# Data-Driven Mergers*

Alexandre de Cornière† and Greg Taylor‡

2 February 2024

Accepted for publication at *Management Science*.

**Abstract**

We study mergers between firms operating in data-connected markets: the data generated as a byproduct of the activity on market $A$ can be used by firms operating on market $B$. The effects of such a merger depend on whether data trade among independent firms is possible, and on whether data use benefits consumers or leads to more surplus extraction. When data increases product $B$'s quality, the merger benefits consumers on both markets if data cannot be traded absent the merger, and harms them otherwise. When data is used to extract consumer surplus on market $B$ the merger increases consumer surplus on market $A$ and reduces it on market $B$.

## 1 Introduction

The continued growth of the digital economy and the emergence of dominant online platforms have lead academics and policy-makers to re-think the tools available to regulators and competition authorities (e.g. Furman et al., 2019; Crémer, de Montjoye, and Schweitzer, 2019). The role of data as a source of market power has attracted particular attention (e.g. Autorité de la Concurrence and Bundeskartellamt, 2016, for instance), resulting in new regulations (see for instance the European Data Act, or articles 5(a), 6(a,h,i,j) of the Digital Markets Act) as well as interventions by competition authorities.[1]

*We are grateful to participants at various seminars and conferences for useful comments and discussions. Taylor acknowledges funding from the Research Programme on AI, Government, and Policy, funded by the Dieter Schwarz Stiftung gGmbH.

†Toulouse School of Economics, University of Toulouse Capitole; alexandre.de-corniere@tse-fr.eu; https://sites.google.com/site/adecorniere

‡Oxford Internet Institute, University of Oxford; greg.taylor@oii.ox.ac.uk; http://www.greg-taylor.co.uk

[1] See for instance the Bundeskartellamt case against Facebook regarding the collection of user data (`https://www.bundesgerichtshof.de/SharedDocs/Pressemitteilungen/DE/2020/2020080.html`), the European Commission investigation of Amazon's practices regarding third-party sellers' data (`https://ec.europa.eu/commission/presscorner/detail/en/IP_19_4291`), or the commitments from Google obtained by the Competition and Markets Authority regarding the implementation of its "Privacy Sandbox" `https://assets.publishing.service.gov.uk/media/62052c52e90e077f7881c975/Google_Sandbox_.pdf`.

One important issue concerns the treatment of so-called “data-driven mergers”, i.e., motivated—at least partially—by the prospect of acquiring and/or combining the merging firms’ datasets. For instance, one reason Microsoft bought LinkedIn in 2016 was to use LinkedIn’s database to improve its own Customer Relationship Management software, Dynamics 365.[2] When Google offered to buy Fitbit in 2020, the European Commission (EC) and others raised concerns about how Google might use the Fitbit data in other markets.[3] In the recent acquisition of iRobot by Amazon, the EC also raised some data-related concerns, namely that the data collected by iRobot could help Amazon strengthen its market power as an online marketplace. Earlier mergers where data was an important part of the transaction include Facebook-WhatsApp, Verizon-Yahoo, Google-DoubleClick, TomTom-Tele Atlas, or Amazon-Goodreads (see Grunes and Stucke, 2016, for a discussion).

In this paper we study the effects of a merger between firms operating in data-connected markets: a primary and a secondary market, respectively denoted $A$ and $B$. Products $A$ and $B$ are neither substitutes nor complements, and the set of potential consumers may be different, but the data collected by firms in market $A$ as a byproduct of their activity can potentially be useful to the firms in market $B$.

Our model emphasizes several features that are specific to data, and that justify a specific analysis of this type of merger. First, the fact that data is collected as a byproduct of a firm’s economic activity implies that the amount of data it collects depends on the size of its customer base, and/or on the intensity of its use by consumers.[4] These in turn depend on the firm’s choices regarding price and quality, among others. Whether the merger increases or reduces incentives to collect data will therefore affect consumer surplus and welfare on the primary market.

Second, unlike an input in a typical vertical merger, data is sometimes subject to trade frictions, such as privacy regulations or reputational concerns (see below for a discussion). It may then be that a merger is the only way to use data in the secondary market. The effects of the merger thus potentially depend on whether data could be shared on the secondary market in the relevant counterfactual scenario.

Third, data can be used in various ways on the secondary market, to the benefit or detriment of consumers. For instance, data can be used to improve products, which tends to benefit consumers, but it can also be used for price-discrimination purposes, which sometimes harms consumers. The consequences of a merger that, say, results in an increase in the collection and use of data will naturally depend on which is the relevant case.

In their assessment of data-driven mergers, competition authorities have considered two main exclusionary theories of harm. In the first, the merger would allow a firm to get access to a new source of data that cannot be replicated by its rivals, thereby raising barriers

[2]See `https://www.reuters.com/article/us-microsoft-linkedin-idUSKBN17Q1FW`.

[3]The merger was eventually cleared subject to behavioral remedies.

[4]For instance, an individual who spends a lot of time on LinkedIn generates more data, by interacting more with his contacts, having a more complete profile, etc.

to entry or expansion and entrenching the firm's market power. The second one is an input foreclosure theory, where the merged entity stops supplying data to its downstream rivals. While plausible and potentially important, these theories are standard from a conceptual point of view, and they do not rely on the properties discussed above. Our main contribution in this paper is precisely to incorporate some features that distinguish data from other inputs and to articulate both a (non-exclusionary) theory of harm and a potential efficiency defense whose applicability depends on the properties of data.

In our model, presented in Section 3, we consider a merger between two firms operating in data related markets $A$ and $B$, where the data collected as a byproduct of the activity on market $A$ can be used by firms on market $B$. We start the analysis in Section 4 by looking at a specific example, where market $A$ is a monopoly, while two firms compete on a Hotelling line on market $B$. The data collected on market $A$ can be used to improve the quality of $B$ products. We show that the key condition for the merger to be socially desirable is the existence of frictions regarding the trade of data. To see this, suppose first that data trade is impossible, but that data can be shared within the same company post-merger. Following the merger, firm $A$ internalizes the value of data that can be used on the secondary market. It therefore finds it profitable to lower its price in order to collect more data, thereby benefiting consumers on the primary market. The merger also benefits consumers on the secondary market, since it allows the merging $B$ firm to offer a product of better quality. However, when data trade between independent firms is possible, the effects of the merger are reversed. Indeed, we show that, absent the merger, the monopolist in market $A$ would find it profitable to sell its data exclusively to one of the two $B$ firms. The value of data, determined through an auction between the $B$ firms, equals the profit difference between the $B$ firm who buys the data and the one who does not. In order to increase the price of data, firm $A$ has an incentive to reduce the profit of the non-buying $B$ firm. Compared to the integrated case, where the merged firm maximizes the sum of its subsidiaries' profits, this incentive to degrade the $B$-rival's profit induces firm $A$ to collect more data. Therefore the merger results in a higher price on the primary market, as well as a lower quality on the secondary one. Consumers on both markets are worse-off.

