Competing for cookies: Platforms' business models in data markets with network effects*

Sarit Markovich† and Yaron Yehezkel‡

April 2025

Abstract

We consider platform competition when platforms can either 1) commercialize users' data and in return offer their services for free (data-based business model); 2) protect users' privacy and charge users for participation (subscription-based model); or 3) offer both options (the hybrid model). We find that competition does not always motivate the incumbent platform to protect users' privacy. When network effects are intermediate, competition can motivate the incumbent to shift from the subscription-based model to the hybrid model; thereby, increasing data commercialization. Yet, the opposite case occurs when network effects are weak. Moreover, allowing the incumbent to adopt the hybrid model is welfare enhancing when network effects are strong, and welfare reducing (or neutral) otherwise.

JEL Classification: L1

Keywords: platforms with network effects, business models, data commercialization.

1 Introduction

In recent years, users' data have become an important asset and an essential element of platforms' strategy. Platforms collect personal consumer information (using cookies and other means) and use it to improve the quality of their service as well as for commercialization

*For helpful comments a discussions we thank the editor Alfonso Gambardella, the associate editor, three anonymous referees, Ben Casner, Federico Etro, Leonardo Madio and seminar participants at the Israeli strategy and entrepreneurship May 2024 seminar, the IIOC 2024, ECOP 2024, CRESSE 2024, the BU Platform Strategy Research Symposium 2024, EARIE 2024, Toulouse Economics of Platforms Seminar and the Israeli Competition Authority. For financial support, we thank the Coller Foundation, the Henry Crown Institute, the Eli Hurvitz Institute on Strategic Management and the NET Institute–www.NETinst.org.

†Kellogg School of Management, Northwestern University (s-markovich@kellogg.northwestern.edu)

‡Coller School of Management, Tel Aviv University (yehezkel@tauex.tau.ac.il)

purposes, such as selling it to third party vendors or to advertisers. This trend is particularly evident among online platforms like Google, Facebook, TikTok, and Spotify which leverage their large stocks of consumer information to enhance their products and offer free services in exchange for data. Not all platforms, however, base their business strategy on their users data. Platforms like OpenAI, Apple Health, and Ride with GPS,[1] rely on subscription revenues rather than data commercialization. Similarly, Netflix has long followed a subscription-based business model for many years and has only recently introduced a data-based option. This gives rise to a third business model we observe – a hybrid model where users can choose between sharing their data and enjoying the service for free, or paying a subscription fee. Likewise, Facebook recently adopted the hybrid model in Europe, by adopting a subscription-based option.

The variance in platforms' choice of business model raises the questions: what determines platforms' choice of business model? How does competition affect this choice? Specifically, does it motivate platforms toward more privacy-focused models that avoid data commercialization? Moreover, should platforms be prohibited from discriminating between users who share their data and those who do not, as seen in the hybrid business model?

To study these questions, we develop a game with two platforms and users that care about their privacy—i.e., bear a cost if their data is commercialized. Users' disutility from the commercialization of their data differs across users. That is, some users are more sensitive to their privacy than others. Users who join a platform enjoy the network effects generated by all other users on that platform. These network effects can be viewed as the benefit derived from the data the platform collects on all users, and used to improve the quality of the service the platform offers (Markovich and Yehezkel, 2024). For example, Netflix leverages data on users' viewing habits and preferences to refine its recommendation algorithm. The data collected from all users improves features like "Play Something," enabling Netflix to better predict and suggest content a user is likely to enjoy. Alternatively, the network effects may stem from the direct interactions that users can have with one another on the platform.

Platforms can choose between three business models: (1) data-based; (2) subscription-based; and (3) hybrid. Under the data-based business model, the platform generates revenue from commercializing its users, either by selling their data to third-party providers or by monetizing their time and attention, such as through advertisements. We define the data-based business model broadly to include both the direct sale of user data and the indirect commercialization of user engagement through advertisements. For clarity, however, our discussion will focus on the case where the data-based model corresponds to directly selling users' data. Users that join the platform know that their data will be commercialized and

[1]Ride with GPS is a social route-planning and navigation tool for cyclists.

the cost this would impose on them.

Under the subscription-based business model, users must pay a subscription fee to participate in the platform, yet know that their data will not be commercialized. The hybrid business model combines the two first business models. That is, the platform allows users to choose whether they want to join the platform for free and share their data, knowing the data would be commercialized. Alternatively, users can pay the subscription fee, in which case their data will not be commercialized. For example, Google, Facebook, TikTok, and X utilize the data-based business model, while Apple has been an avid advocate of the subscription-based model. Likewise, the social apps True and Mastodon, the messaging app Signal, and the search engine DuckDuckGo, explicitly chose not to commercialize their users' data.[2] The hybrid model has become more popular, recently, where OpenAI shifted in February 2023 from a data-based model to a hybrid model, while later that year, in November 2023, Meta launched in Europe a no-ads Facebook subscription service. Accordingly, users can choose between a free service by agreeing to have their data tracked and commercialized through advertising, or choose a subscription model which protects their privacy and offers an ad-free experience. Facebook's shift is controversial in Europe. A coalition of 28 organizations has called for an investigation of this business model, arguing that Meta essentially asks users to pay for their privacy.[3]

In order to capture the advantage that a large, dominant platform may have, we assume a two-stage game with an incumbent and an entrant, where the incumbent enjoys a focality advantage–users believe that the incumbent would be the dominant platform in the market.

We find that platforms' optimal business model depends on the strength of network effects. Specifically, if network effects are strong, the platform chooses the subscription model when the commercial value of data is low and the data-based model when it is high because the market is fully covered under all models, in which case the platform cannot benefit from combining both plans and offering the hybrid model. If network effects are intermediate, an incumbent platform should adopt the hybrid model as it allows it to dominate the market by attracting both privacy sensitive users – with the subscription plan – and privacy insensitive users – with the data plan. When network effects are weak, the incumbent should avoid fierce competition. Hence, the incumbent focuses on the privacy-sensitive users and adopts the subscription-based model which then allows the entrant to differentiate itself and attract with the data-based model users that are not as sensitive to their privacy.

The effect of competition on platforms' optimal business model also depends on the

[2]True plans on making money by charging users for subscription. Mastodon relies on decentralization, Signal on donations, and DuckDuckGo on keywords, rather than targeted, advertising.

[3]See CPI, February 18, 2024. Available at: https://www.pymnts.com/cpi_posts/privacy-advocates-urge-european-regulators-to-oppose-metas-no-ads-subscription-model/

strength of the network effects. In particular, if the commercial value of data is very high or very low, competition has no effect on the incumbent's choice of business model. Intuitively, the incumbent prefers the hybrid model when data has a high commercial value and the subscription-based model when data value is very low, under both monopoly and competition. However, if the commercial benefit of data is moderate, the effect varies with the strength of the network effects. When network effects are weak, competition incentivizes a monopolistic incumbent to move away from the hybrid model and promote privacy by choosing the subscription-based model. In this case, the entrant adopts the data-based model. While the overall choice of plans from the users' perspective remains unchanged, competition may still suppress data collection because the price of the subscription plan under competition is lower than under monopoly, resulting in more users choosing the subscription plan and thus a decreases in data commercialization.

When network effects are intermediate, competition incentivizes a monopolistic incumbent to shift from the subscription based model to the hybrid one. In this case, not only more data is commercialized compared to the monopolistic outcome, but the incumbent is also able to attract all users and monopolize the market. That is, competition does not necessarily promote a more privacy-sensitive market.

To study the welfare implications of prohibiting platforms from discriminating across users who share their data and those who do not, we compare equilibrium market structure when platforms are allowed to offer the hybrid business model to the structure when this model is banned. This latter case corresponds to the EU General Data Protection Regulation (GDPR) which prohibits such discrimination. We find that the effect of the ability to offer a hybrid business model on welfare largely depends on the strength of network effects. When network effects are strong, the availability of the hybrid model enhances welfare because there is a social value in having all users on the same platform. Yet, when network effects are moderate and the commercial benefit of data is high, the hybrid model can reduce welfare because in this parameter space there is less data collection under the hybrid model. Finally, when network effects are weak, the hybrid model is not attractive for the platforms and in equilibrium, none of the platforms offers this model to users.

Our paper is mainly related to the literature on how competition shapes platforms' business models. Casadesus-Masanell and Hervas-Drane (2015) study a competitive market where firms compete in prices and qualities, which can be interpreted as privacy. They show that competition increases privacy, but greater competition intensity does not always enhance privacy further. while increasing competition intensity does not necessarily imply that privacy is further improved. They also show that low privacy firms tend to subsidize consumers, while high privacy firms charge positive prices. Calvano and Polo (2020) study the business

models of two competing platforms that connect between viewers and advertisers, such as a TV channel. They show that ex-ante identical platforms can strategically differentiate themselves by selecting different business models, such that each platform's revenues come from a different side of the market. In a closely related paper, Llanes and Madio (2024) consider a monopolistic AI platform that can adopt a subscription-based model in which it charges a price for its services, a data-monetization business model in which it monetizes the AI's algorithm's quality, or a freemium model in which it offers both options. Users differ in their willingness to pay for the platform's services (with a fraction of users that are only willing to join the free service). In another closely related paper, Casner and Teh (2024) consider platforms that can adopt a "pure discovery" business model, that generates per-viewer ad revenue for the platform, a "pure membership" model, where content creators charge users a fee from which the platform takes an ad-valorem transaction commission, or a hybrid model that combines the two. We contribute to this literature by explicitly modeling network effects between users and studying their effect on the equilibrium business model. Moreover, in our model the difference in business models is driven by users' heterogeneity in their disutility from privacy. Assuming a monopolistic platform, Markovich et. al (2024) study a general model with heterogenous users and network effects. The platform offers a continuous menu of options, varying in price and the extent of data commercialization. Focusing on the effect of the platform's popularity (focality) on the platform's choice of menu, they find that even when the platform can offer a plan with a positive price and positive level of data commercialization, the platform opts to offer a free plan solely based on data commercialization. In a market without network effects, Chen (2025) considers platforms that collect and commercialize user data which generate commercial benefit and also improve the service to the user only. The paper studies optimal business model under monopoly and competition, allowing platforms to charge a positive or negative price when collecting user data.

