



Make or Buy Decisions and Data Sharing

Antoine Dubus
ETH Zürich

Patrick Legros
Université Libre de Bruxelles

February 2026

ECARES working paper 2026-05

Make or Buy Decisions and Data Sharing*

Antoine Dubus[†] Patrick Legros[‡]

February 5, 2026

Abstract

Firms may share data to discover potential synergies between their data sets and algorithms, eventually leading to more efficient mergers and acquisitions (M&A) decisions. However, data sharing also modifies the competitive balance when firms do not merge, and a company may be reluctant to share data with potential rivals. Under general conditions, we show that firms benefit from (partially) sharing data. By doing so, they can merge conditionally based on high synergies. Compared to a laissez-faire situation, the presence of a regulator allowing or refusing the M&A may increase or decrease data sharing, with a concomitant increase or decrease in consumer surplus. Hence, regulation can *lower the surplus* of consumers it is willing to protect. We revisit the Google/Fitbit acquisition through the lens of this interplay between strategic data sharing and antitrust policy.

Keywords: Artificial Intelligence; Synergies; Mergers and Acquisition; Incomplete Information; Antitrust.

JEL: G34; K21; L1; L21; L24; L5; L86.

*We would like to thank Stefan Bechtold, Andreea Cosnita-Langlais, Jacques Crémer, David Gilo, Jan Krämer, Mathieu Parenti, Pierre Régibeau, Shiva Shekhar, Thibaud Vergé, Patrick Waelbroeck, as well as participants to the 2021 CRESSE conference, the 2022 IIOC conference, the 2023 DISS workshop, the 2024 Conference on Regulators and Regulation in the Digital Era, the 2024 IBEO workshop, the 2024 EARIE conference, the 2024 ECARES workshop, the Paris Seminar on the Digital Economics, and to the LIEN seminar for insightful comments and discussions. This project benefited from financial support from FNRS (Project PDR T.01.47.19) and the European Research Council (under the European Union’s Seventh Framework Program (FP7-IDEAS-ERC), ERC Grant Agreement OIO- 339950). Previous versions of this work have circulated under the titles “Data-Driven Mergers and Acquisitions with synergies”, and “The Sale of Data: Learning Synergies Before M&As.”

[†] Department of Management, Technology, and Economics, ETH Zürich Leonhardstrasse 21 Switzerland – 8092 Zürich, Switzerland; antoine1dubus@gmail.com.

[‡] Université Libre de Bruxelles, ECARES and CEPR; patrick.legros@ulb.be.

Pre-merger information gaps – whether from overoptimistic synergy assessments or uncertainty about data complementarities – can cause costly M&A failures. Reducing these gaps before an acquisition is valuable, but while information sharing improves the assessment of potential returns, it can also shift the competitive balance if the merger fails. To our knowledge, the interaction between learning about merger synergies and the effect of this learning on competition has not been studied in the literature.

We examine this issue through the lens of data sharing in the digital economy. Data sharing is increasingly common (Scaria, Berghmans, Pont, Arnaut and Leconte, 2018; Bhargava, Dubus, Ronayne and Shekhar, 2025), even among direct competitors (Choe, Cong and Wang, 2024). Digital firms seek vast amounts of data to train AI systems,¹ generate data-driven network effects (Hagi and Wright, 2023), and benefit from identification externalities across consumer groups (Bergemann, Bonatti and Gan, 2022). Data acquisition has become a major driver of many M&As (de Cornière and Taylor, 2024; Gautier and Maitry, 2024).² For instance, Facebook’s acquisitions of WhatsApp and Instagram were partly motivated by the desire to integrate user profiles across platforms to optimize targeted advertising. Similarly, in Apple’s acquisition of Shazam, the European Commission examined how access to the target’s data could change market power and whether combining the firms’ data sets could generate anti-competitive effects.³

Yet digital M&As are risky, and many fail to deliver expected benefits (Gautier and Lamesch, 2021).⁴ Synergies between firms’ data and algorithms are unknown before the merger – firms may only have a probabilistic estimate – and their actual value can be assessed only after integrating the target’s data and retraining algorithms.⁵ As Experian notes, data quality and data/algorithm synergies are increasingly central to digital M&As but also generate major uncertainty about merger outcomes.⁶ Accurately anticipating these synergies is thus crucial both for merging firms and for regulators seeking to protect consumers.⁷

¹For Data-Guzzling AI firms, the Internet Is Too Small, Wall Street Journal, April 1, 2024.

²For 2015–2021, Gautier and Maitry (2024) estimate that 30% of digital M&As by large tech firms were data-driven, especially among intensive AI users.

³Mergers: Commission clears Apple’s acquisition of Shazam, September 6, 2018.

⁴See also The Three Reasons Why Tech M&A Deals Fail to Deliver Value, Forbes, October 19, 2017. Examples include America Online/Netscape, Twitter/Vine, Yahoo/Tumblr, and Yahoo/Broadcast.com.

⁵While firms like Google and Facebook can infer the type of data owned by other digital firms, they cannot precisely predict synergies without actual integration and analysis. Unity, for example, suffered substantial losses after misjudging acquired data quality. In 2022, it lost \$110 million after training AI algorithms on poor-quality client data integrated into its own datasets (The Impact of Bad Data and Why Observability is Now Imperative, IBM, June 2, 2022.).

⁶The Critical Role of Data Quality in Mergers and Acquisitions, Experian, last accessed March 19, 2024.

⁷Make Data a Strategic Priority in Your Post-Merger Integration, PwC, last accessed August 20, 2024.

We focus on Google’s 2.1 billion USD acquisition of Fitbit, a data-intensive wearable firm, completed in 2021 ([Chen, Choe, Cong and Matsushima, 2022](#)). We study key strategic moves by Google and Fitbit before and after the deal through the lens of data sharing as a way to learn about complementarities between Fitbit’s products and software and Google’s platform ecosystem, as detailed in Section 5. The model also applies to other contexts where data sharing facilitates learning about complementarities.

We study a market with several incumbents selling products, where Google is considering acquiring one competitor, Fitbit. If the acquisition occurs, synergies may arise from combining their data and algorithms within Google’s platform ecosystem. Google is uncertain about profitability: high synergies make the deal attractive, while low synergies generate losses, making continued competition preferable.⁸ To better evaluate the deal, Fitbit can share part of its data with Google before the acquisition.

We offer positive and normative insights on how digital firms should design data and merger strategies, and on corresponding competition policies. Without merger control, firms prefer data sharing to merging under uncertainty. Thus, when data sharing is possible, firms should avoid M&As with incomplete information. Even when firms have a good estimate of potential synergies, they still gain from refining this estimate before merging. This implies that pre-merger data sharing to assess synergies is likely already common in data-driven M&As.

The growing importance of data in digital markets has led to calls for updated regulation to protect competition and consumers, especially in merger control ([Cr mer, de Montjoye and Schweitzer, 2019](#); [Tirole, 2023](#)). Our theoretical analysis identifies when data sharing and pre-merger learning arise and how they could shape optimal merger policy. Strict merger control can reduce consumer surplus relative to laissez-faire because it may deter data sharing and push firms to either merge under incomplete information or compete without sharing. With less data sharing, competition is weaker, harming consumers. The optimal policy combines merger control with mandatory data sharing, as in the European Data Act (Art. 32).⁹

The remainder of this article is structured as follows. Section 1 introduces the model, and Section 2 characterizes the equilibrium. Section 3 examines how regulatory interventions influence data sharing and surplus. Section 4 derives the corresponding policy implications.

⁸In data-driven mergers, losses often stem from privacy concerns when data is shared with firms like Google or Facebook (see [’Tossed my Fitbit in the Trash’: Users Fear for Privacy After Google Buys a Company, The Guardian, November 6, 2019](#)). Losses can also arise from integration costs, cultural clashes, or X-inefficiencies.

⁹[Data Act, December 22, 2023](#).

Section 5 provides an in-depth analysis of the Google/Fitbit acquisition, and Section 6 surveys the literature on data sharing and data-driven mergers and acquisitions (M&As). Section 7 concludes.

1 Model

We examine a market in which Fitbit (indexed by F) supplies wearable devices. Google (indexed by G) is initially not active in this specific product market, but earns profits in several technologically and commercially related markets and provides Cloud services to firms such as Fitbit. This framework reflects the empirical timeline of the Google/Fitbit merger: prior to the acquisition, Google did not operate in the hardware segment of the health-wearables industry (its Google Fit platform was predominantly software-based, and the Pixel product line did not yet comprise devices in direct competition with Fitbit).

At the initial stage of the interaction, firms $i \in \{F, G\}$ earn baseline profits denoted by π_i^0 .¹⁰ We adopt an indirect specification of the profit functions. This nonparametric approach permits a more transparent isolation of the economic mechanisms associated with data sharing. Appendix A derives explicit parametric representations of these profit functions within oligopoly models featuring linear demand, where the demand intercepts are functions of the extent of data sharing. The environment encompasses both (i) a regime in which Google is not yet active in the wearable devices market—corresponding to the actual Google/Fitbit setting and (ii) a counterfactual regime in which Google is an incumbent competitor in that market.

The sequence of actions is summarized in Figure 1. In the first stage, Fitbit chooses whether to migrate, and at what intensity to migrate, a subset of its application and user base—hereafter referred to as “sharing data”—to Google’s platform. The contractual arrangement between the firms specifies a transfer $T(s)$ as a function of the chosen data-sharing level s .

¹⁰This general framework accommodates an arbitrary number of firms competing with Fitbit in the wearable devices market. Fitbit’s profits can be parameterized by the number of competitors, or more generally by any measure of competitive intensity. For expositional clarity, we omit explicit reference to the number of firms in the profit functions.

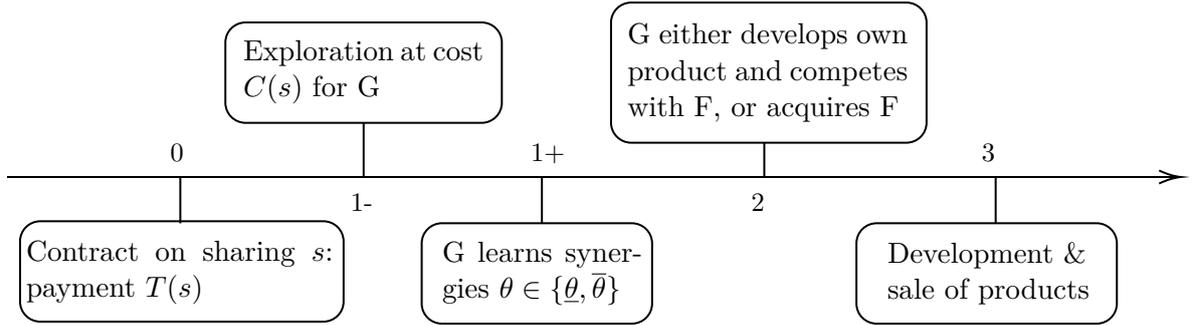


Figure 1: Sequence of events

For the sake of clarity, we decompose Google’s decision to incur an exploration cost, and the subsequent process of learning about synergies, into two sub-stages, denoted 1– and 1+. At stage 2, an information revelation problem arises: Google privately observes the value of the synergies, whereas this value is unknown to Fitbit. We demonstrate that this private information can be credibly disclosed, as Google will choose to propose an acquisition only when synergies are high, thereby enabling Fitbit to infer the value of the synergies from Google’s merger proposal.

Sharing and contracting. Firms possess a stock of data generated in the course of their economic activities. Fitbit can share a proportion $s \in [0, 1]$ of its data with Google, either unilaterally or in exchange of a lump-sum transfer $T(s)$. These data consist, for example, of health-related information produced through consumers’ use of Fitbit wearable devices (“including several billion nights of sleep”¹¹). Other pertinent examples, observed in past merger proceedings, include detailed user profiles derived from social media platforms (e.g., Facebook/WhatsApp) or traffic and navigation data in the context of mapping and navigation applications (e.g., Google/Waze).

Synergies and exploration. The use of data originating from Fitbit on Google’s platform ecosystem generates synergies of value $\theta \in \{\underline{\theta}, \bar{\theta}\}$, with $\bar{\theta} > \underline{\theta}$, which affect both firms’ profits and consumer surplus. These synergies may be operational—arising, for example, from efficiency gains when Fitbit’s systems are run on Google Cloud services—or may stem from positive cross-market data spillovers once Fitbit is integrated into Google’s broader platform ecosystem. The true realization of θ is initially unknown, but each firm shares a common prior according to which $\theta = \underline{\theta}$ with probability α and $\theta = \bar{\theta}$ with probability $1 - \alpha$.¹²

¹¹Fitbit insiders sweat (and dream) as Google swallows it whole. *Wired*, November 21, 2019.

¹²Our results extend to general probability distributions over a continuous support $\theta \in [\underline{\theta}, \bar{\theta}]$. For expositional clarity, we restrict attention to the binary case.

By conducting an exploratory analysis of Fitbit’s data, Google can learn the realized value of the synergies and, conditional on this information, develop a product that leverages Fitbit’s data. Inferring the total synergistic value of Fitbit’s data from a subset s entails an exploration cost $C(\cdot)$, which is a function of the volume or scope of data examined. Upon incurring this cost, Google privately observes the realization of θ .

The literature on statistics and machine learning shows that inferring information from missing data is costly and that models perform better with more data (see, e.g., [Neiswanger and Xing \(2017\)](#)). We therefore assume that the exploration cost $C(\cdot)$ of inferring synergies decreases with the amount of data explored. Assuming $C(1) = 0$ and the Inada conditions is for analytical convenience and can be relaxed without changing the qualitative results.

(H1) Exploration cost: $C(\cdot)$, $C''(\cdot) \geq 0$, $C(0) = +\infty$, $C'(0) = -\infty$ & $C'(1) = C(1) = 0$.

These specifications are consistent with recent empirical evidence on economies of scope and scale in data aggregation ([Hocuk, Martens, Prufer, Carballa Smichowski, Duch-Brown, Kumar and Mulder, 2022](#); [Schaefer and Sapi, 2023](#)). Other costs can arise, such as the structural costs of storing and analyzing the data or developing new algorithms. For simplicity, these development costs are embedded in the profit functions in the different outcomes of the game.

