Collect, Keep, or Share?
Platform Information Management*

Salvatore Piccolo†

April 27, 2026

Abstract

This paper studies how platform business models shape incentives to acquire and disclose marketplace information. It compares a pure intermediation platform, which merely brokers trade between third-party sellers, with a hybrid platform that also competes downstream through a private label. In both settings, the platform invests in demand information, chooses how much of that information to disclose, and provides a marketplace-wide service effort that improves ecosystem quality. The analysis shows that the two business models generate sharply different information policies. A pure intermediary fully discloses to third-party sellers any information it acquires, because exclusive information has no private strategic value. By contrast, a hybrid platform prefers to keep information private and, therefore, treats disclosure mandates as implicit constraints on data acquisition. With a covered-market, consumers prefer information not to be shared irrespective of the business model. In the hybrid-platform scenario, however, they may still prefer the platform to collect information when the platform is sufficiently efficient at providing ecosystem services. Total welfare is maximized when no information is collected. Yet, conditional on some information being available to the platform, disclosure may be second-best optimal when the platform is sufficiently efficient at providing ecosystem services.

JEL Codes: D82, L13, L86.

Keywords: digital platforms, information disclosure, data acquisition, self-preferencing, hybrid marketplaces, asymmetric information, vertical integration, platform governance.

*I would like to thank Michele Bisceglia, Raffaele Fiocco, Jakub Kastl, Jorge Padilla, Federico Vaccari, and Helder Vasconcelos for useful comments. This research was conducted in a purely academic capacity and is entirely unrelated to my professional activity as a consultant. Any views expressed are solely my own and do not reflect those of any client, employer, or affiliated organization.

†University of Bergamo and Econic Partners LLC: salvatore.piccolo@unibg.it

1 Introduction

Digital ecosystems are, at their core, information systems that observe, process, and govern marketplace activity. Modern platforms do more than simply match buyers and sellers: they build data infrastructures that track user behavior and then deploy these data to shape pricing and service provision. Viewed through this lens, the organization of digital ecosystems is inherently a matter of data governance. How much information do platforms collect, combine, and refine from marketplace activity? And, under what terms is that information disclosed to third-party sellers or, instead, withheld for the platform's own downstream operations?

The distinction between data *acquisition* and data *disclosure* is not merely conceptual. Prominent platforms have faced repeated scrutiny from antitrust authorities in recent years. The Amazon investigations in the European Union and the United Kingdom, for instance, focused on the use of non-public marketplace data together with self-preferencing concerns in the Buy Box and Prime ecosystem.[1] Other interventions have targeted, more directly, the collection, combination, or cross-use of data, including Meta's use of rivals' advertising data to support adjacent entry, the cross-service combination of user data in the Facebook/Meta proceedings before the Bundeskartellamt, and the conditions under which gatekeepers may combine or exploit user data under the Digital Markets Act.[2]

Against this background, it is surprising that the literature still lacks a simple framework that explains how a platform's organizational form shapes both the incentive to acquire commercially valuable information and the incentive to disclose it. This gap matters because different business models fundamentally alter the private value of data. A pure intermediary uses marketplace information to improve coordination and service quality. A hybrid platform can do the same, but it can also deploy non-public information to support its private label or other downstream activities.

To shed light on these issues, this paper studies how platforms manage information along two key margins: data collection and data disclosure. I compare two canonical business models: a pure intermediary, which brokers trade between third-party sellers and monetizes marketplace activity through intermediation, and a hybrid platform, which combines intermediation with downstream competition through a private label. In both settings, the platform acquires a private signal about uncertain demand, chooses how much of that information to disclose, and selects a marketplace-wide service

[1] European Commission, *Antitrust: Commission accepts commitments by Amazon* (Press Release IP/22/7777, 19 December 2022), https://ec.europa.eu/commission/presscorner/detail/en/ip_22_7777; Competition and Markets Authority, *Investigation into Amazon's Marketplace*, https://www.gov.uk/cma-cases/investigation-into-amazons-marketplace.

[2] Competition and Markets Authority, *CMA protects competition by curbing Meta's use of ad customers' data* (3 November 2023), https://www.gov.uk/government/news/cma-protects-competition-by-curbing-metas-use-of-ad-customers-data; Bundeskartellamt, *Facebook proceeding concluded* (10 October 2024), https://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2024/10_10_2024_Facebook.html; European Commission, *Digital Markets Act – Commission starts first proceedings against Apple and Meta* (24 March 2024), https://ec.europa.eu/commission/presscorner/detail/en/ip_24_1689; European Commission, *Commission finds Apple and Meta in breach of the Digital Markets Act* (23 April 2025), https://ec.europa.eu/commission/presscorner/detail/en/ip_25_1085.

effort that raises ecosystem quality and benefits all sellers. This effort captures investment in the marketplace's socio-technical layer, including fulfillment intensity, logistics support, customer service, returns processing, fraud prevention, and the broader operational architecture of the ecosystem.

The baseline model develops a simple Hotelling environment with covered market.[3] Within this setting, a pure intermediary does not value exclusive information and fully discloses whatever it learns. Yet, consumers prefer secrecy because public information makes sellers more responsive to realized demand conditions, thereby strengthening surplus extraction. Total welfare maximization delivers an even more restrictive result, requiring no data acquisition ex ante, since information is costly. However, once the platform owns one of the sellers, non-public information can be used to refine both the private label's price and the platform's service effort. A hybrid platform then strictly prefers exclusive to public information and chooses maximal garbling. Mandatory disclosure works as an indirect cap on data acquisition because it reduces the ex ante value of learning. Consumers also prefer secrecy, but they may still value privately held information when service provision is sufficiently efficient — i.e., reductions in costs are passed on to consumers via lower prices. Total welfare again mandates no data acquisition. Yet, conditional on a positive and non-verifiable stock of data that the platform collects despite a mandate prohibiting its collection, full disclosure can become second-best optimal when service provision is sufficiently efficient that the allocative gains from common adjustment outweigh the additional rent extraction induced by public information.