The economic literature on competing platforms (see Jullien et al., 2021, for a review of the literature) extends the work of Katz and Shapiro (1985) on competition with network effects, where the size of the network creates additional value to the customers. Jullien (2011), Hałaburda and Yehezkel (2013; 2016; 2019), and Markovich and Yehezkel (2022) consider platform competition and coordination in the context of a static game. Hagiu (2006) considers sequential competition on two sides of a market. Hałaburda et al. (2020) and Biglaiser and Crémer (2020) consider dynamic competition. Much of this literature focuses on the coordination problem and the role pricing plays in overcoming this problem by using a divide-and-conquer strategy where platforms compete in subsidizing one set of users in order to attract another set.

Our paper is also related to the literature on privacy and network externalities. Most

of this existing literature focuses on the negative externalities associated with users sharing their data where one user's data can help platforms learn and predict the behavior of other users who do not share their data (Fairfield and Engel, 2015; Choi et al., 2019; Acemoglu et al., 2022; Bergemann et al., 2022; Liang and Madsen, 2019). Following Markovich and Yehezkel (2024), our paper recognizes and focuses on the positive externalities—e.g., users that share data help the platform improve the quality of its product and offer higher value to other users. Fainmesser et al. (2022) study how a monopolistic platform's revenue model affects its data policy in terms of data collection and data protection. Considering the net value of network externalities (positive minus negative), they find that relative to the socially desired data strategy, the platform may over- or under-collect users' data and may over- or under-protect it. The authors then show that the inefficiency in data collection can be corrected with taxes or fines imposed on the firms. We add to this literature by focusing on competition and its effect on platforms' business models in terms of commercializing data or charging users for using the platform. O'Brien and Smith (2014) study a model where sellers can commit to privacy policies and consumers have heterogeneous – negative or positive – preferences over privacy. They find that under perfect competition, firms make the socially optimal decision. Furthermore, a positive and sufficiently large correlation between consumers' valuations for the product and privacy is a necessary condition for the under-supply of privacy by firms. Assuming a two-stage game where data accumulated in the first period can be used to customize products in the second stage, Ke and Sudhir (2023) find that in a perfectly competitive market, whether privacy rights lower or increase profits depends on the expected privacy breach costs. Our paper considers imperfect competition between an incumbent and an entrant platforms. We show how the strategic effect of competition and the threat of entry shape the incumbent's and the entrant's business models. Similar to our paper, Hagiu and Wright (2023) study competition between an incumbent and an entrant platform that collect data on their users. The focus of their analysis, however, is on data-enabled learning across- and within-users and on how a platform's competitive advantage is affected by the shape of the learning function.

Our paper is also related to the growing empirical literature studying the impact of the GDPR. Utilizing data from an online travel intermediary, Aridor et al. (2023) find that the GDPR has resulted in an immediate drop in the total number of advertisements clicked and a corresponding immediate decline in revenue. The remaining set of consumers, however, are higher value consumers to the advertisers, compared with the pre-GDPR set of consumers. Focusing on market concentration, Johnson et al. (2023) find that GDPR increased market concentration among technology vendors where the relative market shares of the largest firms—particularly, Google and Facebook—increase post-GDPR. Using data on apps at the

Google Play Store, Janssen et al. (2022) show that GDPR induced the exit and reduced entry of new apps by half, resulting in an overall reduced consumer surplus. We add to this literature by analyzing the effect of banning firms from the ability to using a hybrid business model which price discriminates between users that share their data and those who do not share their data for commercialization.

2 The Model

Consider two competing platforms, an incumbent, I, and an entrant, E, and a mass 1 of users. Each platform can collect data from users and can utilize the data for two benefits. The first is enhancing services to other users. This is the network effect of data and we denote it by β. For example, platforms like Google, Netflix, and Spotify use other users' data to improve the quality of their search and suggestion algorithms. Secondly, the platform can "commercialize the user" by, for example, selling their personal data to advertisers or other platforms. Alternatively, the platform can commercialize users by commercializing their time or attention, for example, with push advertisements. We refer to these options as the platform's "commercial benefit" and denote it by α. Users incur disutility when being commercialized , which we denote by k. User's k's utility from joining platform $i = I, E$ is:

$$U_{ki} = v + \beta n_i - C_i k - p_i, \tag{1}$$

where v is the base benefit from joining a platform,[4] n_i is the number of users that join platform i, $C_i = \{0, 1\}$ is the platform decision on whether not to commercialize the user's data ($C_i = 0$) or to commercialize ($C_i = 1$), in which case the user incurs a costs k. Finally, p_i is the platform's price. Suppose that users differ in their costs from being commercialized: some users are more sensitive to their privacy than others. Likewise, some users suffer higher disutility from observing ads than others. Hence, we assume that k is uniformly distributed on the interval $[0, 1]$.[5] We focus on the interesting case where when the platform commercializes users' data, the market is not fully covered, and thus for most of our analysis restrict the parameter space to: $v < 1$ and $0 < \beta < 1 - v < 1/2$. This parameter space rules out corner solutions. We consider the case of strong network effects, $1 - v < \beta$, in Section 7.

Each platform can choose between three business models: data-based that we denote by D, subscription-based that we denote by S and a hybrid model, denoted by H. In the data-based business model, $C_i = 1$: the use of the platform is free and its source of revenues

[4]Our analysis focuses on the effect of platforms' choice of business model on competition. In order to isolate this effect, we assume that both platforms offer the same base benefit.

[5]See Markovich et. al (2024) for the case where k follows a general distribution.

is from collecting and commercializing users' data. In this case, the platform's profit is $\pi_i(D, B_j) = \alpha n_i(D, B_j)$, where $n_i(D, B_j)$ is the number of users that join it given that platform j adopts business model $B_j = D, S, H$; and recall that $\alpha > 0$ is the data's commercial benefit to the platform. Under the "subscription based" business model, the platform commits not to commercialize users' data ($C_i = 0$) and instead charges users for participation and earns $\pi_i(S, B_j) = p_i n_i(S, B_j)$. The third, hybrid business model is a combination of the two: the platform allows users to choose between a subscription plan in which it commits not to commercialize the user's data and a free plan where it makes no such commitment and hence commercializes users' data. The platform's profit is $\pi_i(H, B_j) = \alpha n_{iD}(H, B_j) + p_i n_{iS}(H, B_j)$, where $n_{iD}(H, B_j)$ and $n_{iS}(H, B_j)$ are the number of users that join the free and subscription plans, respectively.

The timing is as follows. In the first stage, the incumbent chooses its business model: $B_I = D, S, H$. In the second stage, the entrant chooses its business model $B_E = D, S, H$. Then, in the third stage, the two platforms compete on users. As is usually the case in platform competition with network effects, in the third stage of the game there can be multiple equilibria, because each user's decision depends on the beliefs regarding the decisions of other users. To this end, we assume that the incumbent has a "focal" position in that whenever possible, users expect other users to join the incumbent. We elaborate on these beliefs in Section 4.

3 Benchmark: Monopoly

In order to build intuition, we start with a benchmark case in which the incumbent is a monopoly that does not face the threat of entry. The incumbent can choose between $B_I = S, D, H$. If the incumbent chooses the data-based business model, it announces that joining the platform is free and its intention to commercialize users' data. Users will join the incumbent as long as

$$v + \beta n_I - k \geq 0 \quad \Longleftrightarrow \quad n_I(D) = \frac{v}{1-\beta}. \tag{2}$$

Because by assumption $v < 1 - \beta$, not all users join the platform: data-sensitive users prefer to stay out. Yet, as network effects increase, more users join the platform in order to enjoy the network effects generated by other users. The incumbent earns $\pi_I(D) = \frac{\alpha v}{1-\beta}$. If the incumbent adopts the subscription-based business model, because the incumbent benefits from a focal position and users expect other users to join it, the incumbent can attract all users if $p_I \leq v + \beta$. Hence, the incumbent charges $p_I(S) = v + \beta$ and earns $\pi_I(S) = v + \beta$.

If the incumbent chooses the hybrid business model, the incumbent offers a data plan and a subscription plan and allows users to choose their preferred option. All users will join the incumbent and will prefer the data plan if $\beta + v - p_I \leq \beta + v - k$, or $k < p_I$. Hence, users with $k \in [0, p_I]$ join the data plan and the platform earns α on these users, while users with $k \in (p_I, 1]$ join the subscription plan and pay p_I. The platform, thus, sets p_I such that:

$$\max_{p_I} \pi_I(p_I|(H)) = \alpha p_I + (1 - p_I)p_I, \tag{3}$$
$$\text{s.t.} \quad \beta + v - p_I \geq 0 \text{ and } p_I \leq 1.$$

The following lemma summarizes the result. All proofs are in the Online Appendix A.