In Section 5, we generalize our analysis by considering a model of competition in utility on market $B$, which allows us to encompass various uses of data beyond product improvement (e.g. targeted advertising, price-discrimination). Building on the framework developed in de Cornière and Taylor (2023), we show that the insights from the product improvement model generalize to any model where data is *unilaterally pro-competitive* (UPC), that is when more data leads a firm to improve its offer to consumers. In other contexts, data can be *unilaterally anti-competitive* (UAC). Broadly speaking, this happens when data improves firms' ability to extract surplus from consumers. In such situations, we show that the merger benefits consumers on the primary market but hurts consumers on the secondary market, irrespective of the presence of trade frictions.

In the baseline model we assume that consumers do not strategically increase or reduce their usage of product $A$ in order to manipulate the data available on market $B$, but relax this assumption in Section 6. In Section 7 we discuss the robustness and policy implications of our analysis.

## 2 Literature review

A few recent papers study data-driven mergers. Most closely related is Chen et al. (2022a), where the market structure is similar to ours, with data collected in one market to be used in another for product customization and price-discrimination purposes. Unlike us, they do not allow for data to be traded without a merger, and emphasize the long-term exclusionary effects of the merger, whereas we focus on non-exclusionary harms. In a similar set-up, Herresthal, Mayskaya, and Nikandrova (2022) study a model where data can be used by insurers. They show that the merger enables the integrated insurer to cream-skim consumers and improves welfare. Its effect on consumers depends on the competitiveness of the market. Esteves and Vasconcelos (2015) and Kim and Choi (2010) consider horizontal mergers allowing firms to merge their datasets for price-discrimination.[5]

The paper is also related to the literature on data intermediaries who sell data to competing firms, surveyed in Bergemann and Bonatti (2019) and Pino (2022). Unlike papers such as Montes, Sand-Zantman, and Valletti (2018) or Bounie, Dubus, and Waelbroeck (2021) who consider partitioning of the dataset by the intermediary, we do not rely on an explicit information structure to model data, but instead use a more reduced-form approach which allows us to be more general with respect to the potential uses of data, as explained in more detail in de Cornière and Taylor (2023). Our modelling approach is thus closer to the licensing of intellectual property (Katz and Shapiro, 1986). Although de Cornière and Taylor (2023) provides the starting framework for the general version of our analysis, that paper does not consider data-driven mergers which is the focus here.

Our model is also formally related to D'Annunzio (2017), who studies vertical integration between a content producer and a distributor in a model close to our leading example (Section 4) in the case where data trade is possible between independent firms. Due to a different motivation, she does not discuss the case where trade would be impossible absent the merger, nor does she look at welfare on the primary market, and she has no equivalent to our case where data is unilaterally anti-competitive.

[5]Kim, Wagman, and Wickelgren (2018) compare the effects of a horizontal merger when firms can or cannot price-discriminate.

# 3 A model of data-connected markets

**Market structure** We consider an economy with two markets, $A$ and $B$. Market $A$ is monopolized by firm $A$. Market $B$ is a symmetric duopoly, with firms $B_1$ and $B_2$. Products $A$ and $B$ are neither substitutes nor complements, but the markets are *data-connected*: data generated as a byproduct of the activity of firm $A$ can be used by firms on the $B$ market. More specifically, if firm $A$ serves $q_A$ consumers, it generates a data set whose value for $B$ firms is increasing in $q_A$. We use $\delta_A \equiv q_A$ to index the quantity of data collected by $A$, and write $\pi_B(\delta_j, \delta_{-j})$ for firm $B_j$'s equilibrium profit when it has $\delta_j$ units of data and its rival has $\delta_{-j}$. We assume that $\frac{\partial \pi_B(\delta_j,\delta_{-j})}{\partial \delta_j} \geq 0$: better data is always valuable on the $B$ market.[6] We assume that consumers' decision of whether to use product $A$ (and thereby generate data) is independent of the subsequent use of that data on market $B$, either because they are myopic or because data is use at an aggregate level so that a single consumer's decision is irrelevant to the outcome on market $B$ (we can relax this assumption, as we discuss in Section 6).

It is helpful to reflect on the mapping from this setup to some of the cases mentioned in the introduction. For the purpose of evaluating a data-driven merger, the relevant "$B$" market(s) are those where the data will ultimately be used. For example, we have already mentioned that the Microsoft-LinkedIn merger was motivated by the use of LinkedIn's data in Microsoft's ($B_1$'s) customer relationship management software. In that market, the most important rival is Salesforce, which would correspond to firm $B_2$. Likewise, a primary question in the Google-Fitbit case surrounded Google's potential use of Fitbit's data in health markets, which would then constitute market $B$ (e.g., see Spagnolo et al., 2020). In the case of Amazon's acquisition of GoodReads, market $B$ would correspond to the retail market for books, and firm $B_2$ would represent other book retailers.

**Merger and data trade** Our main interest is evaluating the effects of a merger between $A$ and $B_1$ that allows $B_1$ to access the data generated by $A$.[7] However, a merger may not be the only way for a $B$ firm to access $A$'s data and, because of its non-rival nature, data can potentially be used by both $B$ firms.

One feature that distinguishes data from other inputs is that it may or may not be tradeable between independent firms.[8] We will consider two regimes. In the first regime,

[6]This rules out situations where $B_{-j}$ would react so aggressively to $B_j$'s getting better data that $B_j$'s profit would decrease.

[7]In some cases a competition authority may allow a merger only under the condition that data will not be shared between the merging firms. These cease to be data-driven mergers and the relevant concerns are then likely to be those of a more traditional conglomerate merger. But the examples given in the introduction make clear that many mergers *are* driven by the intention to combine data. More generally, CMA (2019) observe that integration can help firms to share data: "The GDPR makes gaining and managing consent [...] within an undertaking, or group of undertakings in common control, an easier exercise than sharing data between undertakings to deliver the same purpose."

[8]See Kim and Wagman (2015) for an empirical study of the effects of regulation preventing data trade

data trade between independent firms is impossible, meaning a merger is the only way for $A$'s data to be used in market $B$. This may, for instance, be due to privacy regulations that prevent the trade of personal data with a third party.[9] Another possible reason is the existence of contractual frictions, such as the inability to restrict what the data buyer can do with the data, which may deter the data holder ($A$) from sharing its data. For instance, if $A$ cannot control $B$ firms' efforts in securing the data, it may fear that a leak of personal data caused by $B_j$'s negligence could damage its reputation (as when Facebook shared user data with Cambridge Analytica). Or $A$ could fear that providing data to an independent $B$ firm could help that firm enter market $A$ in the future.