The extensions show that these conclusions are not tied to a single timing or demand environment, while also clarifying the economic limits of the baseline model. First, when service quality cannot be easily adjusted and must be chosen before the state is realized and before prices are set, the information-sensitive service channel disappears. In that case the pure intermediary continues to disclose everything it learns, whereas the hybrid platform still prefers secrecy, but welfare rankings become simpler because information can no longer improve the state-contingent allocation of platform services. Hence, the effects of information collection on consumer welfare depend on whether platforms can adjust these services in response to the information collected. Second, when I allow for an intensive margin of demand by replacing the Hotelling structure with a differentiated-demand system, the role of disclosure becomes richer. In particular, under hybrid organization disclosure need no longer be unambiguously detrimental to the platform: when products are sufficiently differentiated, making information public can improve the rival seller's response enough for the platform to internalize part of the gain through intermediation revenue, whereas secrecy remains privately optimal when products are closer substitutes and disclosure mainly intensifies competition. The same extension also shows more sharply that public and exclusive information need not affect consumer surplus

[3] The assumption of a covered market is well suited to platform environments in which purchase incidence is largely taken as given and the platform primarily reallocates demand across sellers, such as hotel-booking platforms, food-delivery apps, online marketplaces for standardized goods, app stores, or travel platforms for specific routes and dates. In the Extensions, I consider an alternative demand specification featuring an intensive margin, allowing the model to capture demand-expansion effects.

in the same way, because they operate through distinct channels: exclusive precision changes how effectively the platform can tailor common effort, while public precision changes how aggressively the rival seller responds. Additional remarks on seller participation, seller-specific services, and double marginalization further clarify how the value of transparency depends on whether information affects not only allocation across sellers, but also entry, marketplace variety, and total demand.

Taken together, these results have several implications for the design and governance of information systems in digital platforms. First, they highlight that data governance involves distinct but interdependent decision margins — how much data to collect, how to combine and refine them, and how and with whom to share them. Platforms that both intermediate and compete downstream face fundamentally different trade-offs than pure intermediaries, suggesting that IS research and practice should explicitly account for ownership structure and participation roles when designing data policies. Second, the findings emphasize that data are not only an operational resource for improving service quality, but also a strategic asset that can shape competition within the ecosystem. Finally, the analysis underscores that investments in the platform's socio-technical infrastructure — such as logistics, ranking systems, and customer service — are tightly linked to data governance. Effective platform design requires a coordinated approach in which data policies and system architecture are jointly determined.

The main policy lesson is that regulating data collection and data disclosure are distinct tasks. In the pure-intermediation regime, disclosure obligations mainly affect how information is used, since the platform would in any event disclose what it learns. In the hybrid regime, by contrast, disclosure obligations also affect how much information the platform wants to collect in the first place, because forcing disclosure destroys the value of informational exclusivity. Data-sharing mandates can, therefore, work as indirect limits on data acquisition. A second implication is that consumer protection and total welfare need not point in the same direction. Consumers systematically dislike seller-facing disclosure because it intensifies sellers' responsiveness and rent extraction. But, conditional on data having been collected, total welfare may favor disclosure when this policy improves the allocation of ecosystem quality. This suggests that the desirability of transparency depends on the regulator's ability to control hidden data collection. If bans on collection are imperfectly enforceable, mandatory disclosure may be the best available instrument even when disclosure would not be first-best, although it harms consumers. This consumer-versus-welfare comparison is also distinct from the business-user lens motivating much platform regulation, since third-party sellers care directly about how seller-facing disclosure affects their ability to respond.

More broadly, the results caution against treating transparency as unambiguously pro-consumer or data secrecy as unambiguously anticompetitive. In vertically integrated platforms, privately held data can improve operational decisions, but they can also strengthen the platform's strategic advantage over rivals. The appropriate policy thus depends on which margin is realistically regulable: if data collection can be directly constrained, limiting acquisition is welfare improving; if hidden collection cannot be ruled out, disclosure mandates may serve as a second-best remedy by neutralizing the strategic value of exclusive information. This conclusion is broadly consistent with the recent IS discussion of data sharing and data siloing as distinct governance instruments rather than substitutes (Krämer and Shekhar 2025).

Related literature. This paper sits at the intersection of information-systems (IS) and economics research on data, platforms, and information control.

The IS literature provides a first important building block by treating data as a strategic asset and data governance as an organizational capability. It shows that firms create value not only by collecting data, but also by developing the organizational processes and technical capabilities needed to govern access, integration, and deployment of that information (Mikalef et al. 2020; Engert et al. 2025). Related work studies platform ecosystems more specifically and emphasizes that platforms are not merely transactional infrastructures, but governance systems that allocate visibility, access, and control across heterogeneous participants (Huber, Kude, and Dibbern 2017; von Scherenberg et al. 2024). This perspective is particularly useful here because it highlights the platform's role in shaping who observes what and on what terms.

A related stream of IS research examines how platform governance changes when the platform also competes downstream. These papers show that platform-owner entry, own-brand introduction, and hybrid governance affect seller outcomes, strategic incentives, and the balance of the ecosystem (Zhu and Liu 2018; He et al. 2020; Cheng et al. 2023; Zha et al. 2023). Other contributions focus on the informational instruments through which platforms influence market outcomes, including ranking systems, certification, visibility design, and data-sharing policies (Mao, Dewan, and Ho 2023; Dewan, Kim, and Nian 2023; Krämer and Shekhar 2025). Taken together, this literature provides a rich institutional account of data governance in platform environments. What it does not yet offer, however, is a tractable framework in which the platform jointly chooses how much information to acquire and how much to disclose, and in which these choices depend on the platform's business model.

The economics literature provides a complementary set of tools. A traditional strand studies information sharing in oligopoly and shows that the value of information depends on the mode of competition and on who has access to it (Vives 1984; Gal-Or 1985; Shapiro 1986; Raith 1996). A second strand studies information design and the economics of data more broadly, emphasizing that the value of data depends on ownership, exclusivity, and the allocation of control rights (Bergemann, Bonatti, and Smolin 2018; Bergemann and Bonatti 2019, 2024; Bergemann, Bonatti, and Gan 2022). More recent work brings these insights into platform settings by studying intermediation, vertical integration, hybrid governance, and self-preferencing (Hagiu and Wright 2015; Abhishek, Jerath, and Zhang 2016; de Cornière and Taylor 2019; Hagiu, Teh, and Wright 2022; Long and Amaldoss 2024; Wang and Qiu 2024). This literature shows that organizational form and downstream participation matter for incentives and welfare, but typically takes the informational environment as given or focuses on how an existing informational advantage is used. By contrast, the present paper endogenizes both how much information the platform acquires and how much of that information it shares with third-party sellers.

Overall, the paper brings these two perspectives together in a unified framework. Its contribution is twofold. First, it separates data acquisition from data disclosure and shows that these are distinct strategic margins whose interaction depends on the platform's organizational form. Second, it provides a simple framework to study how restrictions on data collection and seller-facing disclosure mandates affect platform behavior, consumer outcomes, and welfare under pure intermediation and hybrid governance.

Roadmap. The rest of the paper is organized as follows. Section 2 introduces the common environment. Section 3 studies the pure-intermediation model. Section 4 analyzes the hybrid platform. Section 5 discusses extensions and their policy implications. Section 6 concludes. The Appendix reports the proofs, the equilibrium derivations and the calculations of expected platform profits, consumer surplus, and total welfare in the same order in which the corresponding results appear in the text.