Google’s acquisition strategy. Google can acquire Fitbit either with imperfect information if no synergies are learned, or with perfect information after exploring Fitbit data post-acquisition. At the start of the game, firms thus face three options: compete without data sharing; Google acquires Fitbit without prior data sharing; or data sharing followed by exploration and then Google’s acquisition decision once the synergy state is known. We assume acquisitions are final and that divestment is prohibitively costly.¹³

Competition following data sharing. After sharing data and realizing the value of synergies, Google can enter the market with a stronger competitive position. The shared data lets Google improve its product through more accurate algorithmic predictions, personalized rec-

¹³In practice, divestment can follow an acquisition, and the model could be extended to allow for it. For example, Nokia bought Withings in 2016, rebranded its watches, and then divested in 2018 at the request of Withings’ previous owner, with watches again sold under the Withings name. See [Nokia is selling its digital health business back to the co-founder of Withings, The Verge, May 2, 2018](#). However, divestment is harder for data-driven mergers that raise consumer privacy concerns, which strongly affect how they value digital products. Such concerns are unlikely to disappear if, after the merger, synergies are low and firms separate. See, for instance, [‘Tossed my Fitbit in the trash’: users fear for privacy after Google buys a company, The Guardian, November 6, 2019](#).

ommendations, and better integration with its service ecosystem. These improvements increase Google’s appeal to consumers, boosting its demand and profits.

Arrow effect. The introduction of a superior product by Google reduces Fitbit’s market share and lowers its profit (Arrow, 1972; Anton and Yao, 2002). The magnitude of this adverse effect increases with s , since a larger volume of shared data enables Google to train more efficient machine-learning models and to implement more effective product improvements.

In the Google/Fitbit context, the launch of a competing wearable device (the Pixel Watch) was reportedly a serious strategic option for Google. Access to Fitbit’s health-activity data could enable Google to enhance the accuracy of health-monitoring functionalities, to further refine its recommendation and user-interface systems, and to exploit complementarities with its Android ecosystem.

More generally, under competition the profits of firms depend on the quality of Google’s product and, consequently, the extent of data sharing and the value of synergies. The profits under competition are denoted by $\pi_i^C(s, \theta)$, continuous in s for any given θ , and satisfy:

$$(H2) \quad \text{Arrow effect: } \frac{\partial \pi_G^C(s, \theta)}{\partial s} \geq 0 \geq \frac{\partial \pi_F^C(s, \theta)}{\partial s}.$$

Denoting $\pi^C(s, \theta) := \pi_F^C(s, \theta) + \pi_G^C(s, \theta)$ the sum of profits of Fitbit and Google under competition in state θ when Google has acquired s data from Fitbit, we assume that sharing data reduces $\pi^C(\cdot)$ when synergies are low, leaving the possibility that they increase with s when synergies are high.

$$(H3) \quad \text{Aggregate profits loss with low synergies: } \frac{\partial \pi^C(s, \theta)}{\partial s} \leq 0.$$

The machine-learning literature shows that firms’ total profits can fall when they share data. Tsoy and Konstantinov (2023) illustrate this with traffic data among ride-hailing firms, and Wu and Yu (2022) study payment data in banking. Had Google entered the wearable health market with its Pixel products, it would have intensified competition with incumbents like Fitbit, and Fitbit’s competitive losses could easily outweigh Google’s gains.

But, data sharing can also raise participating firms’ profits (Choe et al., 2024), for instance by enabling collusion or coordination. In that case, H3 would not hold and firms would choose to share data because it boosts their competitive profits. The tradeoff between competitive loss

and efficiency gain arises when data sharing lowers both Fitbit’s profit and the joint profit of Fitbit and Google, as guaranteed by [H2](#) and [H3](#). Under these conditions, firms are initially reluctant to share data.

In the limit case $s = 0$, firms make the baseline competitive profits (independent from θ):

$$(1) \quad \forall i, \theta, \pi_i^C(0, \theta) = \pi_i^0.$$

Data–synergy complementarities. When synergies are high, Google can more effectively exploit the shared data to enhance the quality of products and the sophistication of analytics within its platform ecosystem. To formalize this mechanism, we posit that the volume of data shared and the magnitude of synergies are complementary inputs, such that the marginal value of additional shared data increases with the level of synergies, and conversely, the marginal value of synergies rises with the amount of data shared.

$$(H4) \quad \text{Data-synergies complementarity: } \frac{\partial^2 \pi_G^C(s, \theta)}{\partial s \partial \theta} \geq 0.$$

This assumption is intuitive and supported by recent empirical evidence of super-additive returns to data when models are efficiently used across multiple tasks ([Varian, 2014](#)).

Acquisition decision. To acquire Fitbit, Google can either make an offer under incomplete information, or acquire data first, and then make its acquisition decision.

Acquisition with data sharing. If Google Acquires Fitbit after data has been shared, the merged firm makes profits denoted by $\pi^A(\theta)$, which depend on the value of the synergies.

For data sharing to create an option value, we focus on the case where the sum of profits of Fitbit and Google is higher under a merger than under competition only if synergies are high:

$$(H5) \quad \begin{aligned} \text{Option value: } & \forall s, \pi^A(\bar{\theta}) > \pi^C(s, \bar{\theta}), \\ & \& \pi^A(\underline{\theta}) < \pi^C(s, \underline{\theta}). \end{aligned}$$

Note that this assumption combined with (1) implies $\pi^A(\bar{\theta}) > \pi^A(\underline{\theta})$. The decision of the firms becomes trivial outside of [H5](#): Google either always wants to acquire Fitbit (if $\pi^A(\bar{\theta}) > \pi^C(s, \bar{\theta})$ and $\pi^A(\underline{\theta}) > \pi^C(s, \underline{\theta})$), or never does so (if $\pi^A(\bar{\theta}) < \pi^C(s, \bar{\theta})$ and $\pi^A(\underline{\theta}) < \pi^C(s, \underline{\theta})$).

When synergies are low, profit losses reflect merger costs, cultural shocks, and incompati-

bilities between the firms' information systems and data sets. An acquisition is beneficial only if synergies are high enough to outweigh these merger-induced losses.¹⁴ Data sharing therefore enables firms to avoid mergers when synergies are low. Obtaining this option is costly because Google pays the exploration cost $C(s)$ and, when data is shared and synergies are low, the sum of Fitbit's and Google's profits under competition falls under [H3](#).

Acquisition decision with Nash bargaining after data sharing. When Google explores Fitbit's shared data, it privately learns the synergy value $\theta \in \{\underline{\theta}, \bar{\theta}\}$, so acquisition negotiations occur under asymmetric information. There are three candidate equilibria: (i) separating, where Google offers to acquire Fitbit at price $p^A(s)$ only when synergies are high; (ii) pooling on merger, where Google offers $p^A(s)$ for any θ ; and (iii) pooling on competition, where Google makes no acquisition offer.¹⁵

We model profit division using Nash bargaining with incomplete information (following [Harsanyi and Selten, 1972](#); [Myerson, 1979](#); [Samuelson, 1984](#)). Fitbit's bargaining weight is $\mu \in [0, 1]$, and Google's is $1 - \mu$.¹⁶ In a separating equilibrium, the acquisition decision depends on θ and must satisfy an incentive-compatibility condition.

Acquisition under imperfect information. When firms merge under incomplete information, they earn $\pi^A(\bar{\theta})$ when synergies are high (prob. α) and $\pi^A(\underline{\theta})$ when synergies are low (prob. $1 - \alpha$). After the acquisition, Google can access all of Fitbit's data and faces no inference cost ($C(1) = 0$). Thus, the expected payoff from an acquisition under imperfect information is

$$\mathbb{E}[\pi^A; \alpha] := (1 - \alpha)\pi^A(\bar{\theta}) + \alpha \pi^A(\underline{\theta}).$$

Hence, an acquisition under incomplete information is beneficial only if

$$(2) \quad \mathbb{E}[\pi^A; \alpha] - \pi^0 \geq 0.$$

¹⁴For data-intensive digital products, the developer's identity is not neutral: a product developed by Google carries the "Google" label and inherits consumers' privacy concerns, which may lower its valuation after the acquisition.

¹⁵In practice, Google could also send Fitbit hard information about synergies via virtual data rooms, which disclose verifiable information such as algorithm performance and are routinely used in merger cases.

¹⁶For simplicity, we assume that if an acquisition occurs, Fitbit can infer whether Google explored the data by inspecting Google's financial records and observing the associated exploration costs.

As in the data sharing case, the expected merger gains are shared according to Nash bargaining. Fitbit has a bargaining weight μ , and Google a weight $1 - \mu$, and the firms have Nash payoffs:

$$\begin{aligned} u_F(\alpha) &:= \pi_F^0 + \mu(\mathbb{E}[\pi^A; \alpha] - \pi^0), \\ u_G(\alpha) &:= \pi_G^0 + (1 - \mu)(\mathbb{E}[\pi^A; \alpha] - \pi^0). \end{aligned}$$

In practice, Google offers a price of $p^A = u_F(\alpha)$ to Fitbit for its acquisition and gets its Nash payoff $\mathbb{E}[\pi^A; \alpha] - p^A = u_G(\alpha)$. A M&A is efficient for Google and Fitbit if condition (2) is satisfied and the expected payoff of a merger under imperfect information is higher than the sum of profits of Fitbit and Google without sharing. It is clear that without sharing, Google will acquire Fitbit only when the acquisition is ex-ante efficient.

2 Equilibrium Analysis

As given in Figure 1, the timing is as follows: in the first stage, Google chooses to either stay uninformed, or to incentivize Fitbit from sharing s data for transfer $T(s)$, in which case it explores the data at a cost $C(s)$. In the second stage, Google chooses whether to compete or to acquire Fitbit, and in a separating equilibrium, Fitbit infers the value of θ from this decision.

We consider sequential equilibria of this game, and as usual, we proceed backward in our analysis. Hence, we first consider Google that acquires data from Fitbit. We characterize the profits of each firm if synergies are high and firms merge, and in case synergies are low and firms compete. We then derive Google's expected payoff and characterize the optimal amount of data it acquires. Finally, we provide conditions for data sharing to yield greater profits than competing without sharing and than merging under incomplete information.

2.1 Profits with Data Sharing

Stage 2: profits with sharing. Suppose Fitbit shares s with Google, who explores the data (at a cost $C(s)$). Because Google learns θ privately, there are three candidates for equilibrium to consider: (i) separating; (ii) pool-to-acquire; (iii) pool-to-compete.

Even if pool-to-acquire is stable, it is less profitable for the firms than a merger with incomplete information. Indeed, although in both regimes the expected surplus from the acquisition would be identical, Google would incur the exploration cost $C(s)$ if it acquires data from Fitbit. For this reason we ignore this situation and focus on separation or pooling to compete.

The complementarity property H4 is necessary for a separating equilibrium to exist. If a separating equilibrium exists, a pool-to-compete is unstable under standard refinements based on type-dominance because there is a profitable deviation by $\bar{\theta}$ that is not profitable for $\underline{\theta}$. If a separating equilibrium is not possible, a pool-to-compete strategy can be made stable.

By standard arguments, it is simple to show that the necessary condition for separation is $\pi^A(\bar{\theta}) - \pi^A(\underline{\theta}) \geq \pi_G^C(s, \bar{\theta}) - \pi_G^C(s, \underline{\theta})$. The left hand side is positive and the right hand side is an increasing function of s by the complementarity condition H4.

Proposition 1. *There exists $\sigma \in (0, 1]$ such that separation is the unique equilibrium if and only if $s \leq \sigma$, and pool-to-compete is the only equilibrium if $s > \sigma$. If there is separation, letting $p^*(s) = \pi_F^C(s, \bar{\theta}) + \mu(\pi^A(\bar{\theta}) - \pi^C(s, \bar{\theta}))$, the acquisition price in state $\bar{\theta}$ is*

$$p^A(s) = \begin{cases} p^*(s) & \text{if } p^*(s) \geq \pi^A(\underline{\theta}) - \pi_G^C(s, \underline{\theta}), \\ \pi^A(\bar{\theta}) - \pi_G^C(s, \bar{\theta}) & \text{if } p^*(s) \leq \pi^A(\underline{\theta}) - \pi_G^C(s, \underline{\theta}). \end{cases}$$

Proof. See the Appendix. □

When $\sigma \in (0, 1)$, if the firms choose share data they will either share $s < \sigma$ and merge only when synergies are high, or share $s > \sigma$ and compete regardless of the value of θ . The resulting profits are provided in (9) of Section 3.3, where we analyze the equilibrium with a regulator.

For the rest of the present section, we focus on $\sigma = 1$, so that after exploration there is acquisition only if the synergies are high.

Stage 1: sharing decision. Anticipating an acquisition when synergies are high, Fitbit makes expected profits when sharing s data equal to $(1 - \alpha)p^A(s) + \alpha \pi_F^C(s, \underline{\theta})$. According to the Arrow effect, Fitbit may incur a competitive loss from data sharing, in which case Google pays $T(s)$ to compensate for this loss and to incentivize the sharing of s data.

In practice, this monetary transfer can take the form of a discount Google offers to Fitbit for the use of its cloud services. What matters is that, compared to a usual cloud client, Google may need to incentivize Fitbit to share its data because Fitbit anticipates the risk from expropriation and the resulting competitive losses in case Google chooses to enter the market and compete.

The value of this transfer depends on the outside option for Fitbit: either competing and making profits π_F^0 , or merging under incomplete information and making expected profits $u_F(\alpha) = \pi_F^0 + \mu(\mathbb{E}[\pi^A; \alpha] - \pi^0)$.

Lemma 1. *Google pays a transfer $T(s)$ to Fitbit to share s data, equal to:*

$$T(s) = \pi_F^0 + \max\{\mu(\mathbb{E}[\pi^A; \alpha] - \pi^0), 0\} - (1 - \alpha)p^A(s) - \alpha \pi_F^C(s, \underline{\theta}).$$

Google is willing to acquire data only if it can recover the exploration cost $C(s)$ and the price of the data $T(s)$. Note that $T(s)$ can be negative, for instance if the price for the acquisition $p^A(s)$ is high. In such case, Fitbit would be willing to share data for free, and the game simplifies to the acquisition decision stage. The optimal amount of data shared by Fitbit then maximizes the ex-ante sum of profits of Fitbit and Google:¹⁷

$$(3) \quad \Pi(s, \alpha) := (1 - \alpha)\pi^A(\bar{\theta}) + \alpha \pi^C(s, \underline{\theta}) - C(s).$$

Overall, ex-ante sharing creates an option for the firms when making their merger decision. Firms will be able to merge only when synergies are high, and the merger is efficient, thereby avoiding inefficient mergers when synergies are low.

Optimal Sharing. An increase in s has two opposite effects on the expected profits of Google. On the one hand, more data is costly to acquire from Fitbit. Under Assumption H3, $\frac{\partial \pi^C(s, \underline{\theta})}{\partial s} < 0$ and the potential gains for Google in acquiring s data in competition mode when synergies are low are dominated by the loss of Fitbit, captured by transfer $T(s)$. To minimize this loss, Google needs to learn the value of synergies using the smallest possible amount of data. On the other hand, more shared data decreases the exploration cost $C(s)$. Under H1, using only a tiny share of data to learn synergies induces a high cost, which is not profitable for Google.