In the second regime, we allow $A$ to sell the data to independent firms. We model data trade the same way Katz and Shapiro (1986) model the sale of intangible property: $A$ chooses to sell its data to $k$ firms ($k \in \{1, 2\}$), and runs a sealed bid auction with reserve price $\underline{b}$. We do not allow $A$ to sell a subset of its data only (unlike e.g. Bounie, Dubus, and Waelbroeck (2023)).

**Timing and equilibrium** At $t = 0$, $A$ and $B_1$ merge or do not merge. At $t = 1$, $A$ offers a product to consumers in market $A$ and collects data from each consumer it serves. At $t = 2$, if data trade is possible, $A$ can choose to sell its data to any independent $B$ firm. At $t = 3$, the two $B$ firms observe the allocation of data and compete by offering a product or service to consumers in the $B$ market. We look for subgame perfect equilibria.

To complete the description of the model we need to formally specify the strategies and payoffs in markets $A$ and $B$. Our ultimate aim is to study a model (Section 5) that can accommodate a range of demand specifications, downstream business models, and uses for data. But we first demonstrate the main ideas with a specific example.

# 4 A leading example: data used to increase product quality

Suppose that demand for $A$'s product is $D_A(p_A) \equiv \max\{\alpha - p_A, 0\}$, where $p_A$ is its price. Marginal cost is normalized to zero. The amount of data collected is $\delta_A \equiv D_A(p_A)$, and data can be used to increase the quality of the product offered on market $B$.[10] The quality

on the mortgage market.

[9]Even though privacy regulations such as the European Union's General Data Protection Regulation (GDPR) also impose a "purpose limitation principle", that in theory restricts firms' ability to use personal data collected for one purpose for another one, we assume that the enforcement of such a principle is limited so that the merger allows firms to bypass it and to share the data internally.

[10]This formulation is consistent with Hagiu and Wright (2023) among others, and related to themes in Wickelgren (2015), who considers a dynamic model where data collection can improve product quality; and in Lefouili, Madio, and Toh (forthcoming), who study the relationship between data collection and incentives to invest in product quality. See (de Cornière and Taylor, 2023) for a microfoundation.

of $B_j$'s product is $V(\delta_j)$, where $\delta_j$ equals $\delta_A$ if $B_j$ buys data from $A$ and 0 otherwise, and where $V(\cdot)$ is a non-decreasing differentiable function.[11]

The $B$ firms are located at the ends of a Hotelling segment, with a transportation cost normalized to one. The utility of a consumer located at a distance $x$ from $B_j$ is therefore $V(\delta_j) - x - p_j$ if they buys from $B_j$. We assume that $V$ is such that the market is always covered, and that both firms are always active, even if only one of them has access to data. This means that we restrict attention to cases where data is not an essential input, even though it provides a competitive advantage. Our focus will therefore not be on the potential exclusion of rivals following the merger.[12]

## 4.1 Competition on the $B$ market

We begin by studying the subgame starting at $t = 3$, where $B$ firms have access to data $\delta_1$ and $\delta_2$. Demand for firm $B_j$ is given by :

$$D_B(\delta_j, \delta_{-j}, p_j, p_{-j}) = \frac{1 + V(\delta_j) - V(\delta_{-j}) - p_j + p_{-j}}{2}. \tag{1}$$

The equilibrium price and profits for firm $B_j$ are found in the usual way:

$$p_B(\delta_j, \delta_{-j}) = 1 + \frac{V(\delta_j) - V(\delta_{-j})}{3}, \quad \pi_B(\delta_j, \delta_{-j}) = \frac{(3 + V(\delta_j) - V(\delta_{-j}))^2}{18}. \tag{2}$$

Let $U_B(\delta_1, \delta_2)$ denote the equilibrium consumer surplus on market $B$. The following lemma highlights some properties of these payoffs that will play an important role later on:

**Lemma 1.** *In the product improvement model, data has the following properties:*

*(i)* *$B_j$'s profit is increasing in the quantity of its own data:* $\frac{\partial \pi_B(\delta_j, \delta_{-j})}{\partial \delta_j} > 0$.

*(ii)* *$B_j$'s profit is decreasing in the quantity of data held by its competitor:* $\frac{\partial \pi_B(\delta_j, \delta_{-j})}{\partial \delta_{-j}} < 0$.

*(iii)* *Consumer surplus on market $B$ is increasing in the amount of data held by each firm:* $\frac{\partial U_B(\delta_1, \delta_2)}{\partial \delta_{B_i}} > 0$ *for* $i = 1, 2$.

**Proof.** Parts (i) and (ii) follow immediately from (2). If we denote by $\hat{x}$ the indifferent consumer, consumer surplus is:

$$U_B(\delta_{B_1}, \delta_{B_2}) = \int_0^{\hat{x}} V(\delta_{B_1}) - x - p_B(\delta_{B_1}, \delta_{B_2})\, dx + \int_{\hat{x}}^1 V(\delta_{B_2}) - (1 - x) - p_B(\delta_{B_2}, \delta_{B_1})\, dx.$$

Using (1) and (2) yields $\frac{\partial U_B(\delta_{B_1}, \delta_{B_2})}{\partial \delta_{B_j}} = \frac{1}{18} [9 + V(\delta_i) - V(\delta_j)]\, V'(\delta_i)$. A necessary condition for the market to be covered is $V(\delta_j) \geq V(\delta_{-j}) - 3$, which guarantees (iii) holds. ■

[11] In practice firms may be able to obtain data from other sources than $A$: we assume that firms are symmetric with this respect, and normalize the quantity of outside data to zero.

[12] To ensure concavity of the profit functions we assume $V'(\delta)^2 + (3 + V(\delta) - V(0))V''(\delta) < 18$.

In this model, data increases a firm's quality, which allows it to charge a higher price and to generate more profit (property (i)). It also makes a firm a *tougher* competitor (in the sense of Fudenberg and Tirole, 1984) and reduces its rival's profit (property (ii)). Consumers are unambiguously better-off when firms have more data because price increases less quickly than quality.

## 4.2 Merger when data trade is impossible

We first assume that data cannot be traded. If the merger does not take place at $t = 0$ then at $t = 1$ firm $A$ maximizes profit in its primary market: $p_A^I = \text{argmax}_p\{pD_A(p)\}$. Since there is no way to supply data into market $B$, firms charge $p_B(0,0)$ in $t = 3$.