Lemma 1. *Suppose that the incumbent is a monopoly that adopts the hybrid model. Then, the incumbent charges and earns*

$$p_I(H) = \begin{cases} \frac{1+\alpha}{2}, & \text{if } \beta \geq \frac{1+\alpha}{2} - v, \\ \beta + v, & \text{if } \beta < \frac{1+\alpha}{2} - v, \end{cases} \tag{4}$$

$$\pi_I(H) = \begin{cases} \frac{(1+\alpha)^2}{4}, & \text{if } \beta \geq \frac{1+\alpha}{2} - v, \\ (1 + \alpha - \beta - v)(\beta + v), & \text{if } \beta < \frac{1+\alpha}{2} - v. \end{cases} \tag{5}$$

The intuition behind this result is as follows. When α is small, $p_I(H) = \frac{1+\alpha}{2} < \beta + v$ (which is equivalent to $\alpha < 2(\beta + v) - 1$), so the participation constraint is not binding. The price is increasing with α because the incumbent takes advantage of the high commercial value of data and sways users to choose the data-plan over the subscription-plan by charging a higher price for the subscription plan. Once $p_I(H)$ reaches $\beta + v$, the utility that users that join the subscription-plan receive reaches 0. In this case, the incumbent extracts all of the utility users that join the subscription-plan enjoy $(\beta+v)$, and the incumbent cannot keep increasing the price (as a function of α). Comparing all three models we have:

Proposition 1. ***(Monopolist optimal business model)*** *Suppose that the incumbent is a monopolist that can adopt $B_I = \{S, D, H\}$. Then, the incumbent adopts the hybrid model if $\alpha > \alpha^M_{H,S} = \beta + v > 2(\beta + v) - 1$ and adopts the subscription-based model otherwise. Moreover, $\alpha^M_{H,S}$ is increasing in β.*

Figure 1 shows the subscription-based and the hybrid business models. Under monopoly, the hybrid model is always more profitable than the data-based model. The intuition is that because $\beta < 1 - v$, under the data plan, the platform does not cover the entire market. Therefore, the platform can always benefit from offering in addition to the data plan, a

subscription plan targeted at the high cost users. As illustrated in the figure, in comparison with the hybrid model, the advantage of the subscription-based model is that the platform can collect the network effects from *all* users. In contrast, the advantage of the hybrid model is that the platform can collect the commercial benefit from low cost users. Therefore, the subscription-based model becomes more profitable as network effects increase and the commercial benefit decreases. This provides the intuition for why the threshold $\alpha^M_{H,S}$ is increasing with network effects.

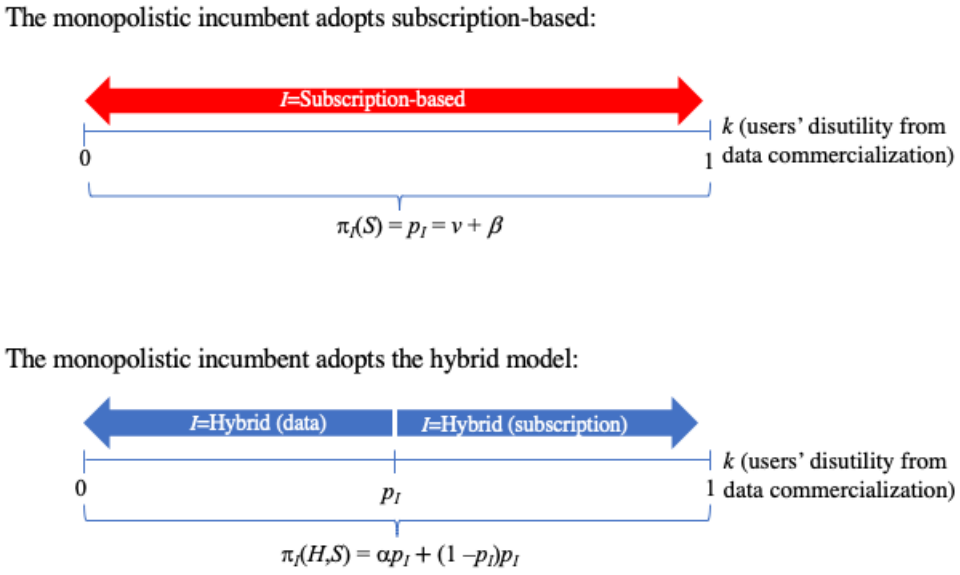

Figure 1: Market share and profits for H and S business models under monopoly

As we show in Section 7, when network effects are strong such that $\beta > 1 - v$, the incumbent can cover the market with the data-based model, and hence adopts this business model if the commercial benefit is high.

4 Platform competition

Moving to competition, we assume that the incumbent chooses its business model first, $B_I = D, S, H$. The entrant observes the incumbent's choice and chooses $B_E = D, S, H$. The platforms then compete for users. We solve the game backwards, and start by solving for the entrant's response to each business model that the incumbent can adopt: starting with the hybrid model, and then the cases where the incumbent adopts the data-based and subscription-based models.

The incumbent adopts the hybrid model

Suppose that the incumbent adopts the hybrid model. The incumbent announces that users can either join for free, conditional on giving their consent to have their data commercialized, or pay a price, p_I, and have their data protected.

We show that in equilibrium, the incumbent dominates the market. The entrant's optimal response is to adopt the subscription-based model and offer it for free. Doing so provides users with the highest alternative utility relative to the utility from joining the incumbent. As the incumbent benefits from a focal position, users expect that all other users join the incumbent, and users' utility from joining the entrant is $0 \times \beta + v - p_E = v$.

Turning to the incumbent, given the price of the subscription plan, p_I, users who join the incumbent choose the subscription plan if $\beta+v-p_I \geq \beta+v-k$, or $k \geq p_I$. Given p_I, users with $k \in [0, p_I]$, i.e., data-insensitive users, join the free plan and the incumbent commercializes their data and earns αp_I. Data-sensitive user with $k \in [p_I, 1]$ join the subscription plan, pay p_I, and the incumbent earns from these users $(1-p_I)p_I$. Hence, the incumbent's maximization problem is to choose p_I that maximizes:

$$\max_{p_I} \pi_I(p_I|(H,S)) = \alpha p_I + (1-p_I)p_I, \tag{6}$$

$$\text{s.t.} \quad \beta + v - p_I \geq v \text{ and } p_I \leq 1.$$

The first constraint requires that the user who is indifferent between joining the incumbent's data-based plan and the subscription-based plan prefers these options over joining the entrant's subscription plan for free. The second constraint requires that there is an internal solution to the indifferent user. The unconstraint solution is $p_I = (1+\alpha)/2$. Notice that users who join the subscription plan gain the utility $\beta + v - p_I = \beta + v - (1+\alpha)/2 < v$, where the inequality follows because $\beta < \frac{1}{2}$. Hence, the maximization problem has a corner solution in which the binding constraint is: $\beta + v - p_I > v$, or $p_I = \beta$. The incumbent earns from the hybrid model:

$$\pi_I(H,S) = (1+\alpha-\beta)\beta,$$

and the entrant earns $\pi_E(H,S) = 0$.

Lemma 2. *If the incumbent adopts the hybrid business-model, and the entrant can choose between $B_E = \{D, S, H\}$, then the incumbent dominates the market, charges $p_I = \beta$ and earns $\pi_I(H,S) = (1+\alpha-\beta)\beta$. The incumbent serves all users and commercializes the data of users with $k < \beta$.*

Notice that, as expected, an incumbent that adopts the hybrid model charges a higher

price under monopoly than under competition.[6]

Figure 2 illustrates the incumbent's benefit and cost of adopting the hybrid model. Intuitively, the benefit is that the incumbent dominates the entire market, allowing all users to fully benefit from network effect. Thus, users who adopt the subscription plan gain network effects not only from other subscribers but also from users on the data plan. This enables the incumbent to collect the full network effects through the price: $p_I = \beta$. The cost of the hybrid model is that the two platforms compete on the entire market. This results in fierce competition which reduces the incumbent's profit.

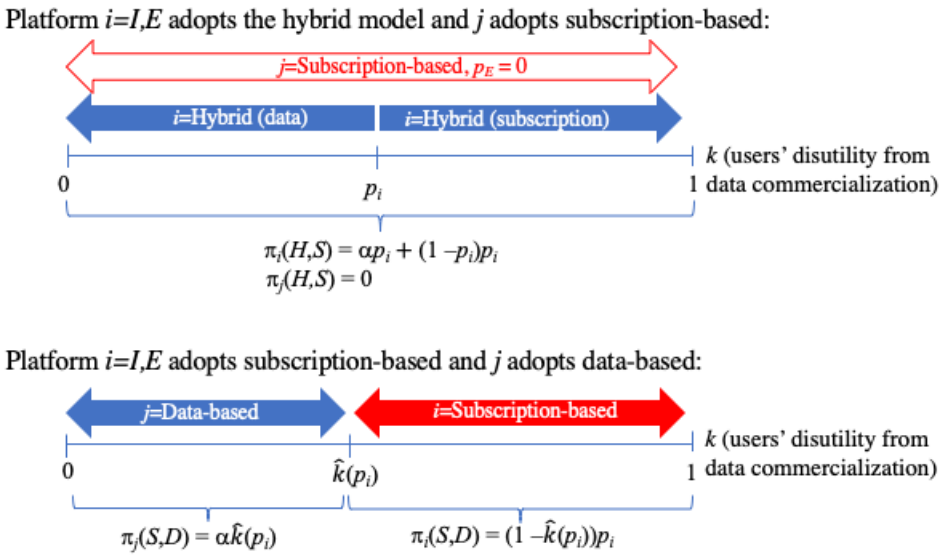

Figure 2: Market share and profits in the various business models configurations under competition

The incumbent adopts the data-based model

Suppose that the incumbent chooses the data-based business model. We solve for the market outcome given each of the entrant's potential business model choices.

Suppose first that the entrant adopts the subscription-based model. Then, there is no equilibrium in which the incumbent dominates the market, which makes focality irrelevant in this business model configuration. If such an equilibrium were to exist, $p_E = 0$ and all users join the incumbent. Yet, even when all users join the incumbent, the utility of the most data-sensitive user with $k = 1$ from joining the incumbent is $1 \times \beta + v - 1$, which is lower than the utility v that the user can gain by joining the entrant, because of our assumption that $\beta < 1 - v$. We, therefore, solve for an equilibrium in which the entrant gains a positive

[6] To see why, we have that $\frac{1+\alpha}{2} > \beta$ whenever $0 < \alpha < 2(\beta+v)-1$ and $\beta+v > \beta$ whenever $\alpha > 2(\beta+v)-1$.

market share. Intuitively, adopting different business models creates differentiation, which enables the entrant to gain positive market share despite the incumbent's focality advantage.