The optimal amount of data shared balances these two opposite effects of data sharing on Google's profits. The equilibrium amount of data shared $s^*(\alpha)$, which depends on α , satisfies

$$(4) \quad \alpha \frac{\partial \pi^C(s^*(\alpha), \underline{\theta})}{\partial s} = C'(s^*(\alpha)).$$

For simplicity, we assume that a unique $s^*(\alpha)$ satisfies this condition, for example, when $\pi^C(s, \underline{\theta})$ is quasi-concave in $s \in (0, 1)$.

¹⁷We show in Appendix C that this aligns with Google's expected profits at the beginning of the game.

2.2 Equilibrium Data Sharing

Data will be shared before the merger if the resulting maximum expected payoff from the industry (with $\Pi^*(\alpha) := \Pi(s^*(\alpha), \alpha)$) is greater than (i) the expected profits of merging under imperfect information $\mathbb{E}[\pi^A; \alpha]$ and (ii) the competitive profits of the firms π^0 :

$$\Pi^*(\alpha) \geq \max\{\mathbb{E}[\pi^A; \alpha], \pi^0\}.$$

Data sharing is profitable if it yields a higher expected payoff than the best alternative for Google, i.e., merging under incomplete information or competing. If competition dominates a merger under incomplete information, sharing can allow M&As when synergies are high. Otherwise, sharing allows firms to avoid mergers when synergies are low. In the rest of this section, we first show that Google's profits are greater with data sharing than when merging under incomplete information. We then provide conditions for data sharing to dominate competition.

Sharing data dominates merging under incomplete information. Sharing data gives Google the option to avoid mergers when synergies are low, yielding profits that are always greater than a merger under incomplete information. When data is shared, the exploration cost and competitive loss in case synergies are low are offset by the gains from merger avoidance. Hence, without merger control, M&As with incomplete information don't happen if data can be shared, and mergers occur with prior data sharing.

Proposition 2. *Data sharing is more profitable than merging under incomplete information.*

Proof. We can rearrange the expression of the expected industry profits with data sharing:

$$\Pi(s, \alpha) = \mathbb{E}[\pi^A; \alpha] + \underbrace{\alpha(\pi^C(s, \underline{\theta}) - \pi^A(\underline{\theta}))}_{\text{Gains from merger avoidance if } \underline{\theta}, \text{ H5}} - \underbrace{C(s)}_{\text{Exploration cost, H1}}.$$

From (15), we know that the equilibrium amount of data shared $s^*(\alpha)$ maximizes $\Pi(s, \alpha)$. Hence, $\Pi^*(\alpha) \geq \Pi(1, \alpha) > \mathbb{E}[\pi^A; \alpha]$, which implies $\Pi^*(\alpha) \geq \mathbb{E}[\pi^A; \alpha]$. \square

Merging vs. competing. While industry profits π^0 are independent of the degree of uncertainty of synergies, this is not the case for $\mathbb{E}[\pi^A; \alpha]$ and $\Pi(s, \alpha)$, which depend on the value of α . In particular, $\mathbb{E}[\pi^A; \alpha]$ and $\Pi(s, \alpha)$ are the highest when $\alpha = 0$, in which case they both are

equal to $\pi^A(\bar{\theta})$.¹⁸ Therefore, a necessary condition for a merger to take place is

$$(5) \quad \pi^A(\bar{\theta}) \geq \pi^0.$$

Otherwise, firms always prefer competition to mergers and data sharing. In our setting this condition is satisfied since $\pi^C(0, \theta) = \pi^0$, and $\pi^A(\bar{\theta}) > \pi^C(0, \theta)$.

When α increases, the probability of low synergies increases and $\mathbb{E}[\pi^A; \alpha]$ and $\Pi(s, \alpha)$ decrease, with $\Pi(s, \alpha) > \mathbb{E}[\pi^A; \alpha]$ for $\alpha > 0$. In the limit case where $\alpha = 1$, **H3** and **H5** ensure that a merger is never profitable and firms always compete with π^0 .

By continuity, there exists α_1 such that firms are indifferent between merging under incomplete information and competing. Moreover, there exists α_2 , with $0 < \alpha_1 < \alpha_2 < 1$ such that firms share data before a merger if and only if $\alpha \leq \alpha_2$. When $\alpha < \alpha_1$, sharing data dominates merging under incomplete information which is the best outside option. In this case, data sharing allows firms to avoid mergers when synergies are low. When $\alpha_1 < \alpha < \alpha_2$ sharing data dominates competition and enables M&As to take place when synergies are high. When $\alpha > \alpha_2$, firms do not merge. These results are illustrated in Figure 2.

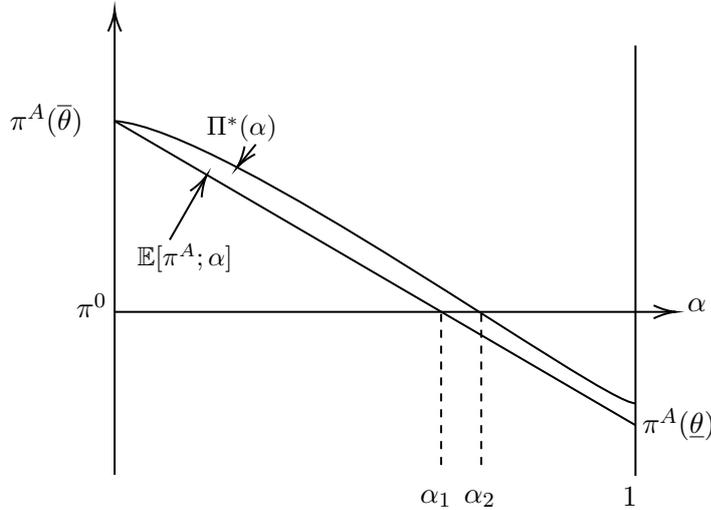


Figure 2: Equilibrium gains from a merger without regulation.

Proposition 3. *There exist α_1 and α_2 , with $0 < \alpha_1 < \alpha_2 < 1$ such that:*

- (i) *data sharing happens if, and only if, $\alpha \leq \alpha_2$,*
- (ii) *sharing prevents firms from merging when synergies are low if $\alpha \leq \alpha_1$ and allows mergers to occur when synergies are high if $\alpha \in (\alpha_1, \alpha_2]$.*

¹⁸When $\alpha = 0$, $\Pi(s, \alpha)$ is equal to $\pi^A(\bar{\theta}) - C(s)$. The firms share all data ($s^*(0) = 1$) and incur an exploration cost equal to $C(1)$, so that $\Pi(1)$ in this case is equal to $\pi^A(\bar{\theta})$.

Firms can gain from data sharing even with an excellent estimate (α near one or zero). In this case, sharing almost all data is optimal because the likelihood that firms compete is low. Sharing and exploring offer the choice not to merge if synergies are low, with minimal exploration cost as $s^*(\alpha)$ is close to one. If the cost to learn synergies is always positive ($C(1) > 0$), merging with incomplete information is optimal for α close to zero only if Google has a high enough bargaining power (μ is small). In such cases, sharing is best for intermediate values of α . Data sharing expands the scope of profitable M&As, enabling mergers when $\alpha_1 < \alpha < \alpha_2$. Thus, pre-merger data-sharing should increase efficient M&A deals.

3 Data sharing in the Shadow of a Regulator

We introduce a competition authority that approves only M&As increasing consumer surplus.¹⁹ We examine how the regulator affects firms' data sharing before a merger, showing that firms might share less data compared to a laissez-faire scenario, which can reduce consumer surplus. We discuss regulatory options to mitigate this negative effect of merger control.

3.1 Consumer Surplus

Without data sharing, CS^0 represents the consumer surplus in the baseline competitive market. When a fraction s of the data is shared and Google and Fitbit engage in competition in state θ , consumer surplus becomes $CS^C(s, \theta)$. We assume that $CS^C(s, \theta)$ increases with s .

$$(H6) \quad \text{Consumer surplus: for all } (s, \theta), \frac{\partial CS^C(s, \theta)}{\partial s} \geq 0.$$

When $s = 0$, the baseline consumer surplus does not depend on the value of the synergies.

$$\forall \theta, CS^C(0, \theta) = CS^0.$$

Surplus after a M&A. A M&A changes competition and allows the merged firm to offer a product quality based on synergies. Hence, surplus post-M&A depends on the value of synergies: $CS^A(\theta)$. Compared to the baseline surplus CS^0 , a M&A benefits consumers only with high synergies; otherwise, it reduces their surplus. (Without this assumption, mergers are

¹⁹A regulator could also include industry profits in its M&A decision. This could capture long-term considerations, as the firms' profits today change their ability to invest in innovation, which will later benefit consumers. Our qualitative results are robust to this type of alternative merger rule.

either always blocked – when $CS^0 > CS^A(\bar{\theta})$ – or authorized – when $CS^A(\underline{\theta}) > CS^0$.)

$$(H7) \quad CS^A(\bar{\theta}) > CS^0 > CS^A(\underline{\theta}).$$

If Fitbit has shared all data and synergies are high, we assume that surplus is greater under competition than after a M&A. In this case, consumers benefit from synergies and competition.

$$(H8) \quad CS^C(1, \bar{\theta}) > CS^A(\bar{\theta}).$$

These assumptions are consistent with our micro-foundations in which consumers benefit from better product quality when synergies are high and from a higher intensity of competition between firms when s increases; see Appendix [A.1](#).

3.2 Merger Control

The regulator assesses M&A proposals by comparing surplus if the merger proceeds or is blocked, considering pre-merger data sharing and the value of synergies. Let

$$G(s, \theta) := CS^A(\theta) - CS^C(s, \theta)$$

be the gain in consumer surplus from a M&A if firms share s and synergies are θ .

Merger control with data sharing. When the regulator assesses a M&A, Fitbit has already shared s with Google. Thus, in a separating equilibrium Google will seek approval for the M&A if synergies are high, and as for Fitbit, the regulator can deduce the value of θ from Google's merger request. Thus, the change in consumer surplus does not depend on the value of α .

To evaluate changes in surplus when data is shared, the regulator must know the value of s . Given that sharing s data requires a monetary transfer, the details of the transfer and the data shared are recorded in the firms' financial statements and accessible to the regulator during the review of the merger.

The regulator approves a merger if $G(s, \bar{\theta})$ is nonnegative. Assumption [H7](#) guarantees that M&As are approved if s is not excessively high. Furthermore, [H8](#) indicates that if the firms share too much data, the surplus under competition will be high, prompting the regulator to

block a M&A. Consequently, as $G(s, \bar{\theta})$ decreases with s , the regulator approves a M&A if,

$$s \leq \hat{s}^A,$$

where $\hat{s}^A \in [0, 1]$ is the minimum quantity of data that satisfies $CS^C(\hat{s}^A, \bar{\theta}) = CS^A(\bar{\theta})$.

Merger control without data sharing. Without sharing, firms and the regulator anticipate an expected gain from merging of:

$$(6) \quad \begin{aligned} \mathbb{E}[G(0, \theta)] &:= (1 - \alpha)G(0, \bar{\theta}) + \alpha G(0, \underline{\theta}) \\ &= (1 - \alpha)CS^A(\bar{\theta}) + \alpha CS^A(\underline{\theta}) - CS^0, \end{aligned}$$

and the regulator allows the merger if $\mathbb{E}[G(0, \theta)]$ is non-negative. This expression decreases with α under [H7](#): it is equal to $CS^A(\bar{\theta}) - CS^0 > 0$ when $\alpha = 0$ and to $CS^A(\underline{\theta}) - CS^0 < 0$ when $\alpha = 1$. Hence, there exists a cut-off value.

$$(7) \quad \alpha^R = \frac{CS^A(\bar{\theta}) - CS^0}{CS^A(\bar{\theta}) - CS^A(\underline{\theta})} \in (0, 1)$$

such that mergers are accepted with incomplete information only when α is below α^R .

As shown in [Proposition 3](#), merging with incomplete information yields higher profits than competition when $\alpha \leq \alpha_1$. Hence, the expected profit of Fitbit and Google without sharing is

$$(8) \quad \Pi_{noshare}^R(\alpha) := \begin{cases} \mathbb{E}[\pi^A; \alpha] & \text{if } \alpha \leq \min\{\alpha^R, \alpha_1\} \\ \pi^0 & \text{if } \alpha > \min\{\alpha^R, \alpha_1\} \end{cases}$$

3.3 Equilibrium Strategies

When data is shared, the expected profits of Fitbit and Google depend on whether the regulator will accept or block the merger.

$$(9) \quad \begin{aligned} \Pi(s, \alpha) &= (1 - \alpha)\pi^A(\bar{\theta}) + \alpha \pi^C(s, \underline{\theta}) - C(s), \quad \text{if the merger is approved,} \\ \Pi_B(s, \alpha) &= (1 - \alpha)\pi^C(s, \bar{\theta}) + \alpha \pi^C(s, \underline{\theta}) - C(s), \quad \text{if the merger is blocked.} \end{aligned}$$

Function Π_B is also the expected profits in a pooling-to-compete equilibrium. Thus, the analysis in the following generalizes our baseline model to $\sigma \leq 1$. In this case, firms can either share

$s \leq \sigma$ such that the equilibrium is separating, and make profits $\Pi(s, \alpha)$. Or they can share $s > \sigma$, leading to a pooling equilibrium where firms compete regardless of θ and make profits $\Pi_B(s, \alpha)$. The presence of a regulator will constrain the strategies of the firms if they anticipate that, by sharing $s < \sigma$, their merger request will be blocked.

Let Σ be the set of levels of data sharing that induce the regulator to approve a merger with high synergies.²⁰ Let $s_A^*(\alpha) := \operatorname{argmax}_{s \in \Sigma} \Pi(s, \alpha)$ and $s_B^*(\alpha) := \operatorname{argmax}_{s \notin \Sigma} \Pi_B(s, \alpha)$. If $s^*(\alpha) \in \Sigma$, regulation has no impact as firms choose $s_A^*(\alpha) = s^*(\alpha)$. If $s^*(\alpha) \notin \Sigma$ (i.e., $s^*(\alpha) > \hat{s}^A$), firms can either follow a ‘‘puppy dog’’ strategy by sharing \hat{s}^A to get approval or accept that the merger will be blocked and choose the best non-approval sharing. Firms will share data leading to a merger when synergies are high if $\Pi(s_A^*(\alpha), \alpha) \geq \Pi_B(s_B^*(\alpha), \alpha)$. Denote

$$\Pi_{share}^R(\alpha) := \max\{\Pi(s_A^*(\alpha), \alpha), \Pi_B(s_B^*(\alpha), \alpha)\}$$

the maximum profits with a regulator. Firms only share data if $\Pi_{share}^R(\alpha) \geq \Pi_{noshare}^R(\alpha)$. With a regulator, firms earn less and prefer to merge under incomplete information when α is small or to compete without sharing when α is near α_2 . Thus, regulation can discourage sharing, leading firms to merge under incomplete information or compete without sharing.