If $A$ and $B_1$ merge at $t = 0$ then the integrated firm takes into account that the data collected on market $A$ can be used by $B_1$. At $t = 1$ its price is therefore given by $p_A^M = \text{argmax}_p\{pD_A(p) + \pi_B(D_A(p), 0)\}$. By Lemma 1 (i), it is easy to see that the integrated firm has stronger incentives to collect data if it merges with $B_1$. This is achieved by serving more consumers on market $A$ and therefore charging a lower price: $p_A^M < p_A^I$. Consumers are thus better-off on market $A$. Moreover, by Lemma 1 (iii), the use of data made possible by the merger also increases consumer surplus on market $B$. We collect these results in the following proposition:

**Proposition 1.** *In the product improvement model, when data trade is not possible, a merger between $A$ and $B_1$ leads $A$ to reduce its price, and both $B$ firms to reduce their quality-adjusted prices. Consumer surplus increases in both markets.*

## 4.3 Merger when data trade is possible

**When is exclusivity optimal?** Here we take a step back an derive a result that applies for general profit functions. Suppose that firm $A$ can sell the data at $t = 2$, and consider the subgame where $A$ and $B_1$ are independent. At $t = 2$, $A$ decides whether to sell the data exclusively ($k = 1$), or to both $B$ firms ($k = 2$). Let $T^k(\delta)$ be the equilibrium price of a dataset of size $\delta$. Since $A$ is a monopoly supplier of the data, this price will extract the $B$ firm's willingness to pay, which is the difference between its profit if it buys the data and its outside option of letting $A$ instead sell the data (only) to its rival. We therefore have $T^1(\delta) = \pi_B(\delta, 0) - \pi_B(0, \delta)$, and $T^2(\delta) = \pi_B(\delta, \delta) - \pi_B(0, \delta)$.[13] Exclusivity is optimal if the revenue from selling data to a single firm at a price of $T^1(\delta)$ is greater than

[13] $T^1(k)$ can be obtained without using the reserve price, while getting $T^2(\delta)$ requires setting $\underline{b} = \pi_B(\delta, \delta) - \pi_B(0, \delta)$. Similar to Katz and Shapiro (1986), if $\pi_B(\delta, 0) - \pi_B(0, 0) < \pi_B(\delta, \delta) - \pi_B(0, \delta)$ then each $B_j$ will accept $T^2(\delta)$ only if it expects $B_{-j}$ to do the same, making this a coordination game. Thus, there is a subgame equilibrium with data trade and one without. We focus on the former, which can be selected by letting $A$ sell data sequentially.

that from selling it to two firms at a price of $T^2(\delta)$, i.e., if $T^1(\delta) > 2T^2(\delta)$. Substituting in the expressions for $T^1(\delta)$ and $T^2(\delta)$, this condition becomes $\pi_B(\delta,0)+\pi_B(0,\delta) > 2\pi_B(\delta,\delta)$.

Now suppose that $A$ and $B_1$ have merged. $A$ can sell the data to $B_2$ at a price $T^2(\delta)$, but this reduces $B_1$'s profit by $\pi_B(\delta,0) - \pi_B(\delta,\delta)$. Exclusivity is thus optimal whenever $\pi_B(\delta,0)+\pi_B(0,\delta) > 2\pi_B(\delta,\delta)$, which is the same condition as above. The following Lemma summarizes.

**Lemma 2.** *Equilibrium $B$-market allocation of data (exclusive or not) is the same with and without the merger. Data is allocated exclusively if and only if $\pi_B(\delta,0) + \pi_B(0,\delta) > 2\pi_B(\delta,\delta)$.*

Because the mechanism through which data is sold is public, the decision of whether to sell it exclusively or not maximizes the profit of the industry, and is therefore independent of the ownership structure. This rules out input foreclosure as a motivation for the merger (unlike Chen et al., 2022b, for instance).

The profit function in (2) satisfies $T^1(\delta) > 2T^2(\delta)$, meaning that $A$ finds it optimal to supply the data exclusively to one $B$-market firm, both before and after the merger (Lemma 2).

**Equilibrium without merger** Let $P_A(\delta)$ be the $A$-market price that generates an amount of data $\delta$, i.e. such that $D_A(P_A(\delta)) = \delta$. We have $P_A'(\delta) < 0$: in order to generate more data, $A$ needs to increase usage of its product, requiring a reduction in its price. Then, at $t = 1$, firm $A$'s optimal pricing satisfies $p_A^I = P_A(\delta_A^I)$, where

$$\delta_A^I = \mathrm{argmax}_\delta\{P_A(\delta)\delta + T^1(\delta)\} = \mathrm{argmax}_\delta\{P_A(\delta)\delta + \pi(\delta,0) - \pi(0,\delta)\}. \tag{3}$$

**Equilibrium with merger** Now suppose that $A$ and $B_1$ have merged at $t = 0$. By Lemma 2, $AB_1$ does not sell its data to $B_2$, and thus the optimal strategy for $A$ solves

$$\delta_A^M = \mathrm{argmax}_\delta\{P_A(\delta)\delta + \pi_B(\delta,0)\} \quad \text{and} \quad p_A^M = P_A(\delta_A^M) \tag{4}$$

Comparing (3) and (4), and using the fact that $\frac{\partial \pi_B(0,\delta)}{\partial \delta} < 0$ (by Lemma 1) as well as the concavity of $\delta \mapsto P_A(\delta)\delta$, we obtain:

**Proposition 2.** *In the product improvement model, when data trade among independent firms is possible, the merger leads to a higher price in market $A$ and to higher quality-adjusted prices in market $B$. Consumer surplus goes down on both markets.*

Intuitively, when firms are independent, the price of data depends on the outside option of not buying it and competing against a rival with a data advantage. $A$ therefore has an incentive to over-collect data to maximize the disadvantage of not buying the data. When

$A$ and $B_1$ merge, the value of the data on market $B$ is fully internalized and the incentive to over-collect data vanishes. Because the incentives to collect data are weakened under the merger, $A$ chooses to serve fewer consumers and the price of product $A$ is higher than with independent firms. Moreover, by Lemma 1(iii), consumer surplus on market $B$ goes down as well: in both cases only $B_1$ has the data, but under the merger it has less data than when firms are independent.

Propositions 1 and 2 reveal that the potential non-tradeability of data has significant implications for the effects of a data-driven merger, causing the consumer welfare effects to reverse. We next turn our attention to another property of data (the diversity of its economic uses) to see how this interacts with the effects of tradeability.

# 5 A more general model: various uses of data

Besides imposing a functional form for demand, Section 4 assumed a particular role of data (increasing product quality) and business model (price competition). But the data economy is characterized by diversity in both dimensions. Data is used to improve algorithms and recommendations, personalize products, target advertisements, resolve information asymmetries, and price discriminate. Meanwhile, firms in the digital economy offer a mix of ad-supported free products and products sold for a positive price.

