy rights of individuals that did not grant *in-situ* data access. In this way, we reduce the risk of information access that goes beyond a specific user's consent to include the information of users with whom she interacts. *In-situ* preserves privacy protection while retaining the value of data and information.

Another major advantage of *in-situ* access to consumer data, compared to portability, is that it enhances the ability to act on data. This is by construction of the *in-situ* right as data is accessed on the platform's infrastructure and therefore, when information is shared it is not necessary to connect the shared data with a new firm's infrastructure. At the same time, merchants and consumers cannot selectively edit unflattering facts and raise moral hazard risks for others.

Table 1 summarizes the main characteristics of the two *in-situ* rights we introduced, the ones for firms and the ones for consumers.

***In-situ* Data Rights**		
Right holder	Sellers	Consumers
Data holder	Platform	Platform
Type of data	Non-personal interaction data, consumer & product characteristics	Individual (personal) interaction data, individual preferences
Information recipients	Sellers of the platform	Other platforms and firms
Market failure addressed	Within platform	Between (outside) platforms

Table 1: Summary of *in-situ* data rights for firms and consumers.

3.4 Further issues

To which platforms would the consumers' *in-situ* data access obligation apply? Making the infrastructure for *in-situ* access available and ensuring the correct application of the right entails substantial responsibilities and costs for the platform. Moreover, the impact of this right will be different on large platform markets when platforms have access to multidimensional information through super datasets, namely, datasets of great volume and variety.

We know from Blackwell (1953) and Blackwell and Girshick (1979) that one superset data pool that contains a smaller data pool is always at least as valuable (case 1). On the other hand, the smaller data pool can be more valuable if it contains unique data absent from the larger pool (case 2). Value depends on the decision problem. There is no decision problem, answerable with the small data pool, that cannot be answered at least as well using the large data pool in case 1. By contrast, there are decision problems, answerable with the small data pool, that cannot be answered at all with the larger data pool in case 2.

Let a big platform with information superset A compete with a smaller firm whose much smaller dataset B has some unique elements that do not belong to superset A. Following the Hotelling model of Section 2, sharing information towards the smaller platform through the *in-situ* mechanism can be welfare improving and increase social value by eliminating market power related failures. Information sharing from the small firm to the big platform can improve the value of service of the latter. However, it may reinforce market power failures by reducing the share of consumers that multi-home and as a result allow the big platform to extract disproportional high rents from its business users.

Hence, while using the *in-situ* mechanism for information sharing towards the small firm can be welfare improving, sharing of information in the opposite direction, from the small firm to the big platform may be welfare reducing.

There are two alternative restrictions that we can impose on the *in-situ* access right in order to overcome this potential socially harmful information sharing equilibrium.

The implementation of the *in-situ* access right could be limited to very large gatekeeper platforms (which have developed extensive data supersets) only. Small firms can access consumers' data that resides on a big platform *in-situ*, but data is not shared in the opposite direction.[27]

Alternatively, we can allow *in-situ* to work in both ways, but we differentiate over the time of the access. Small firms can use the *in-situ* access right in real time, while big platforms can access data from small firms with some delay. Earlier access to the big firm's data means that small firms have a small lead time to access data *in-situ* and improve their services. We can adjust the lead time so that it is proportional to its disadvantage in size and information in comparison to its big platform competitors.

As already discussed, consumer's consent to a platform B for *in-situ* access primarily concerns her interaction data located in one or more platforms. This interaction data is derived from the combination of an algorithmic output and consumer's choices over this output. For example, an online search algorithm "responds" with a ranking of relevant alternatives to a consumers' query. Consumers make specific choices over this ranking of alternatives generating interaction data on the online search platform. When such interaction data is shared with another platform *in-situ*, there may be a risk of reverse engineering of the search platform's algorithm and gaming of the search engine. Since consumers' personal information is only revealed if there is consent, and the *in-situ* access is primarily targeted to help small firms to grow, this risk of reverse engineering is very limited. Big platform algorithms depend on a great number of factors which are difficult to be captured by their smaller competitors that have a narrower market focus and specialization.

There is an additional risk to be considered. When platform B is granted *in-situ* data

[27] Still, big platforms will be able to share data *in-situ* with each other. The restriction targets to improve small firms and their market expansion to create more competitive digital markets. The DMA proposes threshold criteria for the definition of gatekeepers that could be followed. Further threshold criteria exist in new legislative proposals by EU member states like Germany and Greece.

access at platform's A infrastructure, there is a possible risk of *in-situ* data processing because it puts the host platform A in a position to observe algorithm training and *in-situ* service production by platform B. The differential privacy techniques discussed above that enable federated learning can be used to reduce this risk and protect the secrecy of the incoming platform's algorithm.