2 The baseline model

This section introduces a simple environment that accommodates the two business models studied in the paper. In the pure intermediation model, the platform brokers transactions between two third-party sellers and earns a share of their operating profits. In the hybrid model, the platform instead owns a seller and competes against the independent seller through a private label. The demand system, service technology, and information structure are identical across the two scenarios, so the comparison isolates how vertical integration changes the incentives to collect data and disclose them.

Environment. Consider two sellers, denoted by S_1 and S_2, located at the extremes of the Hotelling unit segment. In the pure intermediation case, both are third-party sellers hosted by the platform. In the hybrid case, S_1 is the platform's integrated retail unit, while S_2 remains independent. Seller S_1 is located at 0, while seller S_2 is located at 1. Consumers are uniformly distributed on the unit interval and face quadratic transport costs.[4] A consumer located at $z \in [0, 1]$ obtains utility

$$u_1 = v + \sigma\theta - p_1 - tz^2$$

[4]Since the equilibrium outcome is asymmetric, quadratic transport costs ensure the existence of an equilibrium in which both sellers remain active and the market is fully covered.

if she buys from S_1, and

$$u_2 = v - \sigma\theta - p_2 - t(1-z)^2$$

if she buys from S_2. The parameter $t > 0$ represents the unit transportation cost, or equivalently, the degree of product differentiation. The binary state θ, which tilts market conditions toward one seller and away from the other, satisfies

$$\theta \in \{-1, +1\}, \qquad \Pr[\theta = +1] = \Pr[\theta = -1] = \frac{1}{2},$$

with $\sigma > 0$ measuring the intensity of the demand tilt generated by the binary state.[5] Assuming a covered market — which requires σ sufficiently small — the indifferent consumer is located at

$$x^*(\theta, p_1, p_2) = \frac{1}{2} + \frac{\sigma\theta}{t} - \frac{p_1 - p_2}{2t}.$$

Thus, the demand system is

$$q_i(\theta, p_i, p_j) = \frac{1}{2} + \frac{\mathbb{I}_i \sigma\theta}{t} - \frac{p_i - p_j}{2t}, \quad i, j = 1, 2,$$

with $\mathbb{I}_1 = 1$ and $\mathbb{I}_2 = -1$.

The hypothesis of covered market is appealing if one thinks of a single consumer distributed on the Hotelling line who has to decide from which seller to buy, rather than whether to buy at all. This formulation is well suited to many platform environments in which purchase incidence is essentially taken as given and the platform mainly reallocates demand across sellers, such as hotel-booking platforms where a traveler has already decided to book accommodation, food-delivery apps where the consumer has already chosen to order a meal, online marketplaces for standardized goods where buyers compare alternative sellers of the same product, app stores where users search among competing apps within a given category, or travel platforms where a user needs to book a flight on a specific route and date. In the Extensions, I consider both an uncovered market and an alternative demand specification featuring an intensive margin, which allows the model to capture demand-expansion effects.

Both sellers have the same baseline marginal cost $c > 0$. The platform chooses a service effort $a \geq 0$ that lowers the effective marginal cost of both sellers from c to $c - a$. This effort captures fulfillment intensity, warehouse and customer-support resources, return and refund handling, fraud screening, and related platform services that benefit all sales on the marketplace — i.e., in short, it is a measure of the ecosystem quality. The cost of effort is quadratic:

$$\psi(a) = \frac{a^2}{2b},$$

[5] In the Extensions, I show that results are robust to introducing a continuous state.

where $b > 0$ measures the efficiency of platform service effort.

Information structure. Neither the platform nor the sellers observe the binary state $\theta \in \{-1, +1\}$. However, the platform can acquire a binary demand signal $s_I \in \{-1, +1\}$ such that the posterior mean satisfies

$$\mathrm{E}\left[\theta \mid s_I\right] = \sqrt{\eta_I}\, s_I.$$

It chooses the precision $\eta_I \in [0, 1)$ of this signal. Equivalently, the signal is correct with probability

$$\Pr\left[s_I = \theta \mid \theta\right] = \frac{1 + \sqrt{\eta_I}}{2}.$$

Upon observing the private signal s_I, the platform also releases a binary public signal $s_P \in \{-1, +1\}$ such that

$$\mathrm{E}\left[\theta \mid s_P\right] = \sqrt{\eta_P}\, s_P,$$

with precision $\eta_P \in [0, \eta_I]$. Hence, when $\eta_P = \eta_I$, the platform fully reveals its private information; when $\eta_P = 0$, it reveals nothing. The analysis below is therefore conditioned directly on the two realized signals (s_I, s_P). The signal s_P summarizes the information commonly available in the market, while s_I summarizes the platform's internal (private) information.

By the law of iterated expectations,

$$\mathrm{E}[\theta \mid s_P] = \mathrm{E}[\mathrm{E}[\theta \mid s_I] \mid s_P],$$

so the public posterior mean is the seller's forecast of the platform's posterior mean. In particular,

$$\mathrm{E}\Big[(\mathrm{E}[\theta \mid s_P])^2\Big] = \eta_P, \qquad \mathrm{E}\Big[(\mathrm{E}[\theta \mid s_I] - \mathrm{E}[\theta \mid s_P])^2\Big] = \eta_I - \eta_P,$$

and

$$\mathrm{E}[\mathrm{Var}(\theta \mid s_I)] = 1 - \eta_I.$$

These identities are the only moment restrictions needed in the analysis. They imply that all equilibrium objects can be written directly as functions of (s_I, s_P).[6]

[6] The binary baseline makes the information architecture easy, but the same economic logic extends to a continuous Gaussian model such that

$$u_i = \begin{cases} v + \theta - p_1 - tz^2 & \text{if } i = 1, \\ v - \theta - p_2 - t\,(1 - z)^2 & \text{if } i = 2. \end{cases}$$

The demand state is

$$\theta \sim N(0, \sigma^2),$$

while the private signal and the public signal are respectively

$$s_I = \theta + \varepsilon_I, \qquad \varepsilon_I \sim N\left(0, \sigma^2 \frac{1 - \eta_I}{\eta_I}\right),$$

Following Vives (2010), the information-acquisition cost is

$$K(\eta_I) = k\frac{\eta_I}{1-\eta_I},$$

where $k > 0$ is the cost parameter. This specification is increasing and convex, and satisfies standard Inada conditions.

Timing and equilibrium concept. The acquisition and signal stages are common across the two business models.

1. The platform chooses (η_I, η_P).
2. The taste shock θ realizes. The platform observes s_I, and the public signal s_P is observed by the sellers.
3. Downstream decisions are taken. In the pure intermediation model, the platform chooses service effort a while both sellers set prices. In the hybrid model, the platform chooses service effort a and its private-label price p_1, while the independent seller chooses p_2.
4. Demand is allocated and profits are realized.