What are the implications of this variety? The main results (Propositions 1 and 2) do not depend on the specific modelling choices, but rather on the three properties highlighted in Lemma 1. So the question is how those properties vary across settings. Intuitively, parts (ii) and (iii) of that lemma, namely that a firm is harmed as its rival gets more data and that consumers benefit from firms having more data, stem from the fact that data leads firms to offer more utility to consumers. In other contexts (e.g., when data is used to price discriminate) the main role of data may be to facilitate the extraction of consumers' surplus. One might then expect a firm with more data to leave less utility to consumers, reversing the effect of data on consumer surplus and the rival's profits. In this section we present a more general framework of data use in oligopolistic settings that allows us to discuss the effects of a data-driven merger in settings where parts (ii) and (iii) of Lemma 1 do not hold. This framework was introduced in de Cornière and Taylor (2023).

More specifically, we use Armstrong and Vickers (2001)'s model of competition in utility. This approach involves abstracting from firms' specific strategic choices—such as price, quality, or number of ads to show. Instead, firms are modelled as directly choosing how much utility to offer to consumers. Their choice of utility then determines (i) how attractive the firm is to consumers, and (ii) how much they earn from serving each consumer. This approach nests various underling models of strategic competition and we provide some concrete examples below. Since our aim is to be able to study different situations and business models in which data plays a role, the advantage of the competition in utilities approach is that it nests a variety of business models and strategic choices rather than being tied to any one specific strategic situation.

## 5.1 A model of competition in utility

On market $A$, firm $A$ offers a utility $u_A$, generating an increasing demand $D_A(u_A)$, and a profit in market $A$ of $\pi_A(u_A)$, which we assume is quasi-concave, maximized where $\pi'_A(u_A) = 0$. The quantity of data collected by $A$ is indexed by $\delta_A = D_A(u_A)$.

On market $B$ firm $B_j$ chooses a level of utility $u_j$, leading to a demand $D_B(u_j, u_{-j})$, non-decreasing in its first argument and non-increasing in the second one. For example, such a demand function can be generated using a discrete choice framework: suppose surplus for a consumer who buys from firm $B_j$ is $u_j + \epsilon_j$, where $\epsilon_j$ is a random taste shock specific to each consumer-firm pair, and $\epsilon_0$ is the outside option. Then if the total mass of consumers is normalized to one we would have $D_B(u_j, u_{-j}) = \Pr(u_j + \epsilon_j > \max\{\epsilon_0, u_{-j} + \epsilon_{-j}\})$.

Firm $B_j$'s *per-consumer* profit, if it offers a utility $u_j$, is denoted $r(u_j, \delta_j)$, where $\delta_j$ is an index measuring the quantity of firm $j$'s data. The main assumption is that $r(u_j, \delta_j)$ is increasing in $\delta_j$: data enables a firm to generate a higher profit for a given utility level. We also assume that firm $B_j$'s profit, $r(u_j, \delta_j)D_B(u_j, u_{-j})$, is differentiable and quasi-concave in $u_j$, with the best-response $u_j$ given by the first-order condition. We denote by $\pi_B(\delta_j, \delta_{-j})$ the equilibrium profit of $B_j$ for a given data allocation.[14] We maintain the assumption that $\frac{\partial \pi_B(\delta_j,\delta_{-j})}{\partial \delta_j} \geq 0$, which rules out situations where an increase in the data held by $B_j$ would trigger such an aggressive response by firm $B_{-j}$ that firm $B_j$'s profit would decrease. Lastly, we also assume that $\frac{\partial \ln[D_B(u_j,u_{-j})]}{\partial u_j \partial u_{-j}} \geq 0$, so that utilities are strategic complements.[15]

The timing is the same as in the previous section: At $t = 0$, the merger between $A$ and $B_1$ takes place or not. At $t = 1$, $A$ chooses a utility level and consumers decide whether to use its product. At $t = 2$, $A$ collects the data, and trade takes place if possible. At $t = 3$, $B$ firms observe the quantity of data held by their rival, choose a utility level, and consumers choose which $B$-market product to use. We look for subgame perfect equilibria.

**Examples** One advantage of the competition-in-utility approach is that it allows us to consider different business models and different ways to use data. Let us illustrate this point with two simple examples of the kind of situations in market $B$ that the model nests.

[14]The results would go through with $n \geq 2$ firms in market $B$, albeit at the cost of more cumbersome notation. We can also easily extend the framework to accommodate asymmetries between $B$-market firms. Simply add a $j$ subscript to $B_j$'s profit function, $\pi_{B_j}(\delta_j, \delta_{-j})$, with the corresponding underlying idiosyncratic demand and revenue functions, $D_{B_j}(u_j, u_{-j})$ and $r_{B_j}(u_j, \delta_j)$. The results below then go through unchanged, except in the case that $A$ supplies data exclusively to $B_2$ before merging with $B_1$. In that case the merger may lead the integrated firm to switch to supplying $B_1$ with the data.

[15]To see that this condition guarantees strategic complementarity, note that the standard necessary and sufficient condition is that the cross-derivative of profits is positive, $\frac{\partial^2}{\partial u_j \partial u_{-j}} r(u_j, \delta_j) D_B(u_j, u_{-j}) > 0$. The firm's first-order condition, (6), can be used to eliminate $r(u_j, \delta_j)$ from this cross-derivative, resulting in the condition that depends only on log-demand.

The goal is to take a standard model of strategic competition and rewrite it such that firms' profits can be expressed as the product of per-consumer revenue and demand, both a function of utility.

First, we show how the model of quality improvement from Section 4 can be reformulated in the competition in utility framework. In that example, the revenue per-consumer earned by a firm is simply its price, $r(u_j, \delta_j) \equiv p_j$. Inverting the utility function, $u_j = V(\delta_j) - p_j$, the per-consumer profit is therefore $r(u_j, \delta_j) \equiv p_j = V_B(\delta_j) - u_j$. Additionally, notice that the standard Hotelling demand, given in (1), can be equivalently expressed in terms of utilities as

$$D_B(u_j, u_{-j}) = \frac{1 + u_j - u_{-j}}{2}. \tag{5}$$

Thus, Section 4's price competition problem, $\max_{p_j} p_j D_B(p_j, p_{-j})$, is transformed into the equivalent competition in utilities problem, $\max_{u_j} r(u_j, \delta_j) D_B(u_j, u_{-j})$.

An advantage of the competition in utilities approach is that it applies more generally than just situations of price competition. As a second example, suppose that the two $B$ firms are free-to-access advertising-supported media companies who compete in advertising load, à la Anderson and Coate (2005). The utility offered by firm $B_j$ is $u_j = V - \gamma n_j$, where $n_j$ is the number of ads shown by $B_j$ and $\gamma$ is an ad-nuisance parameter. Suppose that data allows firm $B_j$ to serve targeted ads, which generate more revenue: the price of an ad on firm $j$ is $P(n_j, \delta_j) = 1 - n_j + \delta_j$. The per-consumer profit is $P(n_j, \delta_j) n_j$, which we can rewrite as $r(u_j, \delta_j) = \left(1 - \frac{V - u_j}{\gamma} + \delta_j\right) \frac{V - u_j}{\gamma}$. Paired with a demand function expressed in terms of utilities,[16] we have thus again transformed the firms' problem into one of competition in utilities, $\max_{u_j} r(u_j, \delta_j) D_B(u_j, u_{-j})$.