Lemma 2. *Under regulation, there exist α_0 and α_2^R such that:*

- *firms merge under incomplete information instead of sharing when $\alpha \in [0, \alpha_0)$;*
- *firms compete instead of sharing when $\alpha \in [\alpha_2^R, \alpha_2)$, with $\alpha_2^R < \alpha_2$ when regulation is binding, i.e., when $s^*(\alpha_2) > \hat{s}^A$.*

The regulator may also block M&As under incomplete information but allow them after data has been shared. In this case, firms can convince the regulator of the benefits of a merger by sharing data and ask for a merger only if synergies are high.

Lemma 3. *There exists α^R such that firms can convince the regulator to allow the merger in case synergies are high by sharing data when $\alpha \in (\alpha^R, \alpha_2^R]$ (with $\alpha^R < \alpha_2^R$ if $\Pi^*(\alpha^R) \geq \pi^0$), while without sharing, the regulator would block a M&A.*

Figure 3 below illustrates how regulation changes the firms’ equilibrium strategies.²¹

²⁰For instance, without regulatory oversight, $\Sigma = [0, 1]$. When regulation protects consumer surplus, $\Sigma := \{s : G(s, \bar{\theta}) \geq 0\}$, and there exists $\hat{s}^A \in [0, 1]$ such that $\Sigma = [0, \hat{s}^A]$.

²¹We focus on $\alpha < \alpha_2$, since otherwise, competing without sharing is the dominant strategy anyway and the presence of the regulator does not change the equilibrium outcome.

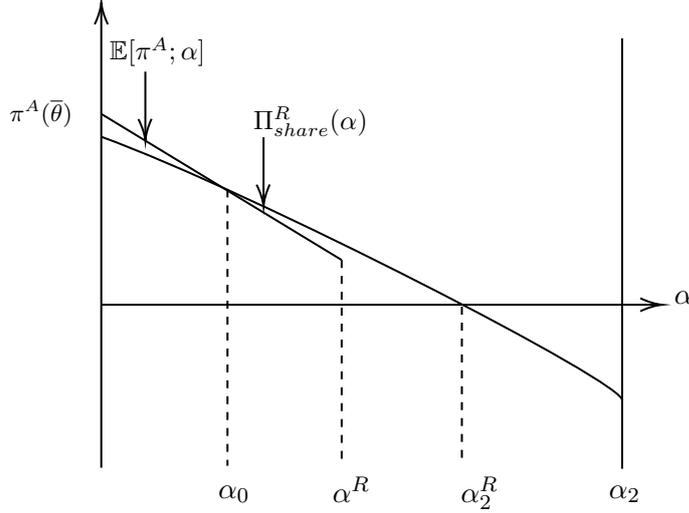


Figure 3: Equilibrium merger gains under regulation.

When firms share data, $\Pi_{share}^R(\alpha)$ depends on whether they share and merge or share and compete. For sharing and competing to be profitable, it must yield higher profits than competing without sharing. This requires $\pi^C(s, \bar{\theta})$ to be greater than π^0 for some values of s .

In turn, firms will prefer to share and merge or share and compete depending on the value of the threshold \hat{s}^A above which mergers are blocked after the data has been shared, which determines the value of profits $\Pi(s_A^*(\alpha), \alpha)$ and $\Pi_B(s_B^*(\alpha), \alpha)$.

Lemma 4. *Firms share and compete only if $\pi^C(s, \bar{\theta}) > \pi^0$ for some s . There exist α^B , $\alpha_2^B < \alpha_2$ and $\hat{s}^A(\alpha)$ such that sharing data and competing is the equilibrium strategy if and only if $\alpha \in (\alpha^B, \alpha_2^B)$ and $\hat{s}^A < \hat{s}^A(\alpha)$.*

Data sharing between firms intensifies competition while allowing Google to provide consumers with a higher utility, and regulation can lower consumer surplus by reducing firms' incentives to share. In contrast, if firms choose to share more data, anticipating that the merger will be blocked, merger control will benefit consumers.

Proposition 4. *Suppose the regulator authorizes a merger only if it increases consumer surplus.*

- *If firms share data and request a merger, the amount of data shared and consumer surplus are smaller than in the absence of a regulator.*
- *If firms do not share data, the expected consumer surplus is smaller than in the absence of a regulator.*
- *If firms share data and the merger is blocked, consumer surplus is higher than in the absence of a regulator if firms share more data.*

By sharing less data, firms reduce consumer surplus under competition and increase the surplus gain from M&A approval. The resulting impact on expected consumer surplus is therefore negative. This effect is the strongest when regulation discourages firms from sharing, such that they either merge under incomplete information or compete instead of sharing. Since firms are constrained in their sharing strategies, the presence of the regulator also reduces the sum of profits of Fitbit and Google. Therefore, no-regulation can Pareto dominate a situation where a regulator is willing to protect consumers through merger control.

Regulation can benefit consumers by increasing firms' incentives to share data when they know a merger will be blocked anyway. On the one hand, more data shared increases the surplus of consumers when firms compete. However, the regulator blocks the merger precisely in cases where competition yields a higher surplus. For both reasons, regulation benefits consumers when the regulator blocks the merger after sharing has taken place.

3.4 Optimal Regulation

We consider different regulatory tools recently discussed or implemented in the case of data-driven mergers: mandated data-sharing rules; variations in the strength of merger control; and data silos. These regulations aim at protecting consumer surplus at different levels of the competition and merger processes. Mandated data sharing rules aim at promoting competition; merger control stringency protects consumers at the M&A stage; data silos aim at protecting consumer surplus once a merger goes through. Our contribution will be to explain the subtle effects of each regulation on the overall competition and merger processes.

3.4.1 Mandated Data Sharing

Merger control can decrease consumer surplus by altering firms' incentives to share data. In this section, we explore how the regulator can influence the amount of data shared by enforcing mandated data-sharing policies. The regulator anticipates how mandating data sharing will affect firms' competitive profits and consumer surplus, as well as how these factors shape its decision to approve or block the merger. Hence, the regulator chooses $s_A \in \Sigma$ to maximize $\mathbb{E}[CS(s_A)]$ if it anticipates that it will allow the merger if synergies are high, with

$$\mathbb{E}[CS(s_A)] := (1 - \alpha)CS^A(\bar{\theta}) + \alpha CS^C(s_A, \theta).$$

Anticipating that it will block a merger, the regulator requires the firms to share $s_B \notin \Sigma$ to optimize $\mathbb{E}[CS(s_B)]$, where

$$\mathbb{E}[CS(s_B)] := (1 - \alpha)CS^C(s_B, \bar{\theta}) + \alpha CS^C(s_B, \underline{\theta}).$$

When the regulator has control over the data shared by firms, it can prevent surplus losses due to changes in s . In particular, the regulator can strategically select s to maximize the expected surplus, which implies that a mandated data sharing policy sets $s = 1$.

Proposition 5. *Regulatory oversight, merger control, and a mandated data-sharing policy yield a greater consumer surplus than a laissez-faire situation. Full data sharing is surplus-maximizing, despite potential industry losses, if the regulator is sufficiently oriented toward protecting consumers.*

Mandated data sharing rules have recently been enacted in the European Data Act (Art. 35).²² The Digital Markets Act recently enacted in Europe also enforces a data access right under Article 61, under which large digital gatekeepers are “required to provide access, on fair, reasonable and non-discriminatory terms, to [...] ranking, query, click and view data in relation to free and paid search generated by consumers on online search engines”.

Our analysis suggests that these regulations can complement merger control when firms are willing to engage in data-driven M&As. Consider Google accessing data from Fitbit through Art. 35 of the Data Act before the merger, and using it to learn the value of synergies. Although the Data Act “prohibit[s] the use of data accessed under this Regulation for developing a competing connected product”, such that Google cannot directly use the data to enhance the quality of its Pixel product line, it can use it to improve its products in other markets where data has cross-market value. Such markets include Google mapping and location apps, as Fitbit’s data sets include precise geo-localization of its users. This generates a data-feedback loop (Bhargava et al., 2025), as, thanks to Fitbit’s data, Google can reach more consumers in other markets, collect more data there and, in turn, enhance the quality of its Pixel products.

Our results then suggest that mandated data sharing regulation can have a causal impact on firms’ ability to learn synergies and engage into a merger.

²²Data Act, December 22, 2023.

3.4.2 Merger Control Stringency

Regulators can vary in leniency regarding the surplus gain requirements for mergers to be approved. They might allow mergers that slightly lower consumer surplus if the increase in the Herfindahl index is minimal, as seen in horizontal merger guidelines.²³ Alternatively, a regulator might be strict, approving only mergers that significantly increase surplus. This prompts the question of the ideal level of “stringency” in this context.

To model this, assume $r \in \mathbb{R}$ such that a merger is approved when the gain $G + r$ is nonnegative. (A lenient regulation has r positive, while a strict one has r negative.) That is, $G(s, \bar{\theta}) + r \geq 0$ when firms share and explore, or $\mathbb{E}[G(0, \theta)] + r \geq 0$ when firms do not share data. As r increases, the set of values of s leading to merger approval $\Sigma(r) := \{s : G(s, \bar{\theta}) + r \geq 0\}$ becomes larger, and the equilibrium bound $\alpha^R(r)$ also increases, making mergers with incomplete information more frequent.

When firms share data, condition $G(s, \bar{\theta}) \geq -r$ influences the decision of firms to share and merge or share and compete, and $\Pi_{share}^R(\alpha)$ decreases with r . When firms share $s_A^*(\alpha)$, a reduction in r diminishes consumer surplus due to a decrease in the extent of data sharing between firms. By continuity of $\Pi_{share}^R(\alpha)$, there exists a value r_0 such that firms opt to share even though the merger will be blocked if and only if $r < r_0$. Therefore, if r becomes less than r_0 , more stringent regulations benefit consumers, as they encourage firms to share more data and compete, regardless of the value of synergies.

Firms may also not share data because of regulation. In line with our previous analysis, there exist thresholds $\alpha_0(r)$ (increasing with r) above which $\Pi_{share}^R(\alpha) > \mathbb{E}[\pi^A; \alpha]$, and $\alpha_2^R(r)$ (decreasing with r) above which $\pi^0 > \Pi_{share}^R(\alpha)$.

When $\alpha < \alpha_0(r)$, firms merge under incomplete information; when $\alpha \in [\alpha_2^R(r), \alpha_2]$, firms compete without sharing. There, the regulator can increase consumer surplus by setting r high enough for mergers to be accepted with $s^*(\alpha)$, so that firms benefit from sharing. We illustrate this result in the quality competition model analyzed in Appendix A.

²³Alternatively, the regulator could internalize the long-term effects of mergers on innovation.

3.4.3 Merger Remedy: Data Silos

To enable positive merger outcomes – e.g. from the complementarity of patent portfolios –, while mitigating adverse effects on privacy and data-driven market dominance, the European Commission has implemented data silo remedies in recent mergers and acquisitions, such as Facebook/WhatsApp and Google/Fitbit. Namely, the merger is allowed, but the combination of data sources is prohibited (Krämer and Shekhar, 2025).

In the following, we investigate how data silos impact the incentives of firms to engage in data sharing. With an exogenous probability γ a silo remedy is required, in which case the data sets of the firms remain separated after a M&A, and the merged firm cannot fully exploit the synergies. We define post-merger profits in this scenario as $\pi^A(s, \bar{\theta})$, where s is the amount of data shared before firms merge and silos are requested, and we assume that $\frac{\partial \pi^A(s, \bar{\theta})}{\partial s} > 0$.

In the presence of high synergies, firms merge their data with a probability of $(1 - \gamma)$, generating a profit of $\pi^A(\bar{\theta})$. They may also be forced to maintain data separated with probability γ , which results in profits equal to $\pi^A(s, \bar{\theta}) > \pi^C(s, \bar{\theta})$. Consequently, Google may have interest to acquire Fitbit when high synergies are anticipated, with expected industry gains equal to:

$$(10) \quad \Pi(s, \alpha) - (1 - \alpha)\gamma [\pi^A(\bar{\theta}) - \pi^A(s, \bar{\theta})]$$

A silo remedy reduces the expected profits of the firms, but enhances the incentives of Google to acquire data, as $\pi^A(s, \bar{\theta})$ increases with s .

Proposition 6. *Data-silos remedies increase the sharing of data.*

4 Policy Implications

Our results contribute to a growing debate among academics and policymakers on regulating data sharing practices among firms and how policies can affect market competition (Jullien and Sand-Zantman, 2021; Bhargava et al., 2025). The regulatory tools that we consider target different stages of the competition (mandated data sharing rules) and merger processes (pre-merger regulation with optimal merger control stringency, and post-merger data silos remedies). An important element of our contribution is to show how these targeted regulations have indirect impacts on welfare at other stages of competition and merger processes by changing firms' data strategies.

Consider first the data silo regulation. While this policy aims at protecting consumers after a merger, it impacts the incentives of the firms to share data before a potential M&A. We have shown that data silos lower the expected payoffs of a merger, and may discourage firms from merging overall. This is to the detriment of consumers if firms choose not to share data.

The Google/Fitbit acquisition illustrates well how silo regulations can also harm consumers after the M&A, by setting constraints on the sharing decisions of a firm. The data collected via Fitbit are subject to the silo regulation, but the data connected through Pixel Watch aren't. So, Google has incentives to promote its Pixel Watch, to the detriment of Fitbit's products. This could explain why Google is progressively turning off prominent features of the Fitbit products, while keeping them on Pixel products.²⁴

Turning to merger control, the novelty of our approach is to highlight a new source of surplus gains that competition authorities should consider when making merger approval decisions, by eliciting a novel rationale for pro-competitive data sharing. To evaluate the impacts of mergers on consumers, policymakers must look for evidence of data-sharing deals between firms. Because data sharing requires a money transfer, the details of the transfer and the amount of data are included in the accounting balance of both firms and should therefore be visible to a competition authority.

When firms engage into data sharing at large scale, typically in the Google/Fitbit case where Fitbit was using Google Cloud Healthcare API before the merger, regulators should carefully assess the terms of such data sharing agreements.²⁵ If these terms leave room for Google to engage into analytics on the firm's data that can allow it to learn the value of synergies, such a data sharing arrangement could be seen as a mean to learn synergies. And indeed, at the time Fitbit moved its data assets to Google's API, this strategic move was seen by specialists as a first step in the direction of the acquisition of Fitbit by Google.²⁶

Our results suggest that classical merger control that ignores data-sharing practices among firms can reduce consumer surplus. Lenient merger control may be good for both firms and consumers because it can give firms incentives to share data, which will intensify competition in case firms do not merge. This is a novel argument that challenges the traditional view in favor of strict merger control. Finally, we also argue that an efficient regulation should combine

²⁴Google is slowly killing Fitbit – should you still buy one? Tom's guide, October 3, 2024. The slow, painful death of Fitbit continues, with Google Assistant the latest feature to be axed. *techradar*, April 17, 2025.