The equilibrium concept is Bayes-Nash Equilibrium. To keep the contracting dimension of the model simple, let $\tau \in (0, 1)$ denote the share of third-party operating profit appropriated by the platform — i.e., an index of its bargaining power. In the pure intermediation model, the platform receives the same share τ from both sellers. In the hybrid benchmark, the platform receives the share τ only from the independent seller S_2, while it fully internalizes the private-label profit of S_1. A natural interpretation (payoff-equivalent) of this reduced-form assumption is that it captures the outcome of an underlying bargaining game between the platform and sellers, as in Chemla (2003) and Nocke and Rey (2018), in which either side makes take-it-or-leave-it offers with some probability. Specifically, with probability $\tau \in [0, 1]$, the platform makes take-it-or-leave-it offers to the sellers.

I assume that the service effort a is chosen simultaneously with prices because it is a short-run, demand-contingent operational decision, rather than a long-term commitment. The platform observes the same information as sellers (its private signal and the public signal) and adjusts service intensity in real time, just as sellers adjust prices. This assumption is particularly appropriate in environments

and

$$s_P = s_I + \varepsilon_P, \qquad \varepsilon_P \sim N\left(0, \sigma^2\frac{\eta_I - \eta_P}{\eta_I\eta_P}\right).$$

Yet, as in other linear-normal models applied to oligopoly theory, the Gaussian support of θ implies that extreme realizations can violate the interior and full-coverage conditions with negligible probability when the variance σ^2 of the random process is small enough. In this case, the model should therefore be interpreted in expected terms (see, e.g., Vives, 2010).

where operational inputs can be adjusted flexibly and rapidly, or where the effects of these investments materialize only after purchase. Examples include fulfillment speed and prioritization (e.g., faster shipping for high-demand products), inventory allocation across warehouses, customer service staffing, fraud screening, and returns processing. In all these cases, the platform can scale effort up or down in response to demand conditions at the same time as sellers revise their prices. Moreover, many of these service dimensions are imperfectly observable or not contractible for sellers. This limits the platform's ability to use service effort as a commitment device and reinforces the idea that a is adjusted contemporaneously with pricing decisions rather than chosen in advance. In the extensions, I consider the case in which a is chosen before price competition.

Since the platform owns private information, the solution concept is Bayes-Nash Equilibrium (BNE). We restrict attention to product-market equilibria in linear strategies.

Technical restrictions. Throughout, I impose the following technical assumptions:

Assumption 1 (Regularity)**.** *The ratio $\frac{v}{\sigma^2}$ is large enough that the market is covered and both sellers make positive sales irrespective of θ. In addition, I also impose*

$$b < \bar{b} \equiv \frac{4t}{(1-\tau)^2},$$

to guarantee positive prices.

3 Pure intermediation platform

I begin with the pure intermediary. After signals are realized, the platform observes (s_I, s_P) and chooses common service effort a, while the two independent sellers set prices conditional only on the public signal s_P. Because the platform earns the same commission rate on both sides and total demand is fixed under full coverage, the extra information contained in s_I beyond s_P has no downstream role in this case.

Proposition 2 (Pure intermediation: market equilibrium)**.** *With a pure intermediary, the unique linear market equilibrium is such that*

$$a^{\mathrm{PI}} = b\tau, \qquad p_i^{\mathrm{PI}}(s_P) = c - b\tau + t + \mathbb{I}_i \frac{2}{3}\sigma\sqrt{\eta_P}\, s_P, \qquad i = 1, 2.$$

A favorable public signal raises $p_1^{\mathrm{PI}}(s_P)$ and lowers $p_2^{\mathrm{PI}}(s_P)$. The platform's internal signal s_I is payoff-irrelevant once s_P is fixed.

Under pure intermediation, the platform benefits from data only insofar as they improve third-party seller profit and, through that channel, the commissions it earns. Since hidden precision has no downstream use, any information the platform acquires is valuable only when it is publicly disclosed. Moreover, since I focus on an equilibrium in which the market is fully covered and both sellers price based on the same information, the platform's effort is constant — i.e., it is independent of the realized signals once the market remains covered.

Using the equilibrium allocation, the platform's ex ante objective can be written as (see the Appendix)

$$\Pi_P^{\mathrm{PI}}(\eta_I, \eta_P) = \Pi_0^{\mathrm{PI}} + \frac{4\tau\sigma^2}{9t}\eta_P - K(\eta_I), \tag{1}$$

where Π_0^{PI} is a constant component that does not depend on (η_I, η_P). Differentiating this function with respect to η_I and η_P subject to $\eta_P \leq \eta_I$, the following holds.

Proposition 3 (Pure intermediation: disclosure and acquisition)**.** *With a pure intermediary, the platform fully discloses whatever it acquires — i.e.,* $\eta_P^{\mathrm{PI}} = \eta_I^{\mathrm{PI}}$*. Moreover, it acquires data if and only if*

$$k < \bar{k}^{\mathrm{PI}}(t, \tau, \sigma) \equiv \frac{4\tau\sigma^2}{9t}.$$

Whenever this inequality holds, the optimal precision is

$$\eta_I^{\mathrm{PI}} = \eta_P^{\mathrm{PI}} = 1 - \sqrt{\frac{k}{\bar{k}^{\mathrm{PI}}(t, \tau, \sigma)}},$$

which is decreasing in t *and* k*, and increasing in* τ *and* σ*.*

In the pure intermediation case, the platform acquires information only because better-informed sellers make choices that are more closely aligned with demand — i.e., the seller facing relatively higher demand can charge a higher price — thereby increasing marketplace performance and the commissions the platform collects. Since the platform has no integrated retail arm, it has no reason to keep information private: all acquired precision is disclosed. The optimal precision is decreasing in t because, when products are more differentiated, competition is weaker and seller decisions become less responsive to demand information. It is decreasing in k because information becomes more expensive, so the platform invests less in it. By contrast, optimal precision is increasing in τ because a higher commission rate allows the platform to appropriate a larger share of the gains generated by better seller decisions. This raises the return to information acquisition. Finally, it is also increasing in σ: the more strongly the demand state affects consumer preferences, the more valuable it becomes to learn that state.

Importantly, these results do not hinge on the assumption of a covered market. When the market is uncovered and sellers behave as local monopolists, the platform still fully discloses any information it acquires, for the same reasons as above.

3.1 Welfare analysis

In this section, I consider two welfare standards: consumer surplus and total welfare, and compare how the information policy that maximizes these objects compares to the policy that maximizes the platform's expected profit. In the Appendix, I show that expected consumer surplus is

$$CS^{\mathrm{PI}}(\eta_I, \eta_P) = CS_0^{\mathrm{PI}} + \left(1 - \frac{8}{9}\eta_P\right)\frac{\sigma^2}{t}, \tag{2}$$

and expected total welfare is

$$W^{\mathrm{PI}}(\eta_I, \eta_P) = W_0^{\mathrm{PI}} + \left(1 - \frac{4}{9}\eta_P\right)\frac{\sigma^2}{t} - K(\eta_I), \tag{3}$$

where CS_0^{PI} and W_0^{PI} are two constant components that do not depend on (η_I, η_P). Differentiating these functions with respect to η_I and η_P subject to $\eta_P \leq \eta_I$, the following holds.