The regulation of *in-situ* rights should be accompanied by compatibility standards. Some minimum standards over how firms can get access to big platforms' infrastructure will be required for this system to work. That essentially requires some standardization over the collected interaction data (Gal and Rubinfeld, 2019) and *in-situ* APIs which will allow firms to design accordingly their algorithms that will run on the data infrastructure. Instead of standards related to data formats, standards should focus on the design of algorithmic systems that are transferred to platform infrastructures for access to data on site.

4 Discussion

In-situ data rights aim at redistributing the value created in digital platform markets towards business users and consumers through effective information sharing that addresses within and between platform markets failures. They can be viewed as an information solution that changes information structures of digital platform markets, reducing the market and bargaining power of big platforms and contributing to a fairer distribution of the created value.

We already see some commercial practices that mirror the *in-situ* proposal. For example, one major healthcare service provider is no longer planning to send data out to its analytics/AI solution partners. Instead, they are providing full compute capabilities and access to real-time data within a cloud environment supported by the company. This idea is designed to solve a data fragmentation problem but it also addresses security and data currency issues. This is taking place amidst the backdrop of a larger migration to cloud technology for healthcare providers all while respecting privacy rights of patients, consistent with the Health Insurance Portability and Accountability Act (HIPAA).[28]

At the same time, the open algorithms (OPAL) project[29] aims to unlock the potential of data collected by private organizations "by bringing the code to the data through open algorithms and safe and fair technological and governance systems for better decisions in support of the sustainable development goals around the globe." The real-world deployment

[28] https://www.healthitanswers.net/why-true-healthcare-data-intelligence-depends-on-an-ai-powered-cloud/

[29] https://www.opalproject.org/home-en.

of OPAL started in mid-2017 in Colombia and Senegal. The main characteristic of this project is that algorithms are used in the data infrastructure of private companies behind the firewall protection with the goal of deriving key indicators *in-situ* that are shared with the users of the ecosystem.

The provision of third party permissioned access offers further benefit in the form of platform oversight. *In-situ* data rights allow invited auditors to review information presented to users, check it for bias, and benchmark it against users' preferences. One study by researchers at New York University obtained permission from 6,500 volunteers to track misinformation spread by political ads on Facebook.[30] Preliminary results showed that "extreme, unreliable news sources get more engagement ... at the expense of accurate posts and reporting," and that "the archive of political ads that Facebook makes available to [the public] is missing more than 100,000 ads."[31] Despite scolding from the Federal Trade Commission, Facebook claimed the research constituted unauthorized data scraping and terminated the researchers' accounts. While defending distribution of false political ads on the basis that users should decide,[32] Facebook thwarts users' capacity to understand the nature of misinformation they receive. Granting users a positive right to analyze their data, where it is resident, enables oversight.

Additionally, new EU privacy rules require technology platforms to limit micro-targeting in cases of political ads. Enforcement, however, requires platform cooperation,[33] which can be hidden information as in the case of the missing political ads on Facebook. Alternatively, monitoring by outside groups could help if monitoring were feasible. *In-situ* data rights provides a mechanism for ensuring oversight and compliance. Indeed, they enable platform competitors, who have both skills and motivation, to aid in compliance oversight. Thus *in situ* data rights provide an avenue for self-enforcement of users' rights and platform obligations via multi-party transparency without need of introducing state level oversight.

Furthermore, *in-situ* rights support the formation of consumer data unions or pools that foster competition and innovation.[34] This expands the possibilities for individuals to

[30] https://www.wsj.com/articles/facebook-cuts-off-access-for-nyu-research-into-political-ad-targeting-11628052204.

[31] https://www.nytimes.com/2021/08/10/opinion/facebook-misinformation.html.

[32] https://www.washingtonpost.com/technology/2019/10/17/zuckerberg-standing-voice-free-expression/.

[33] https://www.politico.eu/article/facebook-google-twitter-european-commission-political-ads/.