In equilibrium, given that n_E users join the entrant and $n_I = 1 - n_E$ users join the incumbent, there is a user, $\widehat{k}$, who is indifferent between joining the incumbent or the entrant. This user solves:

$$\beta(1 - n_E) + v - \widehat{k} = \beta n_E + v - p_E.$$

When there is an internal solution to $\widehat{k}$ (i.e., $0 < \widehat{k} < 1$), users with $k \in [0, \widehat{k}]$ join the incumbent because they are not sensitive to their privacy and therefore prefer a free service, even if the platform commercializes their data. In contrast, data-sensitive users with $k \in [\widehat{k}, 1]$ prefer the platform that charges a membership fee in order to protect their privacy. Hence, the demand function facing the entrant that solves $n_E = 1 - \widehat{k}$ is:

$$n_E(p_E) = \frac{1 - \beta - p_E}{1 - 2\beta}. \tag{7}$$

Because $\beta < 1/2$, the denominator in (7) is positive. Yet, notice that β has two conflicting effects on the demand facing the entrant. To see how, the inverse demand function of (7) is $p_E(n_E) = 1 - \beta - (1 - 2\beta)n_E$, which rotates counterclockwise around $n_E = 1/2$ as β increases, such that the demand increases with β if $n_E > 1/2$ and decreases with β otherwise. The intuition for this feature of the demand function is that when $n_E > 1/2$, the entrant, who does not commercialize users' data, serves more users than the incumbent and thus also collects more data. Hence, as network effects become stronger, the entrant's demand increases. The opposite case occurs when $n_E < 1/2$.

The entrant sets p_E to maximize $\pi_E(p_E) = p_E n_E(p_E)$:

$$p_E(D, S) = \begin{cases} \frac{1-\beta}{2}, & \text{if } \beta \leq \frac{1}{3}, \\ \beta, & \text{if } \beta > \frac{1}{3}, \end{cases} \qquad n_E(D, S) = \begin{cases} \frac{1-\beta}{2(1-2\beta)}, & \text{if } \beta \leq \frac{1}{3}, \\ 1, & \text{if } \beta > \frac{1}{3}, \end{cases} \tag{8}$$

where we note that $n_E(D, S) = \frac{1-\beta}{2(1-2\beta)} \leq 1$ if and only if $\beta \leq 1/3$. As a technical note, recall that the constraint $\beta < 1 - v$ implies that the second row in $p_E(D, S)$ and $n_E(D, S)$ are relevant only when $v < \frac{2}{3}$.

The entrant's price decreases in β while the entrant's market share increases in it.[7] Intuitively, at $\beta = 0$, the two platforms equally share the market. As β increases, the entrant's price decreases while its market share increases, because the entrant can better exploit the increase in network effects for enhancing its demand. Moreover, because the entrant does

[7] We verified that the utility of the indifferent user is always positive because $v > 1/2$, hence all users gain positive utility from joining a platform.

not commercialize users' data, the entrant can fully dominate the market if β is sufficiently high. The profits of the two platforms in the $(B_I, B_E) = (D, S)$ business model configuration are $\pi_E(D, S) = p_E n_E(p_E)$ and $\pi_I(D, S) = \alpha(1 - n_E(p_E))$, or:

$$\pi_E(D,S) = \begin{cases} \frac{(1-\beta)^2}{4(1-2\beta)}, & \text{if } \beta \leq \frac{1}{3}, \\ \beta, & \text{if } \beta > \frac{1}{3}, \end{cases} \qquad \pi_I(D,S) = \begin{cases} \frac{\alpha(1-3\beta)}{2(1-2\beta)}, & \text{if } \beta \leq \frac{1}{3}, \\ 0, & \text{if } \beta > \frac{1}{3}. \end{cases} \tag{9}$$

The following lemma summarizes the features of the (D, S) market configuration.

Lemma 3. *Suppose that the incumbent adopts a data-based business model and the entrant adopts a subscription-based one. Then, the entrant's price decreases with network effects, yet its market share increase with it. Moreover, if network effects are sufficiently strong, the entrant dominates the market.*

Suppose now that the entrant responds by adopting the data-based model. In the case where both choose the data-based model $(B_I, B_E) = (D, D)$, there are two equilibria: all users join the incumbent and all join the entrant. To solve the problem of multiple equilibria, we follow the literature on platform competition (Caillaud and Jullien (2001; 2003), Hałaburda and Yehezkel, 2016) and assume that the incumbent is "focal". Specifically, when there are two equilibria, one in which the incumbent dominates the market and the second in which the entrant dominates the market, users expect all other users to join the incumbent. Notice that focality permits an equilibrium with two active platforms or an equilibrium in which the entrant dominates the market, whenever there is no equilibrium in which the incumbent dominates the market. Focality grants the incumbent a competitive advantage that enables the incumbent to dominate the market whenever possible. This competitive advantage becomes stronger as network effects, β, increase because the benefit of users' expectations that other users will join the incumbent grows significantly.

The case of (D, D) is qualitatively similar to the case of (S, S): because the incumbent is focal, the incumbent dominates the market and the entrant earns 0. For brevity, we analyze this case in the online appendix and state here the following result:

Lemma 4. *Suppose that both platforms adopt the data-based business model. Then, the incumbent dominates the market, serves $n_I = \frac{v}{1-\beta}$ users and earns $\pi_I(D, D) = \frac{\alpha v}{1-\beta}$ while the entrant earns $\pi_E(D, D) = 0$.*

Finally, if the entrant chooses $B_E = H$, the entrant charges a price p_E from users that join the subscription plan, while offering a data plan for free. The analysis of this case is qualitatively similar to the analysis in the previous subsection. For brevity, we relegate it to the proof of lemma 5:

Lemma 5. *Suppose that the incumbent adopts the data-based model ($B_I = D$), and the entrant can choose between $B_E = \{D, S, H\}$. Then, there is a threshold,*

$$\overline{\alpha} = \begin{cases} \frac{1-\beta}{\sqrt{1-2\beta}} - 1, & \text{if } \beta < \frac{1}{3}, \\ 2\sqrt{\beta} - 1, & \text{if } \beta > \frac{1}{3}, \end{cases}$$

such that:

(*i*) *For $0 < \alpha < \overline{\alpha}$, the entrant adopts the subscription-based model, $B_E = S$.*

(*ii*) *For $\overline{\alpha} < \alpha$, the entrant adopts the hybrid model, $B_E = H$, and dominates the market.*

Intuitively, the entrant never responds to the incumbent's data-based model by adopting the same business model because then it loses the market. If data has low commercial value (α is small), the entrant prefers the subscription-based model which does not rely on commercializing user data. In contrast, when α is high, the entrant can leverage the hybrid model: attracting users with high disutility from data commercialization that are not served by the incumbent (given that under the data-based business model, the market is not fully covered) with the subscription plan. Then, using the network effect these users generate to attract the less data-sensitive users with a free plan. This strategy enables the entrant to dominate the market. In other words, because the incumbent does not cover the market with the data-based business model, this business model is vulnerable to entry and market domination by the entrant.

Since (D, S) and (S, D) are symmetric, Figure 2 illustrates also the incumbent's benefit and cost of adopting the data-based model. From the incumbent's perspective, if the entrant responds by adopting the subscription-based model, the incumbent can share the market with the entrant and compete with the entrant only on the marginal user (the user who is indifferent between the entrant's subscription plan and the incumbent's data plan); thereby avoiding fierce competition on the entire market. Yet, adopting the data-based model has an important weakness from the incumbent's viewpoint. The incumbent's potential revenues from the data-based model are high only when data has a high commercial benefit. In this case, however, the entrant adopts the hybrid model and monopolizes the market. In other words, choosing the data-based model is either not very profitable (when the commercial benefit is small), or exposes the incumbent to the threat of losing the entire market when the commercial benefit is high).

The incumbent adopts the subscription-based model

Suppose now that the incumbent adopts the subscription-based model. As in the case of (D, D), if the entrant also adopts the subscription-based model, the incumbent wins the market due to its focal position and $\pi_I(S, S) = p_I \times 1 = \beta$ and $\pi_E(S, S) = 0$. This logic follows to the case where the entrant adopts the hybrid model. Even if the entrant charges $p_E = 0$ and the incumbent charges $p_I = \beta$, there is an equilibrium in which all users join the incumbent and do not share data, because $\beta \times 1 + v - p_I \geq \beta \times 0 + v - p_E$. Finally, the (S, D) business-model configuration is symmetric to the (D, S) configuration discussed above: $\pi_E(S, D) = \pi_I(D, S)$ and $\pi_I(S, D) = \pi_E(D, S)$. The following lemma summarizes:

Lemma 6. *Suppose that the incumbent adopts the subscription-based model ($B_I = S$), and the entrant can choose between $B_E = \{D, S, H\}$. Then, the entrant adopts the data-based model and the platforms' profits are symmetric to the profits in Equation (9) ($\pi_E(S, D) = \pi_I(D, S)$ and $\pi_I(S, D) = \pi_E(D, S)$).*

Unlike the case where the incumbent adopts the data-based model–potentially enabling the entrant to adopt the hybrid model and dominate the market–the entrant cannot dominate the market with the hybrid model, when the incumbent adopts the subscription-based model. This is because the market is partially covered under the data-based model, which makes this business model more vulnerable to entry. The entrant can adopt the hybrid model, attract data-sensitive users, and exploit their network effects to capture the entire market. In contrast, under the subscription-based model, the incumbent can fully cover the market if the entrant adopts the hybrid model. In this case, the incumbent could simply lower its subscription price to block the entrant from attracting any users. As we show in section 7, when $\beta > 1$ and $\alpha > 1$, the incumbent fully covers the market under the data model, in which case the entrant cannot win the market with the hybrid model.