Corollary 4 (Pure intermediation: consumers and welfare)**.** *Under pure intermediation, both consumer surplus and total welfare are maximized when no public information is disclosed — i.e., when* $\eta_P = 0$*. Consumers are indifferent over* η_I*, whereas total welfare is maximized when no private information is collected. Hence, the pure intermediary over-discloses relative to both consumer surplus and total welfare maximization. Whenever* $k < \bar{k}^{\mathrm{PI}}(t, \tau, \sigma)$*, it also over-acquires information relative to total welfare maximization.*

Sharing information intensifies state-contingent pricing: when demand is high, sellers raise prices more aggressively, extracting more surplus from consumers. This reduces consumer surplus, which is therefore maximized when no information is shared. At the same time, consumers are indifferent about data acquisition itself, because data that remain private do not affect market outcomes. Total welfare maximization implies an even more restrictive view. While data can improve the allocation of products to demand conditions, these gains are limited in the pure intermediation setting, where competition is unaffected. By contrast, data acquisition is costly. As a result, the planner prefers no data acquisition and no disclosure. The key takeaway is that, under pure intermediation, the platform has stronger incentives to both acquire and disclose information than is socially optimal. This is because it internalizes the benefits of improved seller performance, but not the redistribution from consumers to sellers nor the cost of acquiring information.

4 Hybrid platform

Consider now a hybrid platform. Relative to the pure intermediation case, a vertically integrated platform can use non-public information not only to improve common marketplace services but also to steer downstream competition toward its own retail arm. That additional use for exclusive information is what drives the hybrid platform's bias toward opacity.

Fix an information structure (η_I, η_P). After signals are realized, the platform knows (s_I, s_P), whereas the third-party seller only knows s_P. For given prices (p_1, p_2), conditional expected (forecasted) demands from the platform's perspective are

$$q_1(p_1, p_2 \mid s_I) = \frac{1}{2} + \frac{\sigma\sqrt{\eta_I}\, s_I}{t} - \frac{p_1 - p_2}{2t},$$

and

$$q_2(p_2, p_1 \mid s_I) = \frac{1}{2} - \frac{\sigma\sqrt{\eta_I}\, s_I}{t} - \frac{p_2 - p_1}{2t} = 1 - q_1(p_1, p_2 \mid s_I).$$

Let $p_2^{\mathrm{H}}(s_P)$ be the independent seller's equilibrium price conditional on the public signal. For every state (s_I, s_P), the platform chooses (p_1, a) to solve

$$\max_{p_1 \in \mathbb{R}_+,\, a \in \mathbb{R}_+} \left\{ q_1(p_1, p_2^{\mathrm{H}}(s_P) \mid s_I)(p_1 - (c-a)) + \tau q_2(p_2^{\mathrm{H}}(s_P), p_1 \mid s_I)(p_2^{\mathrm{H}}(s_P) - (c-a)) - \frac{a^2}{2b} \right\}.$$

The corresponding first-order conditions with respect to p_1 and a are, respectively,

$$q_1(p_1, p_2^{\mathrm{H}}(s_P) \mid s_I) - \frac{p_1 - (c-a)}{2t} + \tau \frac{p_2^{\mathrm{H}}(s_P) - (c-a)}{2t} = 0, \tag{4}$$

and

$$q_1(p_1, p_2^{\mathrm{H}}(s_P) \mid s_I) + \tau q_2(p_2^{\mathrm{H}}(s_P), p_1 \mid s_I) - \frac{a}{b} = 0. \tag{5}$$

An increase in S_1's price has the usual margin and volume effects on the profits of the private label, but it also raises S_2's profit, a fraction τ of which is internalized by the platform. An increase in quality raises demand for both the platform's own product and the third-party product. The platform internalizes both effects: it directly benefits from higher sales of its private label and, through the commission rate, from higher sales of the third-party seller. This must be balanced against the marginal cost of quality. Intuitively, quality provision is not targeted to a single product, but reflects the platform's overall ecosystem incentives: it increases quality up to the point where the combined gains from both its own sales and the commissions it earns are exactly offset by the cost of providing quality.

Next, define the forecasting objects relevant for the third-party seller. Let $p_1^{\mathrm{H}}(s_I, s_P)$ and $a^{\mathrm{H}}(s_I, s_P)$ be the equilibrium price and equilibrium effort set by the platform conditional on state (s_I, s_P). Then,

$$p_1^{\mathrm{H}}(s_P) \equiv \mathrm{E}\left[p_1^{\mathrm{H}}(s_I, s_P) \mid s_P\right], \qquad a^{\mathrm{H}}(s_P) \equiv \mathrm{E}\left[a^{\mathrm{H}}(s_I, s_P) \mid s_P\right]$$

denote the objects forecast by S_2 conditional on the public signal. Given s_P, seller S_2 solves

$$\max_{p_2\in\mathbb{R}_+} \left\{ \mathrm{E}\left[q_2(p_1^{\mathrm{H}}(s_I,s_P),p_2 \mid s_I)\big(p_2-\big(c-a^{\mathrm{H}}(s_I,s_P)\big)\big) \mid s_P \right] \right\},$$

Because choices are simultaneous, seller S_2 treats $p_1^{\mathrm{H}}(s_I,s_P)$ and $a^{\mathrm{H}}(s_I,s_P)$ as fixed random variables when it chooses p_2. Using the fact that

$$\mathrm{E}\left[\theta \mid s_P\right] = \sqrt{\eta_P}\, s_P,$$

the first-order condition is

$$\frac{1}{2} - \frac{\sigma\sqrt{\eta_P}\, s_P}{t} - \frac{2p_2 - p_1^{\mathrm{H}}(s_P) - \big(c - a^{\mathrm{H}}(s_P)\big)}{2t} = 0, \tag{6}$$

which again mirrors the usual margin-volume trade-off.

Solving the above system of first-order conditions, it holds that.