²⁵Fitbit Rides on Google's Cloud, May 1, 2018.

²⁶Why Fitbit could make a good Google acquisition, ZDNET, May 3, 2018.

merger control with mandated data-sharing rules.

Mandated data-sharing regulations have already been implemented to promote the competitiveness of digital markets, as discussed in Section 3.4.1, and as merger remedies, in the Thomson/Reuters case for instance.²⁷ By showcasing how firms can use data exchanges to anticipate the value of synergies, our analysis suggests that recent mandated data sharing rules can prompt firms to better anticipate the synergistic value of data in case of a merger. Therefore, these new types of data regulation can also improve merger efficiency to benefit firms and consumers, establishing a connection between regulation aiming at promoting competition, and their impacts on merger decision and merger efficiency.

5 The Google/Fitbit Merger

We now apply our theoretical framework to Google’s 2021 acquisition of Fitbit, tracing the strategic decisions from pre-merger data migration to post-merger integration.

Market Context and Strategic Positions

Before the merger, the hardware segment of the health wearable market notably included Apple, Xiaomi, Samsung, Huawei, and Fitbit. Despite pioneering the market in 2009, Fitbit’s market share had eroded to approximately 5% of sales volume by the time of acquisition.²⁸

Google occupied a distinctive position: absent from wearable hardware but well-established in adjacent markets through its Google Fit health-tracking platform and Wear OS operating system (used by Fossil, LG, and others).²⁹ Crucially, Google had been developing its Pixel Watch prototype since before 2016, eventually launched in 2022, one year after completing the Fitbit acquisition.³⁰ This made Google both a potential acquirer and a credible competitive threat.

²⁷Case No COMP/M.4726 - Thomson Corporation/Reuters Group.

²⁸How Fitbit Dominated the Wearables Market and Changed Fitness Tracking. Medium, September 30, 2023. Shipment of wearable devices reach 118.9 million units in the fourth quarter and 336.5 millions for 2019. Business wire, March 10, 2020.

²⁹Introducing the new Google Fit. Google Fit, August 21, 2018.

³⁰Google plans two Android Wear smartwatches with Google Assistant. ZDNet, July 7, 2016.

Pre-Merger Data Partnership

In 2018, Fitbit began using Google Cloud Healthcare API to store a subset of its data, while continuing to manage the bulk independently.³¹ The stated purpose was to facilitate integration between wearable data and electronic health records. However, as Google Cloud’s press release noted, the partnership also enabled Google to “explor[e] how Google Cloud’s machine learning APIs can help them uncover deeper insights”; precisely the synergy-learning mechanism our model describes.

Figure 4 illustrates this partial migration structure, where Fitbit shares application data with Google Cloud, enabling Google to learn synergies through experimentation while compensating Fitbit via monetary transfer.

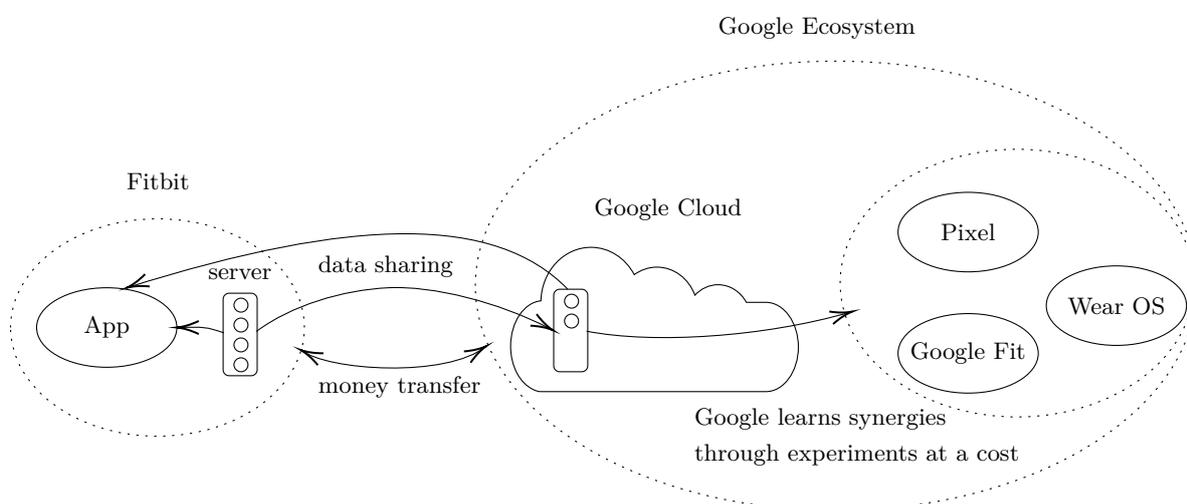


Figure 4: Partial migration of data and learning of synergies.

The Logic of Partial Migration

From a technical standpoint, Fitbit’s engineers faced a choice in migrating their monolithic MySQL-based application to Google Cloud. An all-at-once migration offered efficiency gains but risked service outages if compatibility issues arose. Instead, Fitbit opted for progressive shard-by-shard migration, which provided a backup system and—critically—the ability to reverse the process if needed.³²

This operational choice aligns directly with our theoretical framework. Without any integration, Google could not assess the true complementarity between the firms’ information systems

³¹New collaboration with Fitbit to drive positive health outcomes. Google Cloud, May 1, 2018.

³²As explained by Fitbit’s Principal Software Engineer: [How Fitbit moved its monolithic application to Google Cloud Platform](#). Diginomica, August 12, 2020.

and data assets. Full integration would resolve this uncertainty but would expose Fitbit to expropriation risk: Google could use the shared data to enhance its Pixel products and compete directly (the Arrow effect).

By sharing only a fraction of its data, Fitbit enabled Google to learn synergy values while limiting competitive exposure or operational inefficiencies. The progressive, reversible nature of the migration provided Fitbit with a credible exit option. Our analysis shows this approach was profitable precisely because it resolved uncertainty about merger efficiency, even accounting for the competitive losses Google could inflict by learning from the shared data. Specialists recognized this strategic dimension at the time: Fitbit’s move to Google’s API was widely interpreted as a precursor to acquisition.

Overall, the analysis of the Google/Fitbit acquisition illustrates how partial data sharing can serve as a mechanism for learning merger synergies while reshaping competitive incentives. We now situate these findings within the broader literature on data sharing, data-driven mergers, and competition regulation.

6 Contributions to the Literature

This paper contributes to several strands of the literature by identifying pre-merger data sharing as an endogenous response to uncertainty about complementarities between firms’ products.

A large literature studies data sharing among rival firms, emphasizing its effects on competition and surplus. Data sharing may intensify competition and reduce industry profits, as in [Arrow \(1972\)](#), [d’Aspremont, Bhattacharya and Gérard-Varet \(2000\)](#), and [Anton and Yao \(2002\)](#), or soften competition through improved targeting or coordination (e.g., [Chen, Narasimhan and Zhang, 2001](#); [Liu and Serfes, 2006](#); [Choe et al., 2024](#)). In particular [Bhargava et al. \(2025\)](#) show that data sharing can also be profitable for firms, but this profitability results from a mechanism of co-opetition enabled by cross-market data linkage. We contribute to this literature by showing that data sharing can be profitable even when it reduces competitive profits, because it generates an option value through improved merger decisions.

Our analysis also relates to the literature on data-driven mergers and acquisitions. Existing models typically assume that mergers are ex-ante profitable and focus on their effects on prices, quality, or consumer surplus (e.g., [Chen et al., 2022](#); [Bounie, Dubus and Waelbroeck, 2023](#); [de Cornière and Taylor, 2024](#).) By contrast, we introduce uncertainty over merger outcomes and

study how data sharing can resolve this uncertainty before merging. This approach highlights a previously unexplored interaction between data sharing, merger incentives, and welfare.

Finally, we contribute to the policy debate on regulating data in digital markets. Prior work argues that data-sharing rules can promote competition and contestability (Crémer et al., 2019; Tirole, 2023). We show that such rules also affect firms' incentives to learn about synergies and therefore influence both merger frequency and merger efficiency. Moreover, we analyze post-merger remedies such as data silos and show that they feed back into firms' pre-merger data-sharing strategies. Taken together, these results suggest that merger control and data regulation cannot be evaluated in isolation.

7 Conclusion

This paper studies mergers in environments where the value of complementarities is uncertain ex ante. We show that pre-merger data sharing can dominate mergers under incomplete information by allowing firms to condition acquisition decisions on realized synergies. This mechanism operates even when data sharing intensifies competition and reduces competitive profits.

From a firm perspective, this implies that due diligence in data-driven mergers should place greater emphasis on assessing data and algorithmic complementarities. Partial data sharing provides a mechanism to learn about these factors while limiting exposure to competitive risks.

From a policy perspective, the analysis highlights a tension between merger control and data-sharing incentives. Merger control focusing solely on post-merger outcomes may discourage data sharing and thereby reduce consumer surplus by weakening competition or inducing mergers under uncertainty. In contrast, policies that combine merger control with data-sharing mandates can improve welfare by preserving firms' incentives to learn about synergies before merging.

The Google/Fitbit case illustrates these mechanisms in practice, including the role of partial data migration prior to merger and the effects of post-merger data silo remedies on firms' incentives. These patterns appear in other data-driven acquisitions preceded by data-sharing partnerships.

Several extensions are left for future work. Allowing for divestment after inefficient mergers, introducing uncertainty about regulatory decisions, or modeling data as time-limited would further enrich the analysis. More broadly, improved learning about synergies is likely to affect merger rates, market concentration, and merger failure. Identifying these effects empirically

remains an important avenue for future research.

References

- Anton, J. J. and Yao, D. A. (2002), ‘The sale of ideas: Strategic disclosure, property rights, and contracting’, *The Review of Economic Studies* **69**(3), 513–531.
- Arrow, K. J. (1972), Economic welfare and the allocation of resources for invention, *in* ‘The Rate and Direction of Inventive Activity’, Princeton University Press, pp. 609–626.
- Bergemann, D., Bonatti, A. and Gan, T. (2022), ‘The economics of social data’, *The RAND Journal of Economics* **53**(2), 263–296.
- Bhargava, H., Dubus, A., Ronayne, D. and Shekhar, S. (2025), ‘The strategic value of data sharing in interdependent markets’, *Management Science* .
- Bounie, D., Dubus, A. and Waelbroeck, P. (2023), ‘Competition between strategic data intermediaries with implications for merger policy’.
- Chen, Y., Narasimhan, C. and Zhang, Z. J. (2001), ‘Individual marketing with imperfect targetability’, *Marketing Science* **20**(1), 23–41.
- Chen, Z., Choe, C., Cong, J. and Matsushima, N. (2022), ‘Data-driven mergers and personalization’, *The RAND Journal of Economics* **53**(1), 3–31.
- Cho, I.-K. and Kreps, D. M. (1987), ‘Signaling games and stable equilibria’, *The Quarterly Journal of Economics* **102**(2), 179–221.
- Choe, C., Cong, J. and Wang, C. (2024), ‘Softening competition through unilateral sharing of customer data’, *Management Science* **70**(1), 526–543.
- Crémer, J., de Montjoye, Y.-A. and Schweitzer, H. (2019), ‘Competition policy for the digital era’, *Report for the European Commission* .
- d’Aspremont, C., Bhattacharya, S. and Gérard-Varet, L.-A. (2000), ‘Bargaining and sharing innovative knowledge’, *The Review of Economic Studies* **67**(2), 255–271.
- de Cornière, A. and Taylor, G. (2024), ‘Data-driven mergers’, *Management Science* **70**(9), 6473–6482.
- Gautier, A. and Lamesch, J. (2021), ‘Mergers in the digital economy’, *Information Economics and Policy* **54**, 100890.
- Gautier, A. and Maitry, R. (2024), ‘Big tech acquisitions and product discontinuation’, *Journal of Competition Law & Economics* p. nhae010.
- Hagiu, A. and Wright, J. (2023), ‘Data-enabled learning, network effects, and competitive advantage’, *The RAND Journal of Economics* **54**(4), 638–667.
- Harsanyi, J. C. and Selten, R. (1972), ‘A generalized nash solution for two-person bargaining games with incomplete information’, *Management science* **18**(5-part-2), 80–106.
- Hocuk, S., Martens, B., Prufer, P., Carballa Smichowski, B., Duch-Brown, N., Kumar, P. and Mulder, J. (2022), ‘Economies of scope in data aggregation: Evidence from health data’.
- Jullien, B. and Sand-Zantman, W. (2021), ‘The economics of platforms: A theory guide for competition policy’, *Information Economics and Policy* **54**, 100880.

- Katz, M. L. and Shapiro, C. (1985), ‘Network externalities, competition, and compatibility’, *The American economic review* **75**(3), 424–440.
- Krämer, J. and Shekhar, S. (2025), ‘Regulating digital platform ecosystems through data sharing and data siloing: Consequences for innovation and welfare’, *Management Information Systems Quarterly* **49**(1), 123–154.
- Liu, Q. and Serfes, K. (2006), ‘Customer information sharing among rival firms’, *European Economic Review* **50**(6), 1571–1600.
- Myerson, R. B. (1979), ‘Incentive compatibility and the bargaining problem’, *Econometrica: journal of the Econometric Society* pp. 61–73.
- Neiswanger, W. and Xing, E. (2017), Post-inference prior swapping, *in* ‘International Conference on Machine Learning’, PMLR, pp. 2594–2602.
- Samuelson, W. (1984), ‘Bargaining under asymmetric information’, *Econometrica: Journal of the Econometric Society* pp. 995–1005.
- Scaria, E., Berghmans, A., Pont, M., Arnaut, C. and Leconte, S. (2018), *Study on data sharing between companies in Europe – Final report*, Publications Office.
- Schaefer, M. and Sapi, G. (2023), ‘Complementarities in learning from data: Insights from general search’, *Information Economics and Policy* **65**, 101063.
- Shapiro, C. and Varian, H. (1998), *Information rules: A strategic guide to the network economy*, Harvard Business Press.
- Tirole, J. (2023), ‘Competition and the industrial challenge for the digital age’, *Annual Review of Economics* **15**, 573–605.
- Tsoy, N. and Konstantinov, N. (2023), ‘Strategic data sharing between competitors’, *Advances in Neural Information Processing Systems* **36**, 16483–16514.
- Varian, H. R. (2014), ‘Big data: New tricks for econometrics’, *Journal of economic perspectives* **28**(2), 3–28.
- Wu, X. and Yu, H. (2022), ‘Mars-fl: Enabling competitors to collaborate in federated learning’, *IEEE Transactions on Big Data* **10**(6), 801–811.