[34] In general, the management of a common pool of information can be complicated and costly (Ostrom, 1990; Hess and Ostrom, 2007). It is worth noting that since data is accessed *in-situ*, the pool essentially will not manage information, but only the consents of pool participants for accessing their information on the platform. That significantly reduces the involved management costs (which are also shared with platforms that hold participants' data).

monetize and increase the value of services they receive when they act as a team, encouraging the creation of a market for data (Koutroumpis et al., 2020). As aggregation can improve the generated value in the platform ecosystem, new platforms and firms will be inclined to provide additional benefits to individuals in order to reach sufficient critical mass to provide high quality services. The non-rival property of data makes this feasible. Consumer pools may also reduce the cost of consent and access to personal data. A newcomer platform or firm, instead of trying to approach each individual separately, can contact the management of the pool and make an offer for *in-situ* consent of its members in exchange of a reward or benefit (for example, the provision of a service). If an agreement is reached and some of the members provide their consent, firms can get valuable information from the pool's members (the ones that agree to participate) that can easily aggregate to derive additional competencies and thus compete more effectively with other firms or platforms.

Experimentation to create value is lower risk than porting data to third party firms. If the data pool or its members determine that a startup behaves badly, they can terminate data access confident that the offending firm has forfeited all further use cases and not having to rely on the goodwill of the offender to destroy private copies. Moreover, the pool itself will have incentives to study the preferences and background of its members to provide additional services. Through *in-situ* data access (for the members that provide their consent) it can perform data analysis and get a better picture of the demographics, health-records and preferences of the group. If it finds, for instance, that a certain subgroup of members is not paying sufficient attention to their health (e.g. by not using available medical services), it could devise strategies to remedy this situation, by intervening and negotiating, for example with external providers for better service rates (Hardjono and Pentland, 2019; Hardjono et al., 2019). The value from pool participation can increase further by effectively contributing to a more competitive market but defending users' interests in stronger ways.

5 Conclusions

Online platforms centralise large amounts of information on the behaviour of users on all sides of the market. This allows them to efficiently match users and generate stronger welfare-enhancing network externalities as compared to traditional decentralised offline markets where users must collect their own market information. Policy makers 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 due to information asymmetry between platforms and their users and between competing platforms. Without sufficient data rights, merchants and consumers only receive small excerpts from the pool of comprehensive market information collected by the platform.

We show how the lack of comprehensive market information creates obstacles for merchants selecting an optimal market entry point into the platform. This reduces business revenue and consumer welfare, and reduces market entry for innovative products. We show that consumers and merchants benefit from sharing their networked interaction data from one platform with other platforms.

We explore ways to improve information efficiency in platform markets by granting users an access right to a larger share of the platform data pool. In the EU, existing data portability rights already give consumers the right to access their personal data stored by the platform and transfer that data to another service provider for *ex-situ* processing. However, the scope of that right is limited to one's own interaction data, in effect, a small network radius with reduced context. At present, firms do not have an equivalent right. Data portability and processing outside the platform also has several disadvantages, including stagnation, loss of networked context, data security concerns, and loss of actionability.

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 it is collected, preserving the option value of context, while reducing privacy risk.

While larger firms may have the skills and capacity to operate their marketing algorithms *in-situ* in a platform, smaller firms and consumers may have to engage third-party services for this purpose. We argue that this could create a new and competitive market for data analytics services, while maintaining a sufficient degree of data exclusivity for platforms to ensure their economic sustainability, including their ability to generate welfare-enhancing network effects.

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.[35]

Indeed, the broad implementation of the *in-situ* mechanism can only be achieved through an *ex-ante* regulation that clearly defines obligations, rights and responsibilities of the different actors in digital markets. Moving forward by encouraging better information sharing can address existing market failures and improve market efficiency, competition, and welfare.

[35] https://www.europarl.europa.eu/doceo/document/IMCO-PR-692792_EN.html.

References

D. Acemoglu, A. Makhdoumi, A. Malekian, and A. Ozdaglar. Too much data: Prices and inefficiencies in data markets. *American Economic Journal: Micro*, forthcoming.

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

M. Armstrong. Competition in two-sided markets. *The RAND journal of economics*, 37(3): 668–691, 2006.

M. Armstrong and J. Wright. Two-sided markets, competitive bottlenecks and exclusive contracts. *Economic Theory*, 32(2):353–380, 2007.

Y. Bakos and H. Halaburda. Platform competition with multihoming on both sides: Subsidize or not? *Management Science*, 66(12):5599–5607, 2020.

P. Belleflamme and M. Peitz. Platform competition: Who benefits from multihoming? *International Journal of Industrial Organization*, 64:1–26, 2019.

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

D. Blackwell. Equivalent comparisons of experiments. *The annals of mathematical statistics*, pages 265–272, 1953.

D. A. Blackwell and M. A. Girshick. *Theory of games and statistical decisions*. Courier Corporation, 1979.

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

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

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

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

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

L. Cabral, J. Haucap, G. Parker, G. Petropoulos, T. M. Valletti, and M. W. Van Alstyne. The eu digital markets act: A report from a panel of economic experts. *Cabral, L., Haucap, J., Parker, G., Petropoulos, G., Valletti, T., and Van Alstyne, M., The EU Digital Markets Act, Publications Office of the European Union, Luxembourg*, 2021.