Proposition 5 (Hybrid platform: market equilibrium)**.** *With a hybrid platform, the market subgame has a unique linear Bayes-Nash equilibrium. The platform's price and service effort are*

$$\begin{aligned} p_1^{\mathrm{H}}(s_I,s_P) &= c + \frac{(3+\tau)\,(2t-b)}{2\,(3-\tau)} + \frac{(1-\tau)(2t-b)}{t(3-\tau)}\sigma\sqrt{\eta_P}\, s_P \\ &\quad + \frac{2\big(2t-b(1-\tau)^2\big)}{4t-b(1-\tau)^2}\sigma\Big(\sqrt{\eta_I}\, s_I - \sqrt{\eta_P}\, s_P\Big), \end{aligned}$$

$$a^{\mathrm{H}}(s_I,s_P) = b\left[\frac{3+\tau}{2(3-\tau)} + \frac{1-\tau}{t(3-\tau)}\sigma\sqrt{\eta_P}\, s_P + \frac{2(1-\tau)}{4t-b(1-\tau)^2}\sigma\Big(\sqrt{\eta_I}\, s_I - \sqrt{\eta_P}\, s_P\Big)\right],$$

while the third-party seller charges

$$p_2^{\mathrm{H}}(s_P) = c + \frac{6t-b(3+\tau)}{2\,(3-\tau)} - \frac{2t+b(1-\tau)}{t(3-\tau)}\sigma\sqrt{\eta_P}\, s_P.$$

The public signal affects both sellers' prices, whereas the platform's extra information beyond s_P affects only $p_1^{\mathrm{H}}(\cdot)$ and $a^{\mathrm{H}}(\cdot)$. Moreover,

$$\begin{aligned} \frac{\partial p_1^{\mathrm{H}}(s_I,s_P)}{\partial s_P} &> 0 \quad\Longleftrightarrow\quad b > \frac{1+\tau}{2}\bar{b}, \\ \frac{\partial p_1^{\mathrm{H}}(s_I,s_P)}{\partial s_I} &> 0 \quad\Longleftrightarrow\quad b < \frac{\bar{b}}{2}, \end{aligned}$$

$p_2^{\mathrm{H}}(s_P)$ is decreasing in s_P, while $a^{\mathrm{H}}(s_I,s_P)$ is increasing in both signals.

The key intuition is that the two signals have different strategic meanings. The public signal is observed by both sellers, so it affects the entire competitive environment: when s_P increases, the third-party seller reacts by cutting its price, since better public news strengthens the platform's position and makes competition tougher. The platform also responds to s_P, but its pricing reaction is ambiguous because two effects work in opposite directions: better market conditions tend to raise its optimal price, while the fact that the rival observes the same information pushes the platform to compete more aggressively. This is why S_1's price increases in s_P only when the service dimension is sufficiently important. By contrast, the private signal s_I is extra information available only to the platform, so it affects only the platform's own price and service decision, not the third-party seller's price. A higher s_I makes the platform more optimistic, but whether it translates that optimism into a higher price or into a more aggressive market stance again depends on how valuable service is: when b is small, the platform mainly exploits private good news by charging more, whereas when b is large, it has a stronger incentive to compete through service and a lower price.

Using the equilibrium allocation, the platform's ex ante profit can be written as (see the Appendix)

$$\Pi_P^{\mathrm{H}}(\eta_I, \eta_P) = \Pi_0^{\mathrm{H}} + \left[\frac{4t - b(1-\tau)^2}{2t^2(3-\tau)^2} \eta_P + \frac{2}{4t - b(1-\tau)^2} (\eta_I - \eta_P) \right] \sigma^2 - K(\eta_I), \tag{7}$$

where Π_0^{H} is a constant term that does not depend on (η_I, η_P). Differentiating with respect to η_I and η_P subject to $\eta_P \leq \eta_I$, the following holds.

Proposition 6 (Hybrid platform: private disclosure and acquisition)**.** *The hybrid platform chooses maximal garbling — i.e., $\eta_P^{\mathrm{H}} = 0$. It acquires information if and only if*

$$k < \bar{k}_P^{\mathrm{H}}(b, t, \tau, \sigma) \equiv \frac{2\sigma^2}{4t - b(1-\tau)^2}.$$

Whenever this inequality holds, the optimal precision is

$$\eta_I^{\mathrm{H}} = 1 - \sqrt{\frac{k}{\bar{k}_P^{\mathrm{H}}(b, t, \tau, \sigma)}},$$

which is decreasing in t, k, and τ, and increasing in b and σ.

Unlike a pure intermediary, a hybrid platform has a direct stake in the marketplace through its private-label product. As a result, information has a private strategic value: by keeping it hidden, the platform can use it to adjust its own price and quality more precisely than its rival and thereby shift demand toward its retail arm. This is why the platform chooses maximal garbling: disclosure would make the rival equally responsive to demand and dissipate this advantage. Data acquisition is then driven by a trade-off between this business-stealing benefit and the cost of precision. The comparative statics reflect how strong this strategic value is. Optimal precision is decreasing in t because, when products are more differentiated, competition is weaker and the ability to reallocate demand is less valuable. It is decreasing in k because information is more costly. It is also decreasing in τ, since a higher degree of profit internalization reduces the platform's incentive to shift demand toward its own product. Finally, optimal precision is increasing in b, as greater service efficiency enhances the value of information — i.e., improved forecasts enable more effective adjustment of service effort to demand, thereby increasing the return to information — and increasing in σ because the more strongly the demand state affects consumer preferences, the more valuable it becomes to learn that state.

What happens with an uncovered market? When the two sellers operate under local monopolies, the platform no longer has an incentive to keep information private. Because sellers do not compete for the same customers, withholding information cannot be used to shift demand. As a result, the platform always discloses what it knows. However, it still has strong incentives to acquire data, since information helps improve its own operations and service quality.

4.1 Welfare analysis

I now turn to the welfare analysis. I first solve for the information policy that maximizes consumer surplus, then for the policy that maximizes total welfare, and finally compare both with the platform's privately optimal information policy.

Using the equilibrium values characterized above, in the Appendix I show that expected consumer surplus is

$$CS^{\mathrm{H}}(\eta_I, \eta_P) = CS_0^{\mathrm{H}} + \left[\frac{1}{t(3-\tau)^2}\eta_P + \frac{4t}{\left(4t - b(1-\tau)^2\right)^2}(\eta_I - \eta_P) + \frac{1-\eta_I}{t}\right]\sigma^2, \tag{8}$$

where CS_0^{H} is a constant term that does not depend on (η_I, η_P). Differentiating with respect to η_I and η_P subject to $\eta_P \leq \eta_I$, the following holds.

Proposition 7 (Hybrid platform: consumer surplus)**.** *Consumers always prefer the platform not to disclose information to the third-party rival — i.e., $CS^{\mathrm{H}}(\cdot)$ is maximized at $\eta_P = 0$. Moreover, there exists a threshold $b_C^{\mathrm{H}}(t,\tau) > 0$, increasing in t and τ, such that $CS^{\mathrm{H}}(\cdot)$ is maximized at $\eta_I = 1$ if $b > b_C^{\mathrm{H}}(t,\tau)$; otherwise, it is maximized at $\eta_I = 0$. Hence, from a consumer-surplus standpoint, conditional on $\eta_I^{\mathrm{H}} > 0$, the platform over-acquires information when $b < b_C^{\mathrm{H}}(t,\tau)$ and under-acquires information otherwise.*

Consumers dislike public disclosure because shared information makes both sellers more responsive to realized demand conditions, thereby strengthening state-contingent surplus extraction. As a result, consumers and the platform have aligned incentives for the platform not to disclose its private information to the rival. Consumers, however, may still benefit from the accuracy of the platform's private information. When b is small, concealed data primarily support private targeting, which harms consumers. When b is sufficiently large, however, the induced service-effort response becomes strong enough that consumers begin to benefit from privately held data, as greater investment by the platform is passed on through lower prices.