A Micro-founded Applications

We present two micro-founded models to demonstrate the applicability of our reduced-form methodology to classical modes of competition. In the first framework, we examine price competition in the manner of Bertrand, whereas in the second framework, we investigate quantity competition following the Cournot model. Within both models, we establish the validity of the assumptions underlying our primary analysis. We also delineate the optimal regulatory measures in the context of the price competition model.

In these micro-founded models we focus on a market structure where Fitbit is initially a monopolist in the market, and Google can enter only if some of Fitbit’s data has been shared. Although this polar market structure plays against merger acceptance – as a merger always results in a monopoly – we show that even in this most conservative setting, there is room for merger control design under which firms have incentives to share data and some mergers are accepted. Both applications can be extended to account for an arbitrary number of firms.

A.1 A Price Competition Model

Our first micro-foundation is Bertrand competition among firms offering products of different qualities. Before the data is shared, Fitbit sells a product of quality q_F^0 at a price p_F^0 , and Google sells a product of quality q_G^0 at a price p_G^0 . There is a mass 1 of consumers with quasi-linear utility functions $u(q_i, p_i) = q_i - p_i$. Firms compete à la Bertrand to sell their products, and the firm that wins the market makes profits equal to $\pi_i = p_i$.

At the beginning of the game, only Fitbit can provide positive quality, that is $q_F^0 > 0$, $q_G^0 = 0$. This is consistent with an industry where a monopolist faces a platform that is initially not active in this market. The important element here is that Fitbit initially provides a product of higher quality than Google. Equilibrium prices and profits are $\pi_F^0 = p_F^0 = q_F^0$ and $\pi_G^0 = p_G^0 = 0$.

Data sharing. Assume Fitbit shares $s > 0$, and ignore for now the possibility of a M&A. The exploration cost is $C(s) = \frac{(1-s)^2}{s}$. In this application, we focus on a functional form following a normalized sum of squared deviations to the total amount of data shared corresponding to $s = 1$. The greater the distance to the full set of data, the more costly it is to infer its synergistic value.³³ Decomposing further this functional form, we see that this specific cost captures the reduction in inference costs from obtaining more data through the terms $1 - s$ and $\frac{1}{s}$. As mentioned in the baseline model, these terms capture potential economies of scope and of scale in the inference task. The fact that inferring synergies from a small share of data is very costly is captured by the denominator, as $\lim_{s \rightarrow 0} \frac{1}{s} = +\infty$. Conversely, the exploration cost is equal to zero when all data is shared as $1 - s = 0$ when $s = 1$. Finally, $1 - s$ is squared to ensure that the derivative of $C(s)$ is equal to zero at $s = 1$. These different elements combined in $C(s)$ ensure that the cost function satisfies the Inada conditions of our baseline model.

After learning θ , Google can develop a product with quality $q_G(s, \theta) = s\theta$, while Fitbit’s

³³Similar forms of cost function are typical in the econometrics literature where Pearson’s chi-squared test is defined as the sum of normalized distances of each observation from the corresponding element of a distribution.

product quality remains the same.³⁴ Representing the value enhancements resulting from data access using such a quality term that directly increases with the amount of data is typical in the literature on data sharing, especially in platform ecosystems. For instance, [Bhargava et al. \(2025\)](#) and [Krämer and Shekhar \(2025\)](#) use a similar form of increase in product quality when a firm has more data. We augment this approach by allowing the quality term to depend on the value of synergies through parameter θ .

The assumptions of the baseline model remain valid within this framework under the condition $\bar{\theta} > u > 2\underline{\theta}$. By positing $u > \underline{\theta}$, there is a loss of value from low synergies. This aligns with our narrative where consumers might lose utility from privacy concerns when their data is transferred to Google. For example, we can define $\theta := \nu - \mathcal{P}$, where $\nu \in \{\bar{\nu}, \underline{\nu}\}$ is the gain in product quality from sharing data with $\bar{\nu} > \underline{\nu} > u$; and \mathcal{P} represents the privacy losses encountered by consumers, which increase when more data is shared, yielding $\bar{\theta} = \bar{\nu} - \mathcal{P}$ and $\underline{\theta} = \underline{\nu} - \mathcal{P}$.³⁵ Under this condition, the equilibrium will always be separating.

Consumers purchase the product that yields the highest utility, and Bertrand competition implies that the equilibrium profits of the firms are

$$\begin{aligned}\pi_F^C(s, \theta) &= 0, & \pi_G^C(s, \theta) &= s\theta - u \text{ if } s\theta - u \geq 0, \\ \pi_F^C(s, \theta) &= u - s\theta, & \pi_G^C(s, \theta) &= 0 \text{ if } s\theta - u \leq 0.\end{aligned}$$

The assumptions of the nonparametric model are satisfied with this quality competition model.

$$\begin{aligned}\text{(H2)} \quad \frac{\partial \pi_G^C(s, \theta)}{\partial s} &= \theta \geq 0 \geq \frac{\partial \pi_F^C(s, \theta)}{\partial s} = 0 \text{ if } s\theta - u \geq 0, \\ \frac{\partial \pi_G^C(s, \theta)}{\partial s} &= 0 \geq 0 \geq \frac{\partial \pi_F^C(s, \theta)}{\partial s} = -\theta \text{ if } s\theta - u \leq 0.\end{aligned}$$

Assuming $u > \underline{\theta}$, low synergies make competitive payoffs equal to $\pi_F^C(s, \underline{\theta}) = u - s\underline{\theta}$ and $\pi_G^C(s, \underline{\theta}) = 0$, ensuring [H3](#) is satisfied:

$$\text{(H3)} \quad \frac{\partial \pi^C(s, \underline{\theta})}{\partial s} = -\underline{\theta} < 0.$$

³⁴We assume for simplicity that consumers immediately know the product's value. This aligns with our assumption that, when synergies are unknown to Fitbit, the firm can observe the quality of Google's product at the time of competition. In general, consumers can quickly discover their product value before purchasing it through sampling, free download, freemium, ratings, and reviews ([Shapiro and Varian, 1998](#)).

³⁵Note that this model is isomorphic to a setting where data improves the product quality of Google (and $\bar{\theta} > \underline{\theta} = u$), but merging induces a cost, for instance, administrative or from shock of culture, as we have already explained in the introduction.

The complementarity between data and synergies for Google also takes place in this setting:

$$(H4) \quad \frac{\partial \pi_G^C(s, \theta)}{\partial s \partial \theta} = 0 \text{ if } s\theta - u \leq 0, \quad \frac{\partial \pi_G^C(s, \theta)}{\partial s \partial \theta} = 1 \text{ if } s\theta - u \geq 0.$$

Finally, the competitive profits of Fitbit and Google with sharing and high synergies are:

$$\begin{aligned} \pi^C(s, \bar{\theta}) &= u - s\bar{\theta} \text{ if } s\bar{\theta} - u \leq 0, \\ \pi^C(s, \bar{\theta}) &= s\bar{\theta} - u \text{ if } s\bar{\theta} - u \geq 0. \end{aligned}$$

M&A. When firms merge, the merged entity sells as a monopolist a product of quality equal to $q_G(1, \bar{\theta}) = \bar{\theta}$ if synergies are high and to $q_G(1, \underline{\theta}) = \underline{\theta}$ if synergies are low.³⁶

Alternatively, we can assume that the merged firm sells the product with the highest quality between θ and u , that is, $\bar{\theta}$ if the synergies are high and u if synergies are low. Both approaches are equivalent by assuming that there is a positive cost to merge, even slightly above zero.

As we have assumed that $\bar{\theta} > u$ and $u > 2\underline{\theta}$, Google and Fitbit are willing to merge only when synergies are high and we recover the option value of the baseline model:

$$(H5) \quad \begin{aligned} \pi^A(\bar{\theta}) &= \bar{\theta} > \pi^C(s, \bar{\theta}) = \max\{u - s\bar{\theta}, s\bar{\theta} - u\}, \\ \pi^A(\underline{\theta}) &= \underline{\theta} < u - s\underline{\theta}. \end{aligned}$$

The price of the acquisition is equal to:

$$p^A(s) = \begin{cases} p^*(s) = \max\{u - \bar{\theta}s, 0\} + \mu(\bar{\theta} - \max\{u - \bar{\theta}s, \bar{\theta}s - u\}) & \text{if } p^*(s) \geq \underline{\theta} \\ \bar{\theta} - \max\{0, \bar{\theta}s - u\} & \text{if } p^*(s) \leq \underline{\theta}. \end{cases}$$

Equilibrium data sharing. The expected profit of Google purchasing s from Fitbit is:

$$(1 - \alpha)(\bar{\theta} - u) - \alpha\underline{\theta}s - C(s).$$

³⁶This assumes that the merged firm cannot sell a product of quality equal to u in case of low synergies. This effect captures situations where the identity of the firm selling the product is not neutral, which is in line with our narrative where consumers have privacy concerns. If Google poses a threat to their privacy, for instance, if Google represents a firm like Google or Facebook, consumers will have a lower valuation of the product after the acquisition. (See, for instance, 'Tossed my Fitbit in the trash': users fear for privacy after Google buys a company, *The Guardian*, November 6, 2019.)

By concavity of this expression, the equilibrium amount of data shared is unique and satisfies

$$\alpha\theta = -C'(s^*(\alpha)) = \frac{1 - s^*(\alpha)^2}{s^*(\alpha)^2} \implies s^*(\alpha) = \sqrt{\frac{1}{1 + \alpha\theta}}.$$

Consumer surplus. The surplus of consumers is equal to zero if firms do not share data or if they merge. If firms share data and compete, the consumer surplus equals $\min\{u, s\theta\}$. This weakly increases with s , and is linear on $[0, \frac{u}{\theta}]$ and $[\frac{u}{\theta}, 1]$.

Regulation. A merger with incomplete information does not change consumer surplus:

$$\mathbb{E}[G(0, \theta)] = 0$$

and a M&A is always accepted under incomplete information, since a regulator clears M&As only if they increase the surplus of consumers. This is a polar case where a merger does not induce any gain for consumers under incomplete information.

When firms share s data, a M&A in case of high synergies is detrimental to consumers:

$$G(s, \bar{\theta}) = -\max\{u, s\bar{\theta}\}.$$

This function is negative, and the merger is never approved after data has been shared. Hence, the regulator discourages firms from sharing data. Firms prefer to merge under incomplete information or share data and compete. In the last case, firms anticipate profits equal to:

$$(1 - \alpha)(\max\{u - s\bar{\theta}, s\bar{\theta} - u\}) + \alpha(u - s\theta) - C(s)$$

For sharing to be profitable, it is necessary that $s\bar{\theta} > u$. Moreover, for sharing to dominate competition without sharing, we must have:

$$(1 - \alpha)(s\bar{\theta} - 2u) - \alpha s\theta - C(s) > 0.$$

We can see that a sufficient condition for sharing and competing to be a profitable strategy is when the above condition is satisfied for $s = 1$. In this case, the inequality writes:

$$\bar{\theta} > \frac{\alpha}{1 - \alpha}\theta + 2u.$$

Under this sufficient condition, we see that the presence of the regulator does not necessarily discourage firms from sharing, since their profits in case synergies are high can be significantly greater than those without sharing. In contrast, when $\bar{\theta} < 2u$, $(1 - \alpha)(s\bar{\theta} - 2u) - \alpha s\bar{\theta} - C(s)$ will always be negative and firms will not share data under this sufficient condition.

Optimal stringency. Profits in case of sharing and competing are

$$\Pi_B(s, \alpha) = (1 - \alpha)(s\bar{\theta} - u) + \alpha(u - s\bar{\theta}) - C(s)$$

which are maximized for

$$(11) \quad s_B^*(\alpha) = \begin{cases} 1 & \text{if } \alpha\bar{\theta} \leq (1 - \alpha)\bar{\theta} \\ \sqrt{\frac{1}{1 - (1 - \alpha)\bar{\theta} + \alpha\bar{\theta}}} & \text{if } \alpha\bar{\theta} \geq (1 - \alpha)\bar{\theta} \end{cases}$$

For any α , there exists $\hat{\theta} (> 2u)$ such that firms share and compete if $\bar{\theta} > \hat{\theta}$ and do not share otherwise. Hence, the optimal degree of stringency depends on $\bar{\theta}$. If $\bar{\theta} > \hat{\theta}$, the regulator sets $r = 0$ and the surplus is high with firms sharing data and competing. If $\bar{\theta} \leq \hat{\theta}$, the firms do not share and the regulator can induce them to share by increasing r . Specifically, a M&A will be accepted if

$$G(s, \bar{\theta}) + r \geq 0 \implies r - \max\{u, s\bar{\theta}\} \geq 0$$

Consumer surplus increases when firms share more data, anticipating that they can merge if synergies are high. A regulator can promote such data sharing among firms by implementing a more lenient regulation captured by $r > 0$. Doing so, the regulator initially provides firms with incentives to share data \hat{s}^A , which increases with r . For $r \geq \max\{u, s^*(\alpha)\bar{\theta}\}$, \hat{s}^A is higher than $s^*(\alpha)$, leading firms to share the optimal amount of data $s^*(\alpha)$. Consequently, an optimal regulation has a degree of leniency $r \geq \max\{u, s^*(\alpha)\bar{\theta}\}$.

A.2 A Quantity Competition Model

Our second micro-foundation is Cournot competition. Firms offer products of quality q_F, q_G and compete à la Cournot by setting quantities x_F and x_G to sell their products, and, denoting

total output $X = x_F + x_G$, firms have the following inverse demand functions:³⁷

$$P_F = u - X, \quad P_G = q_G - X.$$

We normalize production costs to zero, and firms make profits:

$$\pi_F^0 = P_F x_F, \quad \pi_G^0 = P_G x_G.$$

Before data is shared, Fitbit's product is of quality u , and Google's product is of quality $q_G^0 = 0$.

Data sharing. Suppose that Fitbit shares $s > 0$, and, for now, let us ignore the possibility of a M&A. As in the Bertrand specification, if Google invests $C(s) = \frac{(1-s)^2}{s}$ and learns θ , it can develop a product of quality $q_G(s, \theta) = s\theta$, while the quality of the product of Fitbit does not change. If Google does not invest $C(s)$, its product quality is zero. As in the Bertrand competition case, this is a typical form of quality improvement in the literature which we augment using the uncertainty over the value of θ .