These examples are the simplest ones for our purpose here. De Cornière and Taylor (2023) provide informational microfoundations to these and other examples that can be nested in our framework.

## 5.2 Competition on the $B$ market

Consider the subgame starting at $t = 3$. Let $\widehat{u}_j(u_{-j}, \delta_j)$ denote firm $B_j$'s reaction function. We say that data is *unilaterally pro-competitive* (UPC) if $\widehat{u}_j(u_{-j}, \delta_j)$ is increasing in $\delta_j$, and unilaterally anti-competitive (UAC) if the reverse holds. In words, data is UPC if it induces a firm to become more aggressive, responding to a given rival strategy with a higher utility offer of its own. It is UAC if the opposite is true.

$B_j$'s best-response, $\widehat{u}_j(u_{-j}, \delta_j) = \text{argmax}_{u_j} \Pi(u_i, u_j, \delta_i) \equiv r(u_j, \delta_j) D_B(u_j, u_{-j})$, satisfies

$$\frac{\partial r(u_j, \delta_j)}{\partial u_j} D_B(u_j, u_{-j}) + r(u_j, \delta_j) \frac{\partial D_B(u_j, u_{-j})}{\partial u_j} = 0. \tag{6}$$

[16] For example, if competition is à la Hotelling then we would have the demand function given in (5) (with $u_j = V - \gamma n_j$). Alternatively, we could generate $D_B(u_j, u_{-j})$ using the discrete choice approach mentioned above, in which case $D_B(u_j, u_{-j}) = \Pr(V - \gamma n_j + \epsilon_j > \max\{\epsilon_0, V - \gamma n_{-j} + \epsilon_{-j}\})$.

It is a standard result of comparative statics that the best-response choice of $u_j$ is increasing in $\delta_j$ if and only if $\frac{\partial \Pi(u_j, u_{-j}, \delta_j)}{\partial u_j \partial \delta_j} > 0$, which, using (6) to eliminate terms involving $D_B$, is equivalent to $\frac{\partial^2 \ln[r(u_j, \delta_j)]}{\partial u_j \partial \delta_j} > 0$ (de Cornière and Taylor, 2023). Such a property is satisfied by the product improvement model. By contrast, in the targeted advertising model presented above data is UAC.

What are the effects of an increase in $\delta_j$ on firm $B_{-j}$ and on consumers? If data is UPC, $u_j$ will increase. This reduces the demand for firm $B_{-j}$, and therefore its profit. The effect on consumers is positive: not only does firm $B_j$ offer more utility, but, by strategic complementarity, firm $B_{-j}$ will also improve its offer. The reverse holds when data is UAC: in that case firm $B_j$ uses the extra data to extract more surplus from consumers, which increases the demand for $B_{-j}$ but harms consumers. Summarizing:

**Lemma 3.** *Data is UPC if and only if* $\frac{\partial^2 \ln[r(u_j, \delta_j)]}{\partial u_j \partial \delta_j} > 0$. *When data is UPC, an increase in* $\delta_j$ *causes* $B_{-j}$*'s profit to decrease and consumer surplus to increase. When data is UAC, an increase in* $\delta_j$ *causes* $B_{-j}$*'s profit to increase and consumer surplus to decrease.*

Lemma 3 gives us a way to connect environments with different uses of data or different business models to the properties in Lemma 1, which are crucial for determining the effects of a merger. *Any* environment in which data is UPC will replicate the strategic properties in Lemma 1. When data is UAC, however, the foundational properties in Lemma 1 change, and this affects the impact of a merger.

## 5.3 Effects of the merger

**When data trade is impossible** For this case the reasoning mirrors that in Section 4. Under the merger, at $t = 1$ the merged firm maximizes $\pi_A(u_A) + \pi_B(D_A(u_A), 0)$, instead of $\pi_A(u_A)$ when the firms are independent. Because $\pi_B(\delta_j, \delta_{-j})$ is increasing in $\delta_j$, the merger leads to more data collection, implying a higher $u_A$, which benefits market $A$ consumers.

The effect on market $B$, however, depends on whether data is UPC or UAC. In the latter case, the use of data on market $B$ following the merger harms consumers.

**Proposition 3.** *In the competition-in-utility model, when data trade is impossible: (i) if data is UPC, the merger increases consumer surplus on both markets; (ii) if data is UAC, the merger increases consumer surplus on market* $A$ *but decreases it on market* $B$.

**When data trade is possible** Recall from Lemma 2 that the equilibrium allocation of data is independent of whether the merger takes place or not. Suppose that it's optimal to have $k$ firms with data and define $U_A(\delta)$ such that $D_A(U_A(\delta)) = \delta$. In words, $U_A(\delta)$ is the utility offer needed to attract enough consumers to accumulate $\delta$ units of data.

When firms are independent, $A$ collects $T^k(\delta)$ in revenue from data sales and chooses $\delta$ to maximize

$$\Pi_A^I(\delta) = \pi_A(U_A(\delta)) + kT^k(\delta) \tag{7}$$

Under the merger, the integrated firm maximizes the sum of its profit on market $A$, the revenue from data sales to $B_2$ (if $k = 2$), and the profit of $B_1$. The objective is therefore to maximize

$$\Pi^M_{AB_1}(\delta) = \pi_A(U_A(\delta)) + (k-1)T^k(\delta) + \pi_B(\delta, (k-1)\delta). \tag{8}$$

Comparing (8) and (7), the difference between the two is $\pi_B(\delta, (k-1)\delta) - T^k(\delta) = \pi_B(\delta, (k-1)\delta) - (\pi_B(\delta, (k-1)\delta) - \pi_B(0, \delta)) = \pi_B(0, \delta)$. The marginal value of data is therefore higher under the merger if and only if $\pi_B(0, \delta)$ is increasing in its second argument, that is if and only if data is UAC (by Lemma 3). In this case, the merger leads firm $A$ to offer more utility to collect more data, which benefits consumers on $A$. However, consumers on $B$ are harmed by the increased quantity of data available to firms. When data is UPC, as in the quality improvement model, the merger leads $A$ to collect less data, which harms consumers on both markets. We thus have the following:

**Proposition 4.** *In the competition-in-utility model, when data trade is possible: (i) if data is UPC, the merger decreases consumer surplus on both markets; (ii) if data is UAC, the merger increases consumer surplus on market $A$ but decreases it on market $B$.*

Table 1: Effect of a data-driven merger on the utility offered on market $A$ ($u_A$), and the utility offered on market $B$ ($u_B$).

| | data is UPC | data is UAC |
|---|---|---|
| Pre-merger data trade | $\downarrow u_A, \downarrow u_B$ | $\uparrow u_A, \downarrow u_B$ |
| No pre-merger data trade | $\uparrow u_A, \uparrow u_B$ | $\uparrow u_A, \downarrow u_B$ |

Table 1 summarizes our results, accounting for both the (non-)tradeability of data and its differing competitive effects arising from different ways of using data. Both properties of data interact in determining the effects of a merger.