Consider now total expected welfare. Using the equilibrium values characterized above, in the Appendix, I show that

$$W^{\mathrm{H}}(\eta_I, \eta_P) = W_0^{\mathrm{H}} + \frac{10t - 4\tau t - b(1-\tau)^2}{2t^2(3-\tau)^2}\eta_P \sigma^2 + \left[\frac{1-\eta_I}{t} + \frac{6\big(2t - b(1-\tau)^2\big)}{\big(4t - b(1-\tau)^2\big)^2}(\eta_I - \eta_P)\right]\sigma^2 - K(\eta_I), \tag{9}$$

where W_0^{H} is a constant term that does not depend on (η_I, η_P). Differentiating with respect to η_I and η_P subject to $\eta_P \leq \eta_I$, the following holds.

Proposition 8 (Hybrid platform: total welfare)**.** *From a total welfare perspective, data acquisition is never optimal — i.e., $W^{\mathrm{H}}(\eta_I, \eta_P)$ is maximized at $\eta_I = 0$. However, conditional on a positive stock of privately acquired data — i.e., if there exists a minimal amount of information that the platform cannot be forced to conceal — there exists a threshold $b_W^{\mathrm{H}}(t, \tau) > 0$, increasing in t and τ, such that full disclosure is second-best optimal if and only if $b > b_W^{\mathrm{H}}(t, \tau)$.*

With a fully covered market, information has only a distributive effect, while data acquisition is itself costly. As a result, from an ex ante perspective, total welfare is maximized when no data are collected. Conditional on data being available, however, the planner faces a second-best trade-off. Disclosure has two opposing effects: it improves coordination and the alignment of decisions with demand (a positive effect), but it also amplifies firms' responsiveness and hence rent extraction (a negative effect). When b is large, quality is cheaper to supply and the allocative gains dominate, so full disclosure is desirable. When b is small, the main effect of disclosure is to intensify price responses, and secrecy is preferred. The threshold $b_W^{\mathrm{H}}(t, \tau)$ captures exactly this trade-off. It is increasing in t and τ because stronger differentiation and greater platform internalization both dampen the allocative benefits of disclosure relative to its rent-extraction effects, making disclosure desirable only when quality responses are sufficiently strong. In these instances, the desirability of disclosure depends on how information translates into real efficiency gains. When platform quality is sufficiently efficient, better information improves the allocation of ecosystem services across demand states.

5 Extensions

This section presents a set of extensions that build on the baseline comparison between pure intermediation and hybrid organization, and show how the policy implications vary with the economic environment. The first four subsections solve formal variants of the model. The remaining ones discuss nearby environments that the baseline helps organize but that are left unmodeled here to keep the paper concise.

5.1 Disclosure floors

Minimum disclosure requirements are often justified on fairness and non-discrimination grounds, as they limit the platform's informational advantage over third-party sellers. However, as the model shows, leveling the playing field is not equivalent to maximizing consumer surplus, at least from an information-management perspective. Greater transparency makes all sellers more responsive to demand, which can intensify rent extraction from consumers — especially when regulating transparency is easier than regulating information collection. Yet, from a total-welfare perspective, transparency mandates can be second-best when direct regulation of the amount of data collected by the platform is difficult because data acquisition is hard to observe, verify, or enforce.

To see how this policy works, suppose that a regulator imposes a proportional disclosure floor of the form $\eta_P \geq \lambda\eta_I$ for some $\lambda \in (0, 1]$. In the pure intermediation case the constraint is slack, because the platform already sets $\eta_P = \eta_I$. The regulatory instrument, therefore, has no bite either for the platform or for the welfare perspective. Yet, the same policy binds in the hybrid model. Private disclosure is chosen at the lower bound. For consumers the floor is unattractive because it pushes the market toward more public responsiveness. For total welfare the floor is valuable only in the region identified by Proposition 8, namely when regulation has a limited impact on private information and service efficiency is large enough that the allocative gains from common adjustment dominate the extra rent extraction generated by public information.

The policy lesson is, therefore, business-model specific: a disclosure floor is redundant for a pure intermediary, but can be a genuine disciplining instrument for a hybrid platform when the standard is total welfare maximization. Yet, this policy may backfire with the incentive to collect information — i.e., it may reduce the value of private information. Suppose for instance that regulation mandates full disclosure — i.e., $\lambda = 1$.

Corollary 9 (Hybrid platform: disclosure as an indirect acquisition cap)**.** *Mandatory disclosure weakly reduces data acquisition.*

This result highlights an important trade-off between the disclosure and the acquisition of information. Mandating disclosure improves the informational environment ex post, by allowing all sellers to condition their decisions on more precise signals. However, it also reduces the platform's incentive to acquire information ex ante, because it limits the private value of exclusive data. This creates a fundamental policy tension. On the one hand, greater transparency improves how information is used once it is available. On the other hand, it weakens the incentives to generate that information in the first place. The optimal second-best policy must, therefore, balance the gains from disclosure against the loss in information acquisition.

The next corollary provides interesting comparative statics.

Corollary 10 (Hybrid platform: revenue sharing and acquisition)**.** *A higher τ lowers information acquisition under platform discretion but raises it under mandatory disclosure.*

An increase in the platform's bargaining power does not uniformly strengthen incentives to acquire data. Absent mandated disclosure rules, it reduces the value of information, because the platform internalizes a larger share of the rival seller's profits and therefore has weaker incentives to reallocate demand toward its own retail arm. Under mandatory disclosure, by contrast, the same increase raises the return to information, as the platform captures a larger share of the gains generated when the third-party seller also responds to the disclosed signal.

5.2 Service effort set before prices

The baseline assumes that common service effort is chosen contemporaneously with downstream prices. This is natural for operational inputs that can be adjusted in real time or that are hard to verify ex ante — e.g., logistic services that take place after purchase. In some environments, however, service design can be a commitment variable: the platform chooses a service level in advance and prices are set afterwards. To isolate that case cleanly, suppose here that the platform chooses a common effort level *before* the demand state and before any signal is observed. Prices are then chosen after information is realized. The following holds.

Proposition 11 (Effort commitment before pricing)**.** *Suppose effort is chosen before the demand realizes. Then, the platform sets zero effort under both pure intermediation and hybrid organization.*

Under pure intermediation, the information-design result is unchanged: the platform fully discloses whatever it learns and cquires information if and only if

$$k < \frac{4\tau\sigma^2}{9t}.$$

Consumer surplus is maximized at zero public disclosure and is indifferent over undisclosed information, while total welfare is maximized when no information is collected.