The main hypotheses of the baseline model hold when $\bar{\theta} > 2\sqrt{u^2 + u}$ and $u > \underline{\theta}$. In particular, the equilibrium will always be separating. Firms make the following competitive profits with data sharing:

$$\pi_F^C(s, \theta) = (u - X) x_F,$$

$$\pi_G^C(s, \theta) = (s\theta - X) x_G.$$

Standard analysis allows us to characterize the equilibrium outputs:

$$\begin{aligned} x_F &= \frac{u}{2}, & x_G &= 0 & \text{if } u > 2s\theta \\ x_F &= \frac{2u - s\theta}{3}, & x_G &= \frac{2s\theta - u}{3} & \text{if } 4u \geq 2s\theta \geq u \\ x_F &= 0, & x_G &= \frac{s\theta}{2} & \text{if } s\theta \geq 2u. \end{aligned}$$

In turn, the profits of the firms are:

$$\begin{aligned} \pi_F^{C^*}(s, \theta) &= \frac{u^2}{4}, & \pi_G^{C^*}(s, \theta) &= 0 & \text{if } u > 2s\theta, \\ \pi_F^{C^*}(s, \theta) &= \frac{1}{9}(2u - s\theta)^2, & \pi_G^{C^*}(s, \theta) &= \frac{1}{9}(2s\theta - u)^2 & \text{if } 4u \geq 2s\theta \geq u, \\ \pi_F^{C^*}(s, \theta) &= 0, & \pi_G^{C^*}(s, \theta) &= \frac{(s\theta)^2}{4} & \text{if } s\theta \geq 2u. \end{aligned}$$

³⁷See [Katz and Shapiro \(1985\)](#) for a micro-foundation of these inverse demand functions.

These Cournot profits satisfy the assumptions of our main model. The Arrow effect holds

$$(H2) \quad \frac{\partial \pi_G^C(s, \underline{\theta})}{\partial s} = \max\{0, 4s\underline{\theta}(2s\underline{\theta} - u)\} \geq 0 \geq \frac{\partial \pi_F^C(s, \underline{\theta})}{\partial s} = \min\{0, 2\underline{\theta}(s\underline{\theta} - 2u)\}.$$

The competitive payoffs are equal to:

$$\begin{aligned} \pi^C(s, \theta) &= \frac{u^2}{4} && \text{if } u > 2s\theta, \\ \pi^C(s, \theta) &= \frac{1}{9} (5s^2\theta^2 - 8s\theta u + 5u^2) && \text{if } 4u \geq 2s\theta \geq u, \\ \pi^C(s, \theta) &= \frac{(s\theta)^2}{4} && \text{if } s\theta \geq 2u. \end{aligned}$$

These profits decrease with s and H3 is satisfied:

$$(H3) \quad \begin{aligned} \frac{\partial \pi^C(s, \underline{\theta})}{\partial s} &= 0 && \text{if } u > 2\underline{\theta}s, \\ \frac{\partial \pi^C(s, \underline{\theta})}{\partial s} &= \frac{2}{9} \underline{\theta} (5s\underline{\theta} - 4u) \leq 0 && \text{if } 2\underline{\theta}s \geq u. \end{aligned}$$

The data-synergies complementarity for Google's profits also takes place:

$$(H4) \quad \frac{\partial^2 \pi_G^C(s, \theta)}{\partial s \partial \theta} = \begin{cases} 0 & \text{if } u > 2s\theta, \\ \frac{4}{9}(4s\theta - u) & \text{if } 4u \geq 2s\theta \geq u, \\ s\theta & \text{if } s\theta \geq 2u. \end{cases}$$

Finally, taking the second-degree derivative of the sum of Fitbit and Google's payoffs with high synergies, we can see that these profits are quasi-concave in s :

$$\begin{aligned} \frac{\partial^2 \pi^C(s, \bar{\theta})}{\partial s^2} &= 0 && \text{if } u > 2s\bar{\theta}, \\ \frac{\partial^2 \pi^C(s, \bar{\theta})}{\partial s^2} &= \frac{10}{9} \bar{\theta} > 0 && \text{if } 4u \geq 2s\bar{\theta} \geq u, \\ \frac{\partial^2 \pi^C(s, \bar{\theta})}{\partial s^2} &= 0 && \text{if } s\bar{\theta} \geq 2u. \end{aligned}$$

M&A. When firms merge, the merged entity sells as a monopolist a quality product equal to $q_G(1, \bar{\theta}) = \bar{\theta}$ if synergies are high and to $q_G(1, \underline{\theta}) = \underline{\theta}$ if synergies are low. As in the price competition model, we could equivalently assume that the merged firm sells a quality product u in the case of low synergies, but there is a cost to merge.

The profits of the merged company are $\pi^A(\theta) = \max_x (\theta - x)x$. Hence, the optimal output

is $x^M(\theta) = \frac{\theta}{2}$, so that the profits after the merger are

$$\pi^A(\theta) = \frac{1}{4}\theta^2.$$

To show that a merger is not profitable with low synergies, it is necessary and sufficient to prove that a merger yields lower profits than competition with the minimizing value of s . As $\pi^C(s, \underline{\theta})$ decreases with s , it is the lowest for $s = 1$. With this amount of data shared, a merger is detrimental with low synergies when $\frac{10u}{11} > \underline{\theta}$, which is implied by our assumption that $u > 2\underline{\theta}$.

In contrast, if a merger with high synergies is profitable for the value of s that yields the highest competitive profits $\pi^C(s, \bar{\theta})$, then mergers are always profitable. As $\pi^C(s, \bar{\theta})$ is convex, it is maximized for $s = 0$ or for $s = 1$. In both cases, simple analysis allows us to verify that a merger is profitable as we have assumed $\bar{\theta} > u$.³⁸ Overall, Google and Fitbit are willing to merge only when synergies are high and we recover the option value of the baseline model:

$$(H5) \quad \begin{aligned} \pi^A(\bar{\theta}) &= \frac{1}{4}\bar{\theta}^2 > \pi^C(s, \bar{\theta}), \\ \pi^A(\underline{\theta}) &= \frac{1}{4}\underline{\theta}^2 < \pi^C(s, \underline{\theta}). \end{aligned}$$

Equilibrium data sharing. The expected profit of Google purchasing s from Fitbit is:

$$(1 - \alpha)\frac{1}{4}(\bar{\theta}^2 - u^2) + \alpha\frac{1}{9}\left(5s^2\underline{\theta}^2 - 8s\underline{\theta}u + 5u^2\right) - C(s).$$

The second-degree derivative of this profits with respect to s is equal to $10\alpha\underline{\theta}^2 - \frac{2}{s^3}$, and is thus negative for all values of s and α as long as $\underline{\theta} > \frac{1}{\sqrt{5}}$. In this case, the equilibrium amount of data shared is unique and satisfies

$$\frac{2}{9}\alpha\underline{\theta}(5s^*(\alpha)\underline{\theta} - 4u) = C'(s^*(\alpha)).$$

Consumer surplus. As usual with Cournot, total output is a sufficient statistic for consumer surplus, and we have $CS = \frac{X^2}{2}$.³⁹ Baseline competition yields a surplus equal to $CS^0 = \frac{u^2}{8}$. After a merger when synergies are high, consumer surplus is equal to $CS^A(\bar{\theta}) = \frac{\bar{\theta}^2}{8}$, while it is equal to $CS^A(\underline{\theta}) = \frac{\underline{\theta}^2}{8}$ in case of low synergies. If firms share data and compete, the total production equals $\frac{1}{3}(s\theta + u)$, which weakly increases with s , and in turn, so does the surplus of

³⁸At $s = 0$, the condition for a profitable merger writes $\bar{\theta} > \frac{2\sqrt{5}}{3}u$, while at $s = 1$. Both conditions are $\bar{\theta} > \frac{10u}{11}$ which is implied by $\bar{\theta} > u$.

³⁹Proof available upon request.

the consumer. This ensures that [H6](#) is satisfied. Assumption [H7](#) holds as $\bar{\theta} > \frac{2u}{3} > \underline{\theta}$.

$$(H7) \quad CS^A(\bar{\theta}) = \frac{\bar{\theta}^2}{8} > CS^0 = \frac{u^2}{8} > CS^A(\underline{\theta}) = \frac{\underline{\theta}^2}{8}.$$

Finally, [H8](#) is also satisfied:

$$(H8) \quad \begin{aligned} CS^C(1, \bar{\theta}) &= \frac{1}{18}(\bar{\theta} + u)^2 > CS^A(\bar{\theta}) = \frac{\bar{\theta}^2}{8} && \text{if } 2u > \bar{\theta}, \\ CS^C(1, \bar{\theta}) &= CS^A(\bar{\theta}) = \frac{\bar{\theta}^2}{8} && \text{if } 2u \leq \bar{\theta}. \end{aligned}$$

To keep the discussion concise we leave the analysis of the equilibrium under regulation available upon request, but it is clear that, as all the baseline assumptions are satisfied, the general results of the main model hold also in this micro-founded application.

B Proof of Proposition 1

Separating equilibrium. Suppose that $\theta = \bar{\theta}$. In equilibrium, Fitbit and Google “share” the belief that $\theta = \bar{\theta}$, in which they respectively get p^A and $\pi^A(\bar{\theta}) - p^A$. The frontier of the set of allocations to respectively Fitbit and Google (u_F, u_G) that dominate the threat point $(\pi_F^C(s, \bar{\theta}), \pi_G^C(s, \bar{\theta}))$ is obtained by varying p^A in the interval

$$D := [\pi_F^C(s, \bar{\theta}), \pi^A(\bar{\theta}) - \pi_G^C(s, \bar{\theta})],$$

with D nonempty by assumption $\pi^A(\bar{\theta}) > \pi^C(s, \bar{\theta})$. However, Google may be tempted to offer acquisition at price $p^A \in D$ even in state $\underline{\theta}$. This will not happen whenever Google’s payoff from acquisition $\pi^A(\underline{\theta}) - p^A$ is smaller than its equilibrium payoff $\pi_G^C(s, \underline{\theta})$ in state $\underline{\theta}$. Hence, incentive compatibility requires that $p^A \geq \pi^A(\underline{\theta}) - \pi_G^C(s, \underline{\theta})$.

Anticipating the incentive compatibility problem, at the time of acquisition, Fitbit and Google bargain over the constrained set

$$(12) \quad \begin{aligned} D^* &:= D \cap [\pi^A(\underline{\theta}) - \pi_G^C(s, \underline{\theta}), \infty] \\ &= \left[\max\{\pi_F^C(s, \bar{\theta}), \pi^A(\underline{\theta}) - \pi_G^C(s, \underline{\theta})\}, \pi^A(\bar{\theta}) - \pi_G^C(s, \bar{\theta}) \right]. \end{aligned}$$

Because $\pi^A(\bar{\theta}) > \pi^C(s, \bar{\theta})$, a necessary and sufficient condition for a separating equilibrium is

then that

$$(13) \quad \pi^A(\bar{\theta}) - \pi^A(\underline{\theta}) \geq \pi_G^C(s, \bar{\theta}) - \pi_G^C(s, \underline{\theta}).$$

Equivalently, when synergies increase, the profits under acquisition must increase by more than the profit to Google under competition. The generalized Nash bargaining outcome is obtained as the solution to

$$\max_{p^A \in D^*} (p^A - \pi_F^C(s, \bar{\theta}))^\mu (\pi^A(\bar{\theta}) - p^A - \pi_G^C(s, \bar{\theta}))^{1-\mu}$$

Note that $D^* = D$ whenever $\pi^A(\underline{\theta}) - \pi_G^C(s, \underline{\theta}) \leq \pi_F^C(s, \bar{\theta})$. In this case, the solution is the same as without the incentive problem, that is

$$p^*(s) = \pi_F^C(s, \bar{\theta}) + \mu(\pi^A(\bar{\theta}) - \pi_G^C(s, \bar{\theta})).$$

More generally, if $p^*(s) \in D^*$, the solution to the constrained problem is $p^A(s) = p^*(s)$. Otherwise the solution is at the lower bound of D^* , that is at $p^A(s) = \pi^A(\bar{\theta}) - \pi_G^C(s, \bar{\theta})$.

Pooling-to-merge equilibrium. If pooling to merge is an equilibrium, ex-ante, Google has a payoff of $\mathbb{E}[\pi^A(\theta)] - p^A(s) - C(s)$ and Fitbit of $p^A(s)$. This is dominated by early acquisition at price p^A since Google has a payoff of $\mathbb{E}[\pi^A(\theta)] - p^A$ and saves $C(s)$.

Pooling-to-compete equilibrium. Suppose that Google is not expected to offer acquisition. At this stage $C(s)$ has been sunk by Google and Fitbit has imperfect information about θ and expects a profit of $\mathbb{E}[\pi_F^C(s, \theta)]$. Suppose first that the set D^* in (12) is non-empty. If Google deviates in state $\bar{\theta}$ and offers a price p^A such that $p^A \in D^*$: then, Google is better off than in the pooling outcome while type $\underline{\theta}$ would be worse off offering acquisition. Using the Never a Weak Best Response (NWBR) refinement (e.g., [Cho and Kreps \(1987\)](#)), Fitbit should put a probability one on the state being $\bar{\theta}$ when Google offers to acquire Fitbit at price $p^A \in D^*$. Hence pooling-to-compete is not stable when D^* is non-empty.

Suppose now that D^* is empty. In order for the pooling-to-compete strategy to be stable, neither $\bar{\theta}$ or $\underline{\theta}$ can benefit from offering acquisition, or acquisition is refused by Fitbit. Since D^* is empty, any beneficial deviation by offering a price p^A by $\bar{\theta}$ is also beneficial by $\underline{\theta}$. The NWBR criterion does not put restrictions on Fitbit's beliefs in this out-of-equilibrium strategy. Let q be the belief that Fitbit assigns to $\underline{\theta}$ out-of-equilibrium that is beneficial to both $\underline{\theta}, \bar{\theta}$; stability

requires that whenever acquisition is profitable for Google, that is when $p^A > \pi^A(\underline{\theta}) - \pi_G^C(s, \underline{\theta})$, Fitbit prefers to reject the offer, that is that $p^A < \mathbb{E}[\pi_F^C(s, \theta)|q]$. Hence, we need⁴⁰

$$(14) \quad q\pi_F^C(s, \underline{\theta}) + (1 - q)\pi_F^C(s, \bar{\theta}) > \pi^A(\underline{\theta}) - \pi_G^C(s, \underline{\theta}).$$

If $\pi_F^C(s, \underline{\theta}) \leq \pi_F^C(s, \bar{\theta})$, (14) holds for any q by our assumption that $\pi^A(\underline{\theta}) - \pi^C(s, \underline{\theta}) < 0$. If $\pi_F^C(s, \underline{\theta}) \geq \pi_F^C(s, \bar{\theta})$, condition (14) holds whenever q is large enough.⁴¹

- (i) Suppose that condition (13) holds. Then, after data exploitation, the unique equilibrium is separation at acquisition price $p^A(s)$ in Proposition 1.
- (ii) Suppose that condition (13) does not hold. Then, a pooling-to-compete continuation equilibrium exists.