Similar to the intuition discussed above, adopting the subscription-based model enables the incumbent to soften competition with the entrant, as the two platforms only compete on the marginal user. The subscription model is also beneficial because its profitability does not depend on the commercial benefit, and can therefore be profitable when the commercial benefit is small. Yet, compared with the hybrid model, when adopting the subscription-based model, the incumbent does not cover the entire market, and users that join the incumbent do not benefit from the network effects of users that join the entrant. In this case, the incumbent cannot collect these network effects, which negatively affects its profits.

Equilibrium business model

We can now turn to solving the equilibrium business models when both platforms can adopt $B_i \in \{D, S, H\}$. The following proposition identifies the optimal business model for the incumbent to adopt:

Proposition 2. ***(Optimal business model)*** *Suppose that both platforms can adopt $B_i \in \{D, S, H\}$. Then, the incumbent adopts the hybrid model when the commercial benefit of data is high, and the subscription-based model otherwise. That is, there is a threshold, $\alpha^C_{H,S}$, where*

$$\alpha^C_{H,S} = \begin{cases} \frac{(1-\beta)(1-5\beta+8\beta^2)}{4\beta(1-2\beta)}, & \text{if } \beta \leq \frac{1}{3}, \\ \beta, & \text{if } \beta > \frac{1}{3}, \end{cases} \tag{10}$$

such that when $\alpha > \alpha^C_{H,S}$, the incumbent adopts the hybrid model and dominates the market. When $\alpha < \alpha^C_{H,S}$, the incumbent adopts the subscription-model while the entrant adopts the data-based model and the two platforms share the market. Moreover, $\alpha^C_{H,S}$ is decreasing in β for $\beta \leq 1/3$ and increasing with β otherwise.

Intuitively, recall from Figure 2, the hybrid model allows the incumbent to monopolize the market. Yet, this comes at the cost of intense competition, because the two platforms compete on the entire market. By adopting the subscription-based model, the incumbent shares the market with the entrant, which in turn results in less intense competition as the two platforms compete only on the marginal user. Additionally, by offering a subscription plan, the hybrid model enables the incumbent to collect the commercial benefit while simultaneously being attractive to users with high disutility from being commercialized.

We therefore have that the hybrid model is more profitable for the incumbent when the commercial benefit is high, as this increases the revenues from the data plan. Moreover, Proposition 2 shows that network effects play a crucial role in the incumbent's decision: $\alpha^C_{H,S}$ is decreasing with β when $\beta < 1/3$, and increases with it otherwise. This implies that given a level of commercial benefit, the incumbent adopts the hybrid model when network effects are strong and the subscription-based model when network effects are weak. The intuition is that when $\beta < 1/3$, stronger network effects reduce the incumbent's concern about competition due to its focality. Moreover, the benefit of having all users on the same platform grows with network effects, incentivizing the incumbent to adopt the hybrid model and dominate the market. Conversely, when network effects are weak, the incumbent prefers to avoid competition and goes for "live and let live" by adopting the subscription-based model and allowing the entrant to enjoy a positive market share. When $\beta > 1/3$, the incumbent can dominate the market even with the subscription-based model, and earn β. Hence, the

incumbent adopts the subscription-based model if $\alpha < \beta$ and adopts the hybrid model otherwise.

Notice that the two platforms never adopt the same business model, because the non-focal entrant cannot gain market share by simply replicating the incumbent's model. If platforms were horizontally differentiated, they could co-exist even when adopting the same business model. Intuitively, in such a case users would join the platform with their preferred characteristics, resulting in two active platforms in the market. By focusing on homogeneous platforms, we can highlight the strategic role business models play in creating differentiation, and thereby relaxing competition and enabling the two platforms to co-exist.

5 The effect of competition

In order to study how competition affects the platforms' choice of business model, in this section we compare the monopolistic case to the competitive case discussed above .

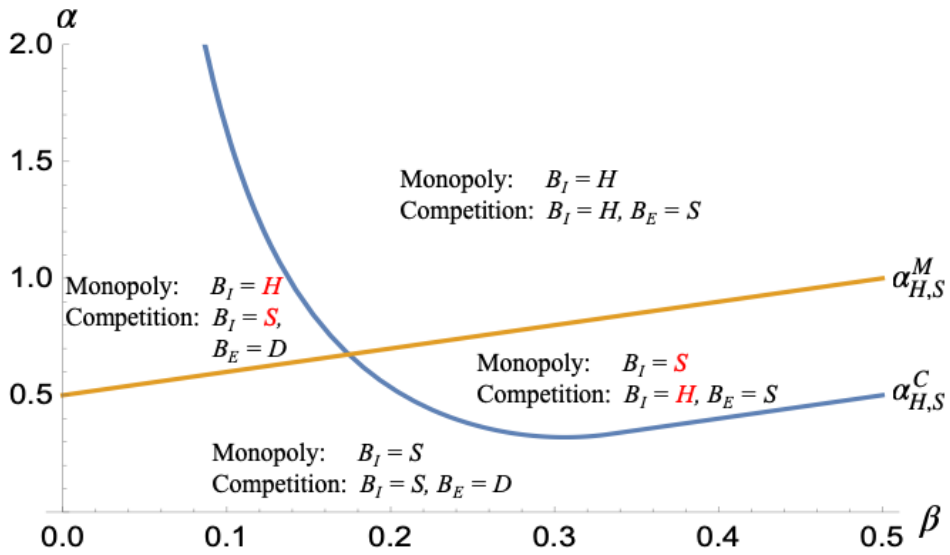

Figure 3: $\alpha^M_{H,S}$ and $\alpha^C_{H,S}$ as a function of β and the equilibrium business models (for $v = \frac{1}{2}$)

Figure 3 illustrates the two thresholds, $\alpha^M_{H,S}$ and $\alpha^C_{H,S}$ as a function of β, where recall that $\alpha^C_{H,S}$ is equal to (10) when $\beta < \frac{1}{3}$ and equals to β when $\beta > \frac{1}{3}$. The figure shows that for low and high values of α, competition does not change the incumbent's behavior. Specifically, if $\alpha < \min\left\{\alpha^M_{H,S}, \alpha^C_{H,S}\right\}$, the incumbent keeps adopting the subscription-based model. Intuitively, for low commercial value of data, it is optimal to avoid data commercialization and instead charge users for the value generated by the platform. For the opposite reason, the incumbent adopts the hybrid model under both monopoly and competition when α is very high, such that $\alpha > \max\left\{\alpha^M_{H,S}, \alpha^C_{H,S}\right\}$.

Yet, competition affects the incumbent's business model for intermediate values of α, when β is either high or low. For low values of β and intermediate values of α, such that $\alpha^M_{H,S} < \alpha < \alpha^C_{H,S}$, competition prompts the incumbent to switch from the hybrid model to the subscription-based model. Here, competition makes the hybrid model less profitable for the incumbent, because it now has to compete with the entrant on all users. With small β, the incumbent lacks a strong focal position. This combined with the aggressive competition with the entrant on the entire market prompts the incumbent to switch from the hybrid model to the less competitive subscription-based model. As a result, the entrant gains a positive market share, and the two platforms compete only for the marginal user. That is, the weaker the network effects the stronger the incentive to avoid competition; driving the incumbent to shift from the hybrid to the subscription-based model.

Conversely, for intermediate β values and intermediate α values, such that $\alpha^C_{H,S} < \alpha < \alpha^M_{H,S}$, competition encourages the incumbent to switch from the subscription-based model to the hybrid model. Here, a monopolistic incumbent can collect all users' network effects and thus leverages the high β to do so. However, under competition, this strategy enables the entrant to adopt the data-based model and steal the data-insensitive users from the incumbent. Anticipating this, the incumbent adopts the hybrid model and monopolizes the market. The strong network effects enhance the incumbent's focality advantage, thereby mitigating the competitive impact of the hybrid model.

Netflix's introduction of an ad-supported plan alongside its subscription model nicely illustrates the shift to a hybrid business approach in response to increasing competition. Netflix's value proposition heavily depends on leveraging user data, such as viewing habits, to enhance recommendations and even inform content development. In a more competitive landscape with strong network effects, Netflix's transition to a hybrid model is consistent with our predictions.

The following Corollary summarizes these results:

Corollary 1. *Consider the change of a monopolistic incumbent's choice of a business model when faced with the threat of competition:*

(*i*) *when the commercial benefit is high,* $\alpha > \max\left\{\alpha^M_{H,S}, \alpha^C_{H,S}\right\}$ *(low,* $\alpha < \min\left\{\alpha^M_{H,S}, \alpha^C_{H,S}\right\}$*), the incumbent adopts the hybrid (subscription-based) model under both monopoly and competition;*

(ii) *when the commercial benefit is intermediate and network effects are weak, such that* $\alpha^M_{H,S} < \alpha < \alpha^C_{H,S}$*, competition motivates the incumbent to shift from the hybrid model to the subscription-based model;*

(iii) *when the commercial benefit is intermediate and network effects are strong, such that* $\alpha^C_{H,S} < \alpha < \alpha^M_{H,S}$*, competition motivates the incumbent to shift from the subscription-based model to the hybrid model.*

Does competition suppress data collection?

In June 2024, a background note by the OECD noted that "...it could be argued that insufficient competition would hinder individual data privacy rights or principles..." The view that competition may provide platforms strong incentives to reduce data commercialization is shared by many. Our model can help shed light on this question. As we show below, we find that competition suppresses data commercialization only when the commercial benefit of data is high.