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

D. Condorelli and J. Padilla. Harnessing platform envelopment in the digital world. *Journal of Competition Law & Economics*, 16(2):143–187, 2020.

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

J. Crémer, Y.-A. de Montjoye, and H. Schweitzer. Competition policy for the digital era. *Report for the European Commission*, 2019.

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

Y.-A. De Montjoye, L. Radaelli, V. K. Singh, and A. S. Pentland. Unique in the shopping mall: On the reidentifiability of credit card metadata. *Science*, 347(6221):536–539, 2015.

J. Furman, D. Coyle, A. Fletcher, D. McAuley, and P. Marsden. Unlocking digital competitionreport of the digital competition expert panel. Technical report, UK Government, 2019.

M. S. Gal and D. L. Rubinfeld. Data standardization. *NYUL Rev.*, 94:737, 2019.

D. Gale and L. S. Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962.

GDPR. Regulation eu 2016/679 of the european parliament and of the council of 27 april 2016. *Official Journal of the European Union*, 2016.

I. Graef and J. van den Boom. Spill-overs in data governance: Uncovering the uneasy relationship between the gdpr's right to data portability and eu sector-specific data access regimes. *Journal of European Consumer and Market Law*, 9(1), 2020.

N. Gu. *Consumer online search with partially revealed information*. PhD thesis, The University of Chicago, 2016.

A. Hagiu and J. Wright. Data-enabled learning, network effects and competitive advantage. In *working paper*. 2020.

T. Hardjono and A. Pentland. Data cooperatives: Towards a foundation for decentralized personal data management. *arXiv preprint arXiv:1905.08819*, 2019.

T. Hardjono, D. L. Shrier, and A. Pentland. *Trusted Data: A New Framework for Identity and Data Sharing*. MIT Connection Science & Engineering, 2019.

F. A. Hayek. The use of knowledge in society. *The American economic review*, 35(4):519–530, 1945.

C. Hess and E. Ostrom. Introduction: An overview of the knowledge commons. 2007.

L. Hill, F. Levy, V. Kundra, B. Laki, and J. Smith. Data-driven innovation for growth and well-being, 2015.

H. Hotelling. Stability in competition. *The Economic Journal*, 39(153):41–57, 1929.

P. Koutroumpis, A. Leiponen, and L. D. Thomas. Markets for data. *Industrial and Corporate Change*, 29(3):645–660, 2020.

J. Krämer, P. Senellart, and A. de Streel. *Making data portability more effective for the digital economy: Economic implications and regulatory challenges.* Centre on Regulation in Europe asbl (CERRE), 2020.

G. Langus and V. Lipatov. Does envelopment through data advantage call for new regulation? 2021.

J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, A. Hung Byers, et al. *Big data: The next frontier for innovation, competition, and productivity.* McKinsey Global Institute, 2011.

E. Ostrom. *Governing the commons: The evolution of institutions for collective action.* Cambridge university press, 1990.

G. Parker and M. Van Alstyne. Innovation, openness, and platform control. *Management Science*, 64(7):3015–3032, 2018.

G. Parker, M. W. Van Alstyne, and X. Jiang. Platform ecosystems: How developers invert the firm. *MIS Quarterly*, 41(1):255–266, 2017.

G. Parker, G. Petropoulos, and M. W. Van Alstyne. Platform mergers and antitrust. *Industrial and Corporate Change*, 00:1–30, 2021.

G. Parker, G. Petropoulos, and M. W. Van Alstyne. Digital platforms and antitrust. *Oxford Handbook of Institutions of Economic Governance and Market Regulation*, forthcoming.

G. G. Parker, M. W. Van Alstyne, and S. P. Choudary. *Platform revolution: How networked markets are transforming the economy and how to make them work for you.* WW Norton & Company, 2016.

A. E. Roth. What have we learned from market design? *The Economic Journal*, 118(527): 285–310, 2008.

S. C. Salop. Monopolistic competition with outside goods. *The Bell Journal of Economics*, pages 141–156, 1979.

F. Scott Morton. Final report of the subcommittee on market structure and antitrust. Technical report, Stigler Committee on Digital Platforms, 2019.

R. M. Ursu. The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions. *Marketing Science*, 37(4):530–552, 2018.

Bruegel, Rue de la Charité 33, B-1210 Brussels
(+32) 2 227 4210
info@bruegel.org
www.bruegel.org