C The Data Shared Maximizes Google's Profits

We note $u_G(s, \alpha)$ the expected value of Google purchasing data s from Fitbit and exploring, net of the non-sharing payoff π_G^0 :

$$(15) \quad \begin{aligned} u_G(s, \alpha) &:= \overbrace{(1 - \alpha)(\pi^A(\bar{\theta}) - p^A(s))}^{\text{Google's profits if } \bar{\theta} \text{ and acquires Fitbit}} + \overbrace{\alpha \pi_G^C(s, \underline{\theta})}^{\text{Google's profits if } \underline{\theta}} - \pi_G^0 - C(s) \\ &\quad - \underbrace{(\pi_F^0 + \max\{\mu(\mathbb{E}[\pi^A; \alpha] - \pi^0), 0\} - (1 - \alpha)p^A(s) - \alpha \pi_F^C(s, \underline{\theta}))}_{\text{Price of data } T(s)} \\ &= \Pi(s, \alpha) - \pi^0 - \max\{\mu(\mathbb{E}[\pi^A; \alpha] - \pi^0), 0\} \end{aligned}$$

where $\Pi(s, \alpha)$ is the ex-ante total profits of Fitbit and Google if s is shared before the acquisition

$$(16) \quad \Pi(s, \alpha) := (1 - \alpha)\pi^A(\bar{\theta}) + \alpha \pi^C(s, \underline{\theta}) - C(s).$$

D Proof of Lemmas 2, 3, and 4

Firms merge under incomplete information instead of sharing. When α is close to 0, a M&A is accepted under incomplete information. On the contrary, firms are constrained on their sharing strategy: the optimal amount of data shared without the regulator when α equals zero

⁴⁰Since (13) is violated, the right hand side is larger than $\pi^A(\bar{\theta}) - \pi_G^C(s, \bar{\theta})$.

⁴¹Precisely, we need that $q > (\pi^A(\underline{\theta}) - \pi_G^C(s, \underline{\theta}) - \pi_F^C(s, \bar{\theta})) / (\pi_F^C(s, \underline{\theta}) - \pi_F^C(s, \bar{\theta}))$.

is $s^*(0) = 1$ (see (4)), but the regulator blocks such a M&A since $G(1, \bar{\theta}) = CS^A(\bar{\theta}) - CS^C(1, \bar{\theta})$ is below zero (see H8). Hence, for small values of α , the presence of the regulator lowers the firms' profits with sharing, but not those if they merge under incomplete information.

In the absence of regulation, the firms' profits for a M&A under incomplete information are arbitrarily close to those with data sharing, and firms prefer to merge under incomplete information instead of sharing when there is a regulator for small values of α . There, regulation discourages firms from sharing. Formally, there exists $\alpha_0 \in (0, 1]$ such that firms prefer to merge under incomplete information instead of sharing when there is a regulator for α below α_0 .

Firms cannot merge under incomplete information but merge after sharing. Consider now the decision of the regulator to allow a M&A under incomplete information. We have shown that M&As are cleared under incomplete information when α is below $\alpha^R = \frac{CS^A(\bar{\theta}) - CS^0}{CS^A(\bar{\theta}) - CS^A(\underline{\theta})}$. For CS^0 sufficiently close to $CS^A(\bar{\theta})$, this threshold is lower than α_1 , and mergers under incomplete information are blocked by the regulator when $\alpha \in [\alpha^R, \alpha_1)$, even though they are more profitable than competing for the firms.⁴²

In this case, firms can still induce the regulator to allow the M&A by sharing data. Indeed, if $s^*(\alpha^R)$ is smaller than \hat{s}^A , firms share $s^*(\alpha^R)$ data and the merger is approved when synergies are high. If $s^*(\alpha^R)$ is greater than \hat{s}^A , firms share data and merge with high synergies if

$$\Pi(\hat{s}^A, \alpha^R) = (1 - \alpha^R)\pi^A(\bar{\theta}) + \alpha^R \pi^C(\hat{s}^A, \underline{\theta}) - C(\hat{s}^A) \geq \pi^0.$$

In other words, when $\alpha \in [\alpha^R, \alpha_1)$ the firms can ensure that a M&A takes place when synergies are high by exploring the data as long as $\Pi(s_A^*(\alpha), \alpha) \geq \pi^0$. This implies that the surplus gains from a M&A will be high too.

Firms compete instead of sharing. When α is close to α_2 , the profits after sharing without regulation are close to π^0 . If a M&A with data sharing is blocked at α_2 , firms can induce the regulator to accept the merger by sharing $\hat{s}^A < s^*(\alpha_2)$, but doing so lowers their profits below π^0 . Hence, by exerting a constraint on the strategies of the firms, the regulator can also induce them not to share data and to compete instead. Formally, when $G(s^*(\alpha_2), \bar{\theta}) \leq 0$, there exists $\alpha_2^R < \alpha_2$ so that firms prefer to compete instead of sharing when there is a regulator for $\alpha \in [\alpha_2^R, \alpha_2)$.

⁴²In terms of primitives, we have $\alpha^R < \alpha_1$ when $\frac{CS^A(\bar{\theta}) - CS^0}{CS^A(\bar{\theta}) - CS^A(\underline{\theta})} < \frac{\pi^A(\bar{\theta}) - \pi^0}{\pi^A(\bar{\theta}) - \pi^A(\underline{\theta})}$.

Sharing and competing can be the firms' equilibrium strategy when it yields greater profits than (i) merging under incomplete information; (ii) sharing and have the M&A accepted; (iii) and than competing without sharing.

Sharing and competing vs. merging under incomplete information. Sharing and competing will yield higher profits than a merger under incomplete information if $\Pi_B(s_B^*(\alpha), \alpha)$ is greater than $\mathbb{E}[\pi^A; \alpha]$. Consider the difference between both profit functions:

$$\Pi_B(s, \alpha) - \mathbb{E}[\pi^A; \alpha] = (1 - \alpha) \underbrace{\left(\pi^C(s, \bar{\theta}) - \pi^A(\bar{\theta}) \right)}_{<0 \text{ H5}} + \alpha \underbrace{\left(\pi^C(s, \underline{\theta}) - \pi^A(\underline{\theta}) \right)}_{>0 \text{ H5}} - C(s)$$

For $s = 1$, the exploration cost equals zero, and this difference is positive for α greater than a threshold α^B . Hence, this difference will also be positive for $s_B^*(\alpha)$ data shared in this case. This implies that sharing and competing yields higher profits than merging under incomplete information for α high enough.

Sharing and competing vs. sharing and merging when synergies are high. $\Pi_B(s_B^*(\alpha), \alpha)$ will be greater than $\Pi(\hat{s}^A, \alpha)$ if \hat{s}^A is small enough. Indeed, for $\hat{s}^A = 1$ all mergers are accepted under high synergies as $s^*(\alpha) \leq \hat{s}^A$. On the contrary, sharing \hat{s}^A is not a profitable strategy when it is close to zero since the exploration cost will be high. Hence, for any value of α , there exists $\hat{s}^A(\alpha)$ such that $\Pi(\hat{s}^A(\alpha), \alpha) = \Pi_B(s_B^*(\alpha), \alpha)$ if $s_B^*(\alpha) > \hat{s}^A(\alpha)$. The existence of $\hat{s}^A(\alpha)$ is guaranteed since the value of \hat{s}^A is determined by the relation $CS^C(\hat{s}^A, \bar{\theta}) = CS^A(\bar{\theta})$, where $CS^A(\bar{\theta})$ can take any value as long as conditions H7 and H8 are satisfied.

Thus, when $\hat{s}^A \leq \hat{s}^A(\alpha)$, regulation exerts a strong constraint on the strategies of the firms, up to a point where sharing and having the M&A blocked is more profitable for the firms than sharing less data to have the M&A accepted. This strong regulatory constraint results from a sharp increase of consumer surplus with s when firms compete, such that firms must share few data, otherwise a M&A will be detrimental to consumers and will be blocked by the regulator.

Competing vs. sharing and competing. Consider $\alpha \in (\alpha^B, \alpha_2)$ and $\hat{s}^A < \hat{s}^A(\alpha)$ defined above, such that sharing and competing dominates sharing and merging, and mergers under incomplete information are blocked. In this case, sharing s and competing can also be

more profitable than competing without sharing if the following difference is positive.

$$\Pi_B(s, \alpha) = (1 - \alpha) \left(\pi^C(s, \bar{\theta}) - \pi^0 \right) + \alpha \underbrace{\left(\pi^C(s, \underline{\theta}) - \pi^0 \right)}_{<0 \text{ H3}} - C(s)$$

As profits with sharing are below π^0 when firms are constrained, sharing and competing will dominate competition for $\alpha < \alpha_2^B < \alpha_2$. This is illustrated by our quality competition rationale.

E Proof of Proposition 4

At stage 0, the expected surplus when $s^*(\alpha)$ data is shared without merger control is:

$$\mathbb{E}[CS^{no}] := (1 - \alpha)CS^A(\bar{\theta}) + \alpha CS^C(s^*(\alpha), \underline{\theta}).$$

When firms share $\hat{s}^A < s^*(\alpha)$ to have the merger accepted by the regulator, the expected surplus at the beginning of the game depends on the realization of θ . If synergies are high with prob. $1 - \alpha$, firms ask for a M&A, which is cleared by the regulator, and surplus equals $CS^A(\bar{\theta})$. If synergies are low with prob. α , firms compete with \hat{s}^A data shared and surplus equals $CS^C(\hat{s}^A, \underline{\theta})$. The expected surplus with sharing at the beginning of the game is:

$$\mathbb{E}[CS(\hat{s}^A)] := (1 - \alpha)CS^A(\bar{\theta}) + \alpha CS^C(\hat{s}^A, \underline{\theta}).$$

Hence, regulation reduces surplus if the following difference is negative:

$$\mathbb{E}[CS(\hat{s}^A)] - \mathbb{E}[CS^{no}] = \alpha(CS^C(\hat{s}^A, \underline{\theta}) - CS^C(s^*(\alpha), \underline{\theta})) < 0.$$

This term is negative since \hat{s}^A is smaller than $s^*(\alpha)$. Hence, the presence of the regulator reduces both consumer surplus and firms' profits.

When firms merge under incomplete information because of the regulator, the expected surplus is $\mathbb{E}[CS^A(\theta)]$, and the variations in surplus due to regulation are equal to:

$$\mathbb{E}[CS^A(\theta)] - \mathbb{E}[CS^{no}] = \alpha(CS^A(\underline{\theta}) - CS^C(s^*(\alpha), \underline{\theta})) < 0.$$

Consumer surplus unilaterally decreases because of the regulator. Consumers lose since firms do not compete with data when synergies are low. Moreover, firms also lose as they no longer

share data, even though this would be the unconstrained optimum for the industry.

When firms prefer to compete because of the regulator, consumer surplus is CS^0 . The variations of surplus due to the presence of the regulator in this case are equal to:

$$CS^0 - \mathbb{E}[CS^{no}] = CS^0 - (1 - \alpha)CS^A(\bar{\theta}) - \alpha CS^C(s^*(\alpha), \underline{\theta}) < 0.$$

In this last case again, the presence of the regulator reduces consumer surplus. Indeed, surplus if firms share data is either realized under a merger with high synergies, or under competition with data shared if synergies are low. In both cases, the resulting surplus is greater than in the baseline competitive mode without sharing.

The presence of the regulator increases surplus if firms share more data. When firms share $s_B^*(\alpha) > s^*(\alpha)$ and compete, the expected surplus with sharing at stage 0 is:

$$\mathbb{E}[CS(s_B^*(\alpha))] := (1 - \alpha)CS^C(s_B^*(\alpha), \bar{\theta}) + \alpha CS^C(s_B^*(\alpha), \underline{\theta}).$$

Hence, regulation increases surplus if the following difference is positive:

$$\begin{aligned} \mathbb{E}[CS(s_B^*(\alpha))] - \mathbb{E}[CS^{no}] &= (1 - \alpha)(CS^C(s_B^*(\alpha), \bar{\theta}) - CS^A(\bar{\theta})) \\ &\quad + \alpha(CS^C(s_B^*(\alpha), \underline{\theta}) - CS^C(s^*(\alpha), \underline{\theta})) > 0. \end{aligned}$$

When firms share more data under regulation, $s_B^*(\alpha)$ is greater than $s^*(\alpha)$ and the difference $CS^C(s_B^*(\alpha), \underline{\theta}) - CS^C(s^*(\alpha), \underline{\theta})$ is positive. Moreover, $CS^C(s_B^*(\alpha), \bar{\theta})$ is greater than $CS^A(\bar{\theta})$ since the regulator blocks the merger. Overall, in this case, the presence of the regulator increases consumer surplus when it blocks the merger.

When firms share less data in the presence of the regulator, $s_B^*(\alpha)$ is smaller than $s^*(\alpha)$ and the difference $CS^C(s_B^*(\alpha), \underline{\theta}) - CS^C(s^*(\alpha), \underline{\theta})$ is negative. On the other hand, $CS^C(s_B^*(\alpha), \bar{\theta})$ is still greater than $CS^A(\bar{\theta})$. The resulting impact of regulation on consumer surplus is ambiguous.

We can anticipate whether firms share more or less data when they anticipate that the merger will be blocked by the regulator by considering the general class of joint profit functions that are quasi-convex in s , as in the two micro-foundations analyzed in the Appendix. This class of profit functions also includes any monotonous profits, increasing or decreasing in s .

We show that $s_B > s^*(\alpha)$, such that firms share more data if the M&A is blocked. A necessary condition for sharing and competing to be an equilibrium strategy is that $\pi^C(s, \bar{\theta}) > \pi^0$

for at least some values of s . Otherwise, $\Pi_B(s, \alpha) < \pi^0$.

This implies that $\pi^C(s_B^*(\alpha), \bar{\theta})$ increases as we have assumed it is quasi-convex. This means that sharing and competing can be the best strategy for the firm when $\alpha \in (\alpha^B, \alpha_2)$ and $\hat{s}^A < \hat{s}^A(\alpha)$ and the profits with sharing and high synergies are high enough. In this case, firms expect an increase in profits from the synergies when synergies are high.

We can show that $s_B^*(\alpha) > s^*(\alpha)$ using simple first-order conditions on Π_B : either Π_B is increasing at $s^*(\alpha)$, such that $s_B^*(\alpha)$ is greater than $s^*(\alpha)$; or Π_B is decreasing at $s^*(\alpha)$, which also implies that $s_B^*(\alpha)$ is greater than $s^*(\alpha)$ since it means that $\pi^C(s^*(\alpha), \bar{\theta})$ has a downward slope, and we have shown that $s_B^*(\alpha)$ is necessarily on the upward slope of $\pi^C(s, \bar{\theta})$. Overall, firms share more data since doing so increases the competitive profits when synergies are high and the merger is blocked.