As Figure **??** shows, when $\alpha < \alpha^M_{H,S}$, competition introduces a data-based model that was not offered under monopoly, as a monopolist prefers the subscription model under these conditions. Thus, competition in this parameter space increases data commercialization, which was non-existent under monopoly. Only when $\alpha > \alpha^M_{H,S}$, competition suppresses data commercialization. In particular, when network effects are strong, the monopolist remains with the hybrid model that it also offers under monopoly. Still, because the entrant enters with a free subscription plan, the incumbent's subscription fee under competition is lower than under a monopoly.[8] This in turn, increases the number of users that prefer the subscription plan over the data-plan and thereby reduces the overall amount of data commercialized. A similar argument holds when network effects are weak and α is high. In this case, under competition, the incumbent switches from choosing the hybrid model as a monopolist to the subscription model and the entrant enters with the data plan. That is, in general, users have the same choice in terms of plans offered. However, under competition the two plans are offered by two competing platforms, rather than by one platform that allows users to choose their plan. Moreover, competition drives the subscription fee down so prices of the subscription plan are lower than in the case of a monopolist offering the hybrid model. Thus, again, more users choose the subscription plan over the data plan and less data is commercialized. To see why, notice that under monopoly, the incumbent collects data from $p_I = v + \beta$ users (it is possible to show that $\alpha^M_{H,S} > 2(\beta + v) - 1$). Under competition, we have from (8) that the entrant collects data from $n_E(S, D) = 1 - \frac{1-\beta}{2(1-2\beta)} = \frac{1-3\beta}{2(1-2\beta)}$ users. Yet, $v + \beta > \frac{1-3\beta}{2(1-2\beta)}$, implying that the incumbent collects more data under monopoly than the amount of data that the entrant collects under competition. The analysis above suggests that subscription prices are potentially a useful lever to encourage platforms to adopt more

[8] See Footnote 6.

privacy-focused business models. The following Corollary summarizes:

Corollary 2. *Competition suppresses data commercialization if and only if* $\alpha > \alpha^M_{H,S}$.

6 Should the hybrid model be banned?

The competitive effect of the hybrid model has been recently a topic of a strong debate in the European Union (EU). Specifically, in response to regulatory changes in the EU, in November 2023 Meta introduced a paid option for its EU users of Facebook and Instagram where users can choose between (i) paying a monthly fee for an ad-free version of these social networks; or (ii) enjoy a free-of-charge access to a version of these social networks with personalized ads. On July 1, 2024, the European Commission informed Meta that its hybrid business model of "pay or consent" fails to comply with the Digital Markets Act (DMA). Below, we analyze the competitive effect of the availability of the hybrid model. We focus on the case where $\beta < \frac{1}{3}$ (or, when $v > \frac{2}{3}$, consider the case where $\beta < 1 - v$), such that the incumbent indeed adopts the hybrid model and consider two scenarios: one where neither platform can adopt the hybrid model, and another where only the incumbent, being the "gatekeeper", is banned from adopting it (which aligns with the DMA's approach). The analysis below provides the main results, while the full analysis appears in Online Appendix B. In the first case, there are 4 market configurations: $(B_I, B_E) = \{(D,S),(D,D),(S,S),(S,D)\}$, all of which have been discussed above. We can, therefore, directly discuss the equilibrium business models.

The impact of a ban on the hybrid model on platforms' business models

Suppose first that both platforms cannot adopt the hybrid model. Intuitively, the incumbent adopts the data-based model if the commercial benefit is high and the subscription mode otherwise. Because the entrant loses the market if it chooses the same business model as the incumbent, it responds by choosing the opposite business model. If, however, only the incumbent cannot adopt the hybrid model, the incumbent no longer wants to adopt the data-based model because now the entrant would respond with the hybrid model and dominate the market. Hence, in this case, the incumbent adopts the subscription-based model for all values of α and the entrant adopts the data-based model. We, thus, have:

Proposition 3. ***(A ban on the hybrid model)***

(*i*) *Suppose the hybrid model is banned on both platforms. Then, there is a cutoff* $\alpha^C_{D,S} = \frac{(1-\beta)^2}{2(1-3\beta)}$ *such that the incumbent chooses the data-based business model iff* $\alpha > \alpha^C_{D,S}$*, and chooses the subscription-based model otherwise.*

(*ii*) *Suppose that the hybrid model is banned only on the incumbent. Then, the incumbent always adopts the subscription-based model and the entrant adopts the data-based model.*

Notice that when only the entrant can adopt the hybrid model, the entrant never chooses the hybrid model in equilibrium; its only role is to deter the incumbent from adopting the data-based model.

Does the hybrid model suppress data commercialization?

The hybrid model is controversial from a policy viewpoint because it arguably discriminates between users who are willing to share their data and users who refuse to share their data. Below, we comment on the implications of the hybrid model for social welfare. The main conclusion of the analysis is that when network effects are strong, allowing platforms to adopt the hybrid model can be in fact welfare enhancing.

We start by asking whether banning the hybrid model under platform competition leads to a reduction in data commercialization. If $\alpha < \alpha^C_{H,S}$, the availability of the hybrid model has no effect on the amount of data commercialized either because it does not affect the incumbent's behavior ($\alpha < \min\left\{\alpha^C_{H,S}, \alpha^C_{D,S}\right\}$), or because the platforms swap their business models (the incumbent shifts from S to D, and the entrant from D to S). In the latter case , due to symmetry, prices remain unchanged, and thus so does the amount of data commercialized.

For intermediate and high levels of α and β, $\alpha > \alpha^C_{H,S}$, banning the hybrid business model shifts the market structure from one where the incumbent dominates the market with a hybrid model to a structure where one platform adopts the data-based model and the other opts for the subscription-based model.[9]

In the former case, the incumbent charges $p_I = \beta$ and hence β users join the incumbent's data plan. In the latter case, we have from equation (8) that the platform that adopts the data-based model collects data from $n_i(S, D) = 1 - \frac{1-\beta}{2(1-2\beta)} = \frac{1-3\beta}{2(1-2\beta)}$ users. Hence, we have that, when $\beta > \frac{1-3\beta}{2(1-2\beta)}$, or $\beta > 1/4$, the hybrid model results in more data commercialized, relative to the (S, D) market configuration. The following corollary summarizes:

Corollary 3. *Under platform competition, banning the hybrid model has no effect on data commercialization when $\alpha < \alpha^C_{H,S}$. Otherwise, banning the hybrid model decreases data commercialization if $\beta > 1/4$ and increases it otherwise.*

The intuition for the second part of Corollary 3 is that if network effects are strong, the incumbent adopts the hybrid model and charges a high price for the subscription plan, as

[9]This result follows to the case where only the incumbent is banned from adopting the hybrid model.

all users join its platform. This drives more users to adopt the data plan, resulting in more data commercialized than under the (S, D) market configuration.

Is the hybrid model welfare enhancing?

Finally, we ask whether social welfare is higher when platforms can adopt the hybrid model, in comparison with the case in which competition authorities forbid platforms from discriminating users based on whether they are willing to have their data commercialized. Note that we focus our comparison on the case where $\beta < 1/3$, such that when the platforms adopt different business models, both platforms are active in the market.

The following proposition compares social welfare with and without the hybrid model:

Proposition 4. ***(Is the hybrid welfare enhancing?)*** *Suppose that $\alpha > \alpha^C_{H,S}$. Then, if $\beta < 1/4$, there is a threshold,*

$$\alpha^C_W = \frac{1 - \beta(1 - 4\beta(4\beta - 3))}{4(1 - 6\beta + 8\beta^2)},$$

such that the hybrid model is welfare enhancing if $\alpha > \alpha^C_W$. Moreover, α^C_W approaches infinity as β approaches $1/4$. If $\beta > 1/4$, the hybrid model is welfare enhancing.

Notice that the comparison is unaffected by v because the market is covered in both cases. Given that the comparison is affected only by α and β, Figure 4 illustrates the regions in which the hybrid model is welfare enhancing or reducing, given α and β.

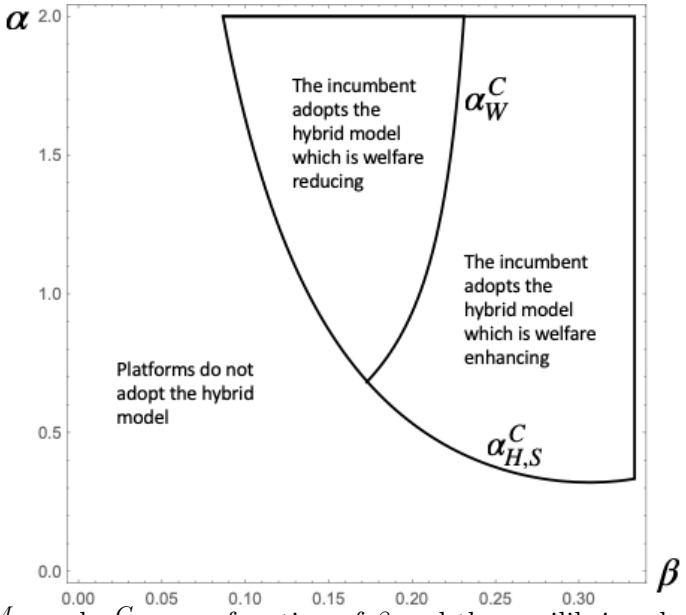

Figure 4: $\alpha^M_{H,S}$ and $\alpha^C_{H,S}$ as a function of β and the equilibrium business models

The figure shows that when the hybrid model is relevant (i.e., when $\alpha > \alpha^C_{H,S}$), its welfare implications depend on the strength of network effects. Specifically, there is a threshold, α^C_W, such that the hybrid model is welfare enhancing when β is high, and welfare reducing otherwise. The threshold, α^C_W, approaches infinity as β approaches 1/4. For $\beta > 1/4$, the hybrid model is welfare enhancing for all values of α, while if $\beta < 1/4$, there is a region in which the hybrid model is welfare reducing which expands with α. Given that users can opt out of data commercialization under both market configurations, the welfare comparison depends on two effects: users' network effects and the commercialization benefit to the platform. Under the hybrid model, all users join the same platform, thereby enhancing their network effect. From the platform's perspective, recall from Corollary 3 that the hybrid model increases total commercial benefit if $\beta > 1/4$, and decreases it otherwise. Thus, if $\beta > 1/4$, both effects are positive and the hybrid model is always welfare enhancing. If $\beta < 1/4$, the second effect is negative and outweighs the first effect if data has a high commercial benefit.