# 6 Strategic consumers

In the analysis so far we have assumed that consumption decisions on market $A$ do not depend on how data is used on market $B$. While this may be a reasonable assumption in various contexts,[17] we now consider the alternative where consumers are fully strategic.

Suppose that data collected about each of firm $A$'s customers can be used by firms on market $B$ to make personalized offers. We assume that there are no data externalities, so that the surplus offered to a consumer on market B only depends on whether firms

[17]For instance, it may be that consumers are atomistic and the data about one consumer does not affect the quality of the product, or that consumers fail to understand how data is transferred across firms.

have access to data about that consumer.[18] Firm $A$ sets a price $p_A$ and collects a unit of data about each of its customers. As in the general model, firm $B_j$ offers utility $u_j$. Denote $B_j$'s equilibrium utility offer by $u(\delta_j, \delta_{-j})$, where $\delta_j = 1$ if firm $B_j$ has access to the consumer's data, and 0 otherwise.

**No pre-merger trade** First suppose that pre-merger data trade is impossible. Prior to the merger, $A$ focuses only on its primary market, solving $\max_{p_A} \{p_A D_A(p_A)\}$. Let $p_A^I$ be the solution. Consumer surplus on market $B$ is $U_B^I = E[\max\{\epsilon_0, u(0,0) + \epsilon_1, u(0,0) + \epsilon_2\}]$, where the $\epsilon$ are idiosyncratic taste shocks (and $\epsilon_0$ is the value of the outside option for the consumer).

Post-merger, the data will be shared with $B_1$. The use of data by $B_1$ results in a change in the consumer's equilibrium $B$-market utility. Let $\beta \equiv E[\max\{\epsilon_0, u(1,0) + \epsilon_1, u(0,1) + \epsilon_2\}] - U_B^I$ be the additional $B$-market utility resulting from data use ($\beta > 0$ if data is UPC and $\beta < 0$ if data is UAC). Because a consumer can always obtain $U_B^I$ by not buying $A$, a price $p_A$ generates a demand $D_A(p_A - \beta)$. Meanwhile, firm $A$ anticipates an extra $\pi_B(1,0) - \pi_B(0,0)$ profit from handing data to its subsidiary. The new problem for firm $A$ is therefore $\max_{p_A} \{[p_A + \pi_B(1,0) - \pi_B(0,0)]\, D_A(p_A - \beta)\}$. We can rewrite it as firm $A$ choosing an effective price $\widetilde{p_A} = p_A - \beta$: $\max_{\widetilde{p_A}} \{[\widetilde{p_A} + \beta + \pi_B(1,0) - \pi_B(0,0)]\, D_A(\widetilde{p_A})\}$. Consumer surplus increases with the merger if and only if the optimal effective price is below $p_A^I$, which is the case if $\beta + \pi_B(1,0) - \pi_B(0,0) > 0$.

If the profit generated by data is larger than the loss in surplus on market $B$ (i.e. $\beta + \pi_B(1,0) - \pi_B(0,0) > 0$), firm $A$ shares the extra surplus with consumers so as to attract more of them. If, on the other hand, the disutility is larger than the extra profit, the merger is equivalent to increasing the marginal cost of product $A$, which harms consumers and the firm. Because in that case the merger would not be profitable, we assume that it would not take place, so that a profitable merger would always benefit consumers.

**Pre-merger trade is possible** By Lemma 2, we know that the merger will not change which firm(s) use the data in market $B$. Therefore consumers' decision regarding the purchase of $A$ is unaffected by the merger, conditional on the price of $A$. Compared to the baseline model, introducing strategic consumers simply shifts the demand function for $A$ by $\beta$,[19] but the change in incentives to collect data is governed by the same forces as in the baseline model: the merger leads to an increase in $p_A$ if data is UPC, and to a decrease in $p_A$ otherwise.

Consumers face the same decision both pre- and post-merger: whether to spend an effective $p_A - \beta$ to get the surplus from consuming product $A$. Since the allocation of

[18] Introducing externalities would require additional notations without fundamentally changing the insights.

[19] In this case, $\beta \equiv E[\max\{\epsilon_0, u(1,1) + \epsilon_1, u(1,1) + \epsilon_2\}] - E[\max\{\epsilon_0, u(0,0) + \epsilon_1, u(0,0) + \epsilon_2\}]$ if data is sold non-exclusively, and is equal to the expression in the previous paragraph otherwise.

data—and hence $\beta$—does not change with the merger, the change in $p_A$ determines the overall change in consumer surplus, so that the merger benefits consumers if and only if data is UAC. This perhaps counterintuitive result is explained as follows: because the value of data is higher with the merger when data is UAC, the firm will seek to collect more of it. To do so, it will have to compensate consumers for their lower surplus on $B$ by reducing $p_A$ enough.

We summarize the above discussion with the following proposition:

**Proposition 5.** *Suppose that consumers are strategic and that there are no data externalities across consumers. When data trade is impossible, a (profitable) merger increases consumer surplus. When data trade is possible, the merger benefits consumers if and only if data is UAC.*

# 7 Discussion and policy implications

We have provided an analysis of data-driven mergers in a model where data is an input with the following properties: (i) it is collected as a byproduct of a firm's activity in one market and used on another; (ii) it is non-rival; (iii) it has many potential uses, which may benefit or harm consumers; (iv) frictions may prevent its trade.

When data is unilaterally pro-competitive, that is in markets where data is used to the benefit of consumers, our analysis offers both an efficiency argument in favor of a data-driven merger (it enables data uses in adjacent markets, Proposition 3 (i)) and a new, non-exclusionary, theory of harm (the merger reduces incentives to collect data, leading to a lower utility in the primary market, Proposition 4(i)). The key condition is whether data trade is possible absent the merger.

Absent pre-merger trade, and if there are no indications that such trade might take place in the near future, it is important to identify the source of the friction: a merger allowing firms to bypass regulations may undermine other public objectives, and the efficiency argument should be given less weight. One could even interpret the existence of regulations as an indication that the use of data does not increase consumers' utility, an argument in favour of blocking the merger. If, however, trade of data is hindered by other types of (e.g., contracting) frictions, our analysis suggests that the merger is more likely to benefit consumers.