Next, consider consumer surplus. Recall that when the incumbent adopts the hybrid model, it charges $p_I = \beta$. All users join the incumbent and benefit from each-other's network effects. When the hybrid model is banned, one of the platforms adopts the subscription-based model and charges $\frac{1-\beta}{2} > \beta$, while the other adopts the data-based model. That is, users joining the subscription plan pay a higher price than when the incumbent adopts the hybrid model, and because users are split across two platforms, network effects are weaker. Consequently, a ban on the hybrid model reduces consumer surplus of all users.

We summarize with two notes: (1) All the results in this subsection follow to the case where only the incumbent is banned from adopting the hybrid model. In this case, the market configuration remains the same with one platform choosing the subscription-based model while the other chooses the data-based model; and (2) our consumer surplus analysis does not account for the long-term effect in which the hybrid model may deter entry by making market entry prohibitively costly for potential entrants. We further discuss this point below.

Hybrid model and efficient entry deterrence

The hybrid business model may have long term negative effects that we so far ignored. Specifically, to focus on platforms' choice of business models, we assume that, with the exception of focality, the two platforms are identical: same base quality, v, same network effects, β, and the same marginal costs (normalized to 0). Below, we comment on the role that the hybrid model plays in deterring efficient entry, i.e., blocking an entrant with a superior quality.

The results above show that when $\alpha > \alpha^C_{H,S}$, the hybrid model provides the incumbent, and not the entrant, with a competitive advantage. That is, allowing both platforms to adopt the hybrid model has an asymmetric effect: it increases the incumbent's competitive advantage, and not the entrant's. Under the hybrid model, the incumbent earns strictly positive profits when facing an equal-quality entrant. By continuity, this advantage extends—up to a certain point—to the case where the entrant offers a higher quality. In such a case, the hybrid model has another welfare reducing effect: it helps the incumbent deter efficient entry and dominate the market. Hence, our results suggest that restricting "gatekeepers" from adopting the hybrid model can enhance welfare-by allowing superior quality entrants to compete effectively. However, this welfare enhancing effect occurs only when the hybrid model provides the incumbent with a competitive advantage, i.e., when $\alpha > \alpha^C_{H,S}$, as identified in this section.

7 Extension: strong network effects

Our base model focuses on the case of weak network effects: $\beta < 1 - v$, such that the incumbent cannot cover the market with the data-based model. That is, network effects are not strong enough to convince the most data-sensitive user (with $k = 1$) to join the platform and agree to data commercialization. In this parameter space, we find that the incumbent never adopts the data-based model, because the hybrid model allows the incumbent to attract all users while avoiding data commercialization of highly data-sensitive users and instead charging them a positive price. This raises the question: what are the platforms' optimal business models when network effects are strong?

Online Appendix C provides the full characterization of the equilibrium business models under monopoly and competition for all values of β, including the case where $\beta > 1 - v$. For brevity, in this section we describe the main results and the intuition behind them.

Starting with the monopolistic case, the following lemma characterizes the optimal business model:

Corollary 4. *Suppose that $\beta > 1 - v$. Then, the monopolistic incumbent never adopt the hybrid model. Moreover:*

If $\alpha < \beta + v$, the incumbent adopts the subscription-based model,

If $\alpha > \beta + v$, the incumbent adopts the data-based model.

The intuition for this result is that when the market is fully covered under both the subscription model and the data-based one, the platform cannot benefit from combining

both plans. Hence, the platform chooses the subscription model when the commercial value of data is low and the data-based model when it is high.

In contrast to the monopolistic case, under competition, the hybrid model can help the incumbent win the market. As expected, the value for the incumbent from the hybrid model depends on the strength of the network effects. The following corollaries characterize the case of intermediate network effects.

Corollary 5. *Suppose that $1 - v < \beta < 1$. Then, competition motivates the incumbent to adopt the hybrid model. Specifically: the incumbent switches from the data-based model under monopoly to the hybrid model under competition when $\alpha > v + \beta$, and from the subscription-based model under monopoly to the hybrid model under competition when $\beta < \alpha < \beta + v$.*

Next, consider the case of strong network effects. When $\beta > 1$, network effects are strong enough for the incumbent to cover the market with the data-based model under competition. In this case, the incumbent may choose to forgo the hybrid model and instead go with either the data-based or the subscription-based models instead.

Corollary 6. *When $\beta > 1$, the incumbent never adopts the hybrid model. Moreover:*

1. *if $\alpha < \beta$, competition does not affect the incumbent's behavior: the incumbent chooses the subscription-based model under both monopoly and competition,*
2. *if $\beta < \alpha < \beta + v$, competition motivates the incumbent to switch from the subscription-based model under monopoly to the data-based model under competition,*
3. *if $\beta + v < \alpha$, competition does not affect the incumbent's behavior: the incumbent adopts the data-based model under both monopoly and competition.*

Recall that under monopoly, the subscription-based model allows the incumbent to extract the full value from users, $\beta + v$, while under competition, the maximum the incumbent can extract is β. In contrast, the maximum the incumbent can get with the data-based model under both competition and monopoly is α. Consequently, with strong network effects and full market coverage, the incumbent prefers the data-based model if the commercial value of data is higher than the network effects it can collect under the subscription model. Conversely, when the commercial value of data is lower, the subscription model is preferred. This holds regardless of the presence of competition. Competition becomes relevant for intermediate commercial benefit. In this case, competition motivates the incumbent to switch from the subscription-based model that requires price competition with the entrant to the data-based model, where there is no price competition and the incumbent can enjoy the commercial value, α.

Corollaries 5 and 6 have important policy implications. Regulations like the GDPR aim to limit platforms' ability to collect user data, weakening the network effects that many dominant platforms rely on. As network effects weaken, our model predicts that, under competition, platforms should switch from the data-based business model to the hybrid one. This, for example, may shed light on Meta's decision to shift to the hybrid model in the EU.

8 Managerial Implications

In today's information age, where data plays an increasingly important role in platforms' value creation, platforms are faced with the value capture dilemma of whether to base their business model on the "traditional" practice of charging users for their services, adopt the newer model of monetizing user data, or do both. Our analysis provides guidelines with respect to when it is optimal for platforms to adopt each business model, and thus has important managerial implications both for competing and monopolistic platforms.

What determines platforms' choice of business model? Our analysis offers direct insights into the profitability of the different business models for platforms. Most importantly, we find that when choosing their business model, platforms should consider not only the commercial value of data but also the strength of network effects. While it might seem intuitive that commercializing data would be the profitable business model if the commercial benefit of data is high, our model reveals that for intermediate commercial value, the strength of network effects is crucial for determining the optimal business model. This is particularly important for many of today's most popular platforms, where network effects are often driven by the benefits that data collected on users provides to other users. For example, in the case of a navigation app such as Waze which collects information on drivers' location, the data collected is crucial to other users that use the app and, in fact, is the core of the service that the app provides. The same principle applies to platforms like Netflix and Spotify, where the value of the service is heavily dependent on user data rather than direct user interactions. Our results indicate that it is imperative for platforms to assess the strength of network effects to determine their optimal business model.

Our findings suggest that as long as the commercial value is high enough, platforms should choose the data-based model when network effects are strong, and should go with the hybrid model when network effects are moderate, because it allows them to attract all users in the market, whether in a competitive environment or under monopoly. Conversely, when network effects are weak, or the commercial value is low, the subscription-based model should be chosen. Indeed, many of today's platforms offer the hybrid model, including

platforms in the music streaming market (e.g., Pandora and Spotify), and video streaming market (Netflix, Hulu, Disney+), while social media platforms with high commercial value and strong network effects (e.g., Facebook and TikTok) offer only a data-based plan. On the other hand, platforms with weak network effects tend to choose the subscription-based model. For example, the network effects in apps like Ride with GPS, a social app that provides route directions to cyclists, are relatively weak. Although the app collects data on rider's location, the data is mainly used to provide directions to the individual rider rather than to offer real-time information on the location of other users, as in the case of Waze.

How does the threat of competition affect platforms' choice of business model? The market for many of today's most dominant platforms is becoming increasingly competitive. TikTok is challenging Facebook's dominance, Netflix faces strong competition from Hulu, Disney+, and others, and even Google is concerned about Microsoft's integration of ChatGPT with Bing. Our model provides valuable insight for incumbent platforms facing such competition. Specifically, according to our results, platforms under threat of entry should consider changing their business model only if the commercial value of data is moderate. If the value is high or low, a monopolistic platform facing competition should maintain its current model: data-based if the value is high, hybrid if it is moderate, and subscription-based if it is low.

When the commercial value of data is moderate and network effects are strong, a monopolist facing competition should shift away from the subscription-based model, that is optimal under monopoly, to the hybrid model. Given the strong network effects, the hybrid model would allow the incumbent to attract both privacy-sensitive and non-privacy-sensitive users, making it harder for the entrant to gain a foothold in the market. More generally, our results indicate that in competitive environments, the hybrid model can help incumbents deter entry or prevent entrants from dominating the market.

However, the incumbent should adopt the hybrid model only as long as it is profitable for the incumbent to deter entry. If network effect are not strong enough to make entry deterrence profitable, the platform should shift to a subscription-based model. Remaining with the hybrid model could lead to intense competition with the entrant over the entire market. In contrast, shifting to the subscription model allows an entrant platform to differentiate itself and offer users a data-based plan; thereby, competing with the incumbent only over the marginal users. Moreover, shifting to the subscription-based model is preferable to the data-based model because the latter would prompt the entrant to adopt the hybrid model and monopolize the market. That is, the subscription model softens competitions and prevents the entrant from dominating the market.

This dynamic is nicely demonstrated by the music- and video streaming markets. Specifically, when Pandora first introduced its music streaming service in the U.S., it launched with a hybrid model. This allowed Pandora to fend off competition for a while. When Spotify launched in the U.S. two years later, it entered with a low price in its subscription plan, intensifying competition. Conversely, Netflix's subscription-based business model led Hulu to differentiate itself by entering the video streaming market with a data-based plan, attracting users with lower privacy concerns. Still, competition in the video streaming market pushed competitors like Hulu, Amazon, and even Netflix to switch to the hybrid model. Interestingly, recently in November 2024, Amazon reversed course, shutting down its ad-supported service, FreeVee, and reverting back to a subscription only model. Relative to other streaming services, Amazon has weaker network effects and lower commercial value, positioning it as a good candidate to switch away from the hybrid model to a subscription-based one.[10]

Does competition suppress data commercialization? Our model also has important policy implications, offering clear guidelines on when competition might encourage platforms to adopt more privacy-focused business models and reduce data commercialization. Our findings indicate that whether competitions suppresses data commercialization mainly depends on the commercial value of data and the degree of network effects. When the commercial value of data is low, then under monopoly, platforms adopt the subscription model and no data is being commercialized. Competition introduces the data-based model or the hybrid model, leading to greater data commercialization. Moreover, the region in which competition leads to greater data commercialization increases with network effects.

Surprisingly, it is when the commercial value is high that competition decreases data commercialization. In this case, while competition does not change the plans offered in the market – both subscription and data plans are available – competition drives subscription prices down. This, in turn, makes the subscription plan attractive to more users and thus results in less data being commercialized. This suggests that regulators can encourage more privacy by regulating prices rather than business models.

Should discrimination based on data-sharing be banned? Our analysis of a ban on the hybrid model offers valuable insights for both managers and policymakers. From a managerial perspective, if the hybrid model is unavailable–whether due to regulation, implementation complexity, or simply lack of popularity–the optimal business model depends on the commercial value of data and the strength of network effects. Specifically, if the com-

[10]Note that Amazon has always offered free content to non-Prime members. However, this content has never been ad-supported.

mercial value of data is high and network effects are not too strong, the incumbent should adopt the data-based model; otherwise, the subscription model is preferable. For example, when Netflix first launched, the commercial value of its data was likely low because Netflix had few users, limited knowledge about them, and network effects were weak. Thus, it made sense for Netflix to launch with the subscription model rather than a data-based plan. In contrast, Google's high commercial value at launch justified its adoption of a data-based business model.

In terms of policy implications, our finding that the hybrid model enhances welfare when network effects are strong suggests that banning it—whether for all platforms or only for dominant ones— could negatively affect welfare. Therefore, decisions on whether to allow platforms to discriminate based on data-sharing should be made on a case-by-case basis, considering the strength of network effects in the market.

9 Conclusion

Data is becoming an essential asset for platforms and an important determinant of platforms' monetization strategies. We develop a tractable model to study how competition affects platforms' optimal business model in a market with network effects and when data has a commercial benefit to the platforms. Platforms can choose between three business models: data-based, subscription-based, and hybrid. We find that the effect of competition on platforms' optimal business model depends on the interaction between the strength of the network effects and the commercial benefit of data.

We establish three main result. First, competition can have an important effect on platforms' business strategy. Specifically, for intermediate commercial value, the threat of competition motivates an incumbent platform to switch from the subscription-based model to the hybrid model when network effects are moderate, and from the hybrid model to the subscription-based model if network effects are weak.

Second, when network effects are strong, competition may in fact increase data commercialization by motivating platforms to adopt the hybrid model rather than the subscription one. That is, competition does not necessarily promote a more privacy-sensitive market.

Our third key result relates to the hybrid business model. Allowing platforms to discriminate across users based on whether they share their data for commercialization–i.e., to offer the hybrid model– may lead to a more concentrated market where the incumbent can deter the entry of a new (and perhaps superior) platform. Still, the hybrid model can be welfare enhancing, if network effects are strong enough, but reduces welfare for intermediate values of network effects.

References

[1] Acemoglu, Daron, Ali Makhdoumi, Azarakhsh Malekian, and Asu Ozdaglar. 2022. "Too Much Data: Prices and Inefficiencies in Data Markets." *American Economic Journal: Microeconomics*, 14 (4): 218-56.

[2] Aridor, Guy, Yeon-Koo Che, and Tobias Salz. 2023. "The effect of privacy regulation on the data industry: Empirical evidence from GDPR." *The Rand Journal of Economics* 54(4): 695-730.

[3] Bergemann, Dirk, Alessandro Bonatti and Tan Gan. 2022. "The Economics of Social Data." *The Rand Journal of Economics* 53(2): 263-296.

[4] Biglaiser, Gary, and Jacques Crémer. 2020. "The value of incumbency when platforms face heterogeneous customers." *American Economic Journal: Micro* 12(4):1-43.

[5] Caillaud, Bernard, and Bruno Jullien. 2001. "Competing cybermediaries." *European Economic Review* 45 (4-6): 797–808.

[6] Caillaud, Bernard, and Bruno Jullien. 2003. "Chicken & egg: Competition among intermediation service providers." *The RAND Journal of Economics* 34 (2): 309–328.

[7] Calvano, Emilio, and Michele Polo. 2020. "Strategic differentiation by business models: Free-to-air and pay- TV." *The Economic Journal* 130(625): 50–64

[8] Casadesus-Masanell, Ramon, and Andres Hervas-Drane. 2015. "Competing with Privacy." *Management Science*, 61(1): 229–246.

[9] Casner, Ben and Teh, Tat-How. 2024. "Content-hosting platforms: discovery, membership, or both?". *The RAND Journal of Economics* (*forthcoming*).

[10] Chen, Zhijun. 2025, "Paying Consumers for Their Data: An Economic Analysis of Data Acquisition and Digital Privacy." Working paper.

[11] Choi, Jay Pil, Doh-Shin Jeon, and Byung-Cheol Kim. 2019. "Privacy and personal data collection with information externalities." *Journal of Public Economics* 173:113-124.

[12] Fairfield, Joshua AT, and Christoph Engel. 2015. "Privacy as a public good." *Duke Law Journal* 65: 385.

[13] Fainmesser, Itay, Andrea Galeotti, and Ruslan Momot. 2022. "Digital Privacy." Management Science 69(6): 3157-3173.

[14] Hagiu, Andrei. 2006. "Pricing and commitment by two-sided platforms." *The RAND Journal of Economics* 37 (3): 720–737.

[15] Hagiu Andrei and Julian Wright. 2023. "Data-enabled learning, network effects and competitive advantage." *RAND Journal of Economics* 54(4): 638-667.

[16] Hałaburda, Hanna, and Yaron Yehezkel. 2013. "Platform competition under asymmetric information." *American Economic Journal: Microeconomics* 5 (3): 22–68.

[17] Hałaburda, Hanna, and Yaron Yehezkel. 2016. "The role of coordination bias in platform competition." *Journal of Economics and Management Strategy* 25 (2): 274–312.

[18] Hałaburda, Hanna, and Yaron Yehezkel. 2019. "How beliefs affect platform competition." *Journal of Economics and Management Strategy* 28 (1), 49-49.

[19] Hałaburda, Hanna, Bruno Jullien, and Yaron Yehezkel. 2020. "Dynamic platform competition: how history matters?" *The RAND Journal of Economics* 51 (1): 3-31.

[20] Johnson Garrett A., Scott K. Shriver, and Samuel G. Goldberg. 2023. "Privacy and Market Concentration: Intended and Unintended Consequences of the GDPR." *Management Science* 69 (10): 5695–6415

[21] Janßen, Rebecca, Reinhold Kesler, Michael E. Kummer, and Joel Waldfogel. 2022. "GDPR and the Lost Generation of Innovative Apps." NBER Working Paper 30028.

[22] Jullien, Bruno. 2011. "Competition in multi-sided markets: Divide and conquer." *American Economic Journal: Microeconomics* 3 (4): 186–220.

[23] Jullien, Bruno, Alessandro Pavan, and Marc Rysman. 2021. "Two-sided markets, pricing, and network effects." In Handbook of Industrial Organization, vol. 4, no. 1, pp. 485-592. Elsevier.

[24] Katz, Michael L., and Carl Shapiro. 1985. "Network Externalities, Competition, and Compatibility." *American Economic Review*, 75: 424-140.

[25] Ke, Tony T., and K. Sudhir. 2023. "Privacy rights and data security: GDPR and personal data driven markets." *Management Science* 69 (8): 4363-4971

[26] Liang, Annie, and Erik Madsen. 2019. "Data Sharing and Incentives." Available at SSRN 3485776.

[27] Llanes, Gastón and Leonardo Madio, 2024. "Business strategy and regulation of generative AI firms." Working paper.

[28] Markovich, Sarit, and Yaron Yehezkel. 2022. "Group Hug: Platform Competition with User-groups." *American Economic Journal: Micro,* 14 (2): 139-175.

[29] Markovich, Sarit, and Yaron Yehezkel. 2024. "For the public benefit": data policy in platform markets." *Journal of Economics and Management Strategy,* 33: 652 – 685.

[30] Markovich, Sarit and Luis Rayo. 2024. "What's on the menu: platform focality & data commercialization." CEPR Working Paper DP19792.

[31] O'Brien, Daniel and Doug Smith. 2014. "Privacy in online markets: A welfare analysis of demand rotations." Working Paper No. 323