Suppose, on the other hand, that pre-merger data trade occurs. Authorities should then obviously lend less credence to the above efficiency argument. However, before accepting the theory of harm that we have proposed, several conditions should be checked: First, firm $A$ must have enough market power on the market for data. Second, data trade is (or has the potential to be) an important part of firm $A$'s activity. Indeed, the main driving force of our result is that the incentive to manipulate the price of data is strong enough to affect $A$'s behaviour in its primary market. Third, the value of the dataset of firm $A$ depends positively on the utility it offers to its primary customers. The idea is that a firm offering higher utility attracts more consumers and therefore gathers more data.

When data is unilaterally anti-competitive, that is when it is used in exploitative ways, the possibility of trade is irrelevant: in either case the mergerś effect is positive on market $A$ and negative on $B$. We must therefore consider the merger's welfare implications not only via the use of data in market $B$, but also by distorting the transactions in market $A$ that generate data as a by-product. The key question is then whether efficiencies on one market ($A$) can be used to offset anti-competitive effects on another ($B$).

Note that, if a competition authority uses a "separate effects" approach, that is does not allow consumer losses in one market to be offset by gains on another, then our model indicates that a sufficient condition for the merger to be anticompetitive is that data trade is possible absent the merger, an easily observable characteristic of the market.

Let us now briefly remark on two assumptions that underlie the analysis. The first is that data collection necessarily benefits consumers on market $A$, through the improved offers necessary to attract more consumers. Under an alternative scenario, firm $A$ could harvest data from its existing customers through privacy-invading technologies. If consumers have strong privacy concerns we could then have $U'_A(\delta) < 0$. In such a case, the implications of the merger on market B are unchanged, but those on market A are reversed, so that the effects of the merger on each market are of opposite sign when data is UPC, while the merger is always harmful when data is UAC.

Secondly, we have assumed that the firm on market A is a monopolist. There are two aspects to this: it does not face competition when it comes to collecting data, nor when it comes to selling data. Relaxing this assumption might affect some results, and should be the subject of future research.

# References

Anderson, Simon P. and Stephen Coate (2005). "Market Provision of Broadcasting: A Welfare Analysis". *The Review of Economic Studies* 72.4, pp. 947–972.

Armstrong, Mark and John Vickers (2001). "Competitive price discrimination". *RAND Journal of Economics* 4, pp. 1–27.

Autorité de la Concurrence and Bundeskartellamt (2016). "Competition Law and Data".

Bergemann, Dirk and Alessandro Bonatti (2019). "Markets for Information: An Introduction". *Annual Review of Economics* 11.1, pp. 85–107.

Bounie, David, Antoine Dubus, and Patrick Waelbroeck (2021). "Selling strategic information in digital competitive markets". *The RAND Journal of Economics* 52.2, pp. 283–313.

Bounie, David, Antoine Dubus, and Patrick Waelbroeck (2023). "Competition Between Strategic Data Intermediaries with Implications for Merger Policy". *CER-ETH Working Paper 23/383*.

CMA (2019). "Online platforms and digital advertising: Market study interim report". URL: https://assets.publishing.service.gov.uk/media/5dfa0580ed915d0933009761/Interim_report.pdf.

Chen, Zhijun, Chongwoo Choe, Jiajia Cong, and Noriaki Matsushima (2022a). "Data-Driven Mergers and Personalization". *RAND Journal of Economics* 53.1, pp. 3–31.

— (2022b). "Data-driven mergers and personalization". *The RAND Journal of Economics* 53.1, pp. 3–31.

Crémer, Jacques, Yves-Alexandre de Montjoye, and Heike Schweitzer (2019). *Competition Policy for the Digital Era*. Report. European Commission.

D'Annunzio, Anna (2017). "Vertical integration in the TV market: Exclusive provision and program quality". *International Journal of Industrial Organization* 53, pp. 114–144.

De Cornière, Alexandre and Greg Taylor (2023). "Data and Competition: A Simple Framework". *Working Paper*.

Esteves, Rosa-Branca and Helder Vasconcelos (2015). "Price discrimination under customer recognition and mergers". *Journal of Economics & Management Strategy* 24.3, pp. 523–549.

Fudenberg, Drew and Jean Tirole (1984). "The fat-cat effect, the puppy-dog ploy, and the lean and hungry look". *American Economic Review* 74.2, pp. 361–366.

Furman, Jason, Dianne Coyle, Amelia Fletcher, Derek McAuley, and Philip Marsden (2019). *Unlocking Digital Competition*. Report of the Digital Competition Expert Panel. HM Government.

Grunes, Alan and Maurice Stucke (2016). *Big data and competition policy*. Oxford University Press.

Hagiu, Andrei and Julian Wright (2023). "Data-enabled learning, network effects, and competitive advantage". *The RAND Journal of Economics* 54.4, pp. 638–667.

Herresthal, Claudia, Tatiana Mayskaya, and Arina Nikandrova (2022). "Data Linkage between Markets: Does the Emergence of an Informed Insurer Cause Consumer Harm?" *CEPR Discussion Paper 17947*.

Katz, Michael L and Carl Shapiro (1986). "How to license intangible property". *The Quarterly Journal of Economics* 101.3, pp. 567–589.

Kim, Byung-Cheol and Jay Pil Choi (2010). "Customer information sharing: Strategic incentives and new implications". *Journal of Economics & Management Strategy* 19.2, pp. 403–433.

Kim, Jin-Hyuk and Liad Wagman (2015). "Screening incentives and privacy protection in financial markets: a theoretical and empirical analysis". *The RAND Journal of EConomics* 46.1, pp. 1–22.

Kim, Jin-Hyuk, Liad Wagman, and Abraham L. Wickelgren (2018). "The impact of access to consumer data on the competitive effects of horizontal mergers and exclusive dealing". *Journal of Economics and Management Strategy.*

Lefouili, Yassine, Leonardo Madio, and Ying Lei Toh (forthcoming). "Privacy Regulation and Quality-Enhancing Innovation". *The Journal of Industrial Economics.*

Montes, Rodrigo, Wilfried Sand-Zantman, and Tommaso Valletti (2018). "The Value of Personal Information in Online Markets with Endogenous Privacy". *Management Science.*

Pino, Flavio (2022). "The microeconomics of data – a survey". *Journal of Industrial and Business Economics*, pp. 635–665.

Spagnolo, Giancarlo, Otto Toivanen, Nicolas Schutz, Michelle Sovinsky, Thomas Rønde, Monika Schnitzer, Paul Heidhues, Cristina Caffarra, Tomaso Duso, Christos Genakos, Gregory Crawford, Chongwoo Choe, Zhijun Chen, Tommaso Valletti, Martin Peitz, Thibaud Vergé, and Marc Bourreau (2020). "Google/Fitbit will monetise health data and harm consumers". VoxEU. URL: `https://cepr.org/voxeu/columns/googlefitbit-will-monetise-health-data-and-harm-consumers`.

Wickelgren, Abraham L. (2015). "An economic analysis of internet privacy regulation". Working Paper.