Under hybrid organization, the platform still chooses maximal garbling, and it acquires information if and only if

$$k < \frac{\sigma^2}{2t}.$$

Consumer surplus and total welfare are maximized when no information is collected and no information is disclosed.

The logic is straightforward. When effort is chosen before signals and prices, and is common to both sellers, it acts like a common cost shifter that cannot be adjusted to realized demand conditions. The subsequent pricing game fully passes that shift through into retail prices, so the platform cannot use common effort either to create a separate strategic advantage or to improve the state-contingent allocation of marketplace services. The pure intermediary, therefore, continues to disclose all acquired information, while the hybrid platform still values exclusive information for its direct effect on downstream pricing. What disappears is the information-sensitive service channel. The welfare implications thus become simpler: under pure intermediation consumer-surplus and welfare rankings are unchanged compared to the baseline model, whereas under a hybrid organization, information no longer generates any offsetting service-adaptation benefit and, as a result, both consumers and total welfare are maximized under no information acquisition and no disclosure.

Remark. Allowing effort to be set after signals but before prices does not restore an operational role for this choice. If effort is publicly observed, it merely becomes an additional communication device. In the hybrid case, however, the platform strictly dislikes any endogenous disclosure of its private information, so the relevant equilibrium is pooling in effort; since common effort is otherwise payoff-irrelevant and costly, the platform again sets $a = 0$.

5.3 Intensive margin

In this section, I move from a Hotelling framework to a Singh-Vives (1984) specification in order to adopt a richer and more flexible representation of differentiated-product competition. While Hotelling offers a very intuitive spatial interpretation of horizontal differentiation, under full market coverage it is best suited to environments in which firms mainly compete for market shares, since total demand is fixed and changes in prices or information primarily reallocate consumers across sellers. By contrast, the Singh-Vives structure allows for a demand system in which product differentiation remains central, but firms' strategies may affect not only the division of demand across sellers but also the overall intensity of market demand. This is particularly useful in our setting, where platform decisions such as information management or service provision may do more than simply shift consumers from one seller to another: they may also influence the overall attractiveness of the marketplace.

To make the analysis simple, suppose that

$$q_i\left(\theta, p_i, p_j\right) = \frac{\left(1+\sigma \mathbb{I}_i \theta\right)(1-\gamma) - p_i + \gamma p_j}{1-\gamma^2}, \qquad i, j = 1, 2,$$

where, as before, $\mathbb{I}_1 = 1$ and $\mathbb{I}_2 = -1$. The parameter $\gamma \in [0, 1)$ measures product substitutability — i.e., the higher γ, the closer substitutes products are. The shock θ remains purely relative: it shifts demand toward one seller and away from the other without directly changing market size. To make the analysis tractable, throughout, I assume that $b = 1$, so the cost of service effort is $a^2/2$, and that $\tau = \frac{1}{2}$ so that each side has equal bargaining power. Results are qualitatively robust to non-degenerate distributions of bargaining power. The extreme cases clarify the role of profit internalization. When $\tau = 0$, the platform does not internalize the rival's profit and therefore has a stronger incentive to garble information, as the steering motive is more prominent. By contrast, when $\tau = 1$, the steering effect disappears, as the platform fully internalizes the rival's profit, thereby strengthening its incentive to disclose information. The rest of the model structure is as in the baseline model. The full analysis is offered in the Appendix. Yet, a useful observation is that

$$\sum\nolimits_{i=1,2} q_i\left(\theta, p_i, p_j\right) = \frac{2 - p_1 - p_2}{1+\gamma},$$

so the state θ changes the composition of demand but not its total size. Hence, information matters through state-contingent pricing and market steering, not because it directly shifts aggregate demand.

Under pure intermediation, no additional regularity condition is needed: for every $\gamma \in [0,1)$ the downstream game has a unique interior linear equilibrium. Under hybrid organization, the platform's problem is strictly concave provided that

$$\gamma < \gamma_{\max} \equiv \frac{2+2\sqrt{10}}{9}.$$

With a pure intermediary, the following holds.

Proposition 12 (Intensive margin and pure intermediation)**.** *Under pure intermediation, the platform fully discloses whatever it acquires. It chooses positive precision if and only if k is sufficiently small. The incentive to acquire information is decreasing in γ. Consumer surplus is maximized at zero public precision and is indifferent over η_I. Total welfare is maximized when no information is collected.*

The logic mirrors the baseline model. Exclusive information has no private strategic use. The only privately valuable part of information is the component disclosed to sellers, which explains why the incentive to collect information falls with product substitutability. From the standpoint of consumers and total welfare, public precision still mainly sharpens state-contingent pricing, while undisclosed information has no independent value.

With a hybrid platform, the picture is richer.

Proposition 13 (Intensive margin and hybrid organization)**.** *There exists a unique disclosure threshold $\gamma_P \in (0, \gamma_{\max})$ such that the platform fully discloses for $\gamma \leq \gamma_P$ and fully garbles for $\gamma > \gamma_P$. Under disclosure, the platform acquires information if and only if k is sufficiently small. When the platform discloses, it acquires more precise information than when it garbles.*

There exists a unique disclosure threshold $\gamma_D \in (0, \gamma_{\max})$ such that public precision raises consumer surplus if and only if $\gamma > \gamma_D$. Moreover, there exists a unique threshold $\gamma_E \in (\gamma_D, \gamma_{\max})$ such that exclusive precision raises consumer surplus if and only if $\gamma > \gamma_E$. Putting the two margins together, there exists a threshold $\gamma_C \in (\gamma_D, \gamma_{\max})$ such that consumer surplus is maximized by

$$\begin{cases} \textit{No acquisition} & \textit{if } 0 \leq \gamma \leq \gamma_D, \\ \textit{Acquisition and full disclosure} & \textit{if } \gamma_D < \gamma \leq \gamma_C, \\ \textit{Acquisition and full garbling} & \textit{if } \gamma_C < \gamma < \gamma_{\max}, \end{cases}$$

Total welfare is maximized by $(\eta_I, \eta_P) = (0, 0)$*. Conditional on a positive stock of data, there exist two thresholds* γ_W^- *and* γ_W^+*, with* $0 < \gamma_W^- < \gamma_W^+ < \gamma_{\max}$*, such that secrecy dominates disclosure for* $\gamma \in (\gamma_W^-, \gamma_W^+)$*, while disclosure is optimal otherwise.*

The economic forces are related to those discussed in the baseline model, but the intensive-margin extension makes the disclosure trade-off considerably richer. In the baseline model, the hybrid platform always chooses full garbling: once information has been acquired, disclosure only weakens the platform's informational advantage over the third-party seller. Here, instead, disclosure may become privately valuable when products are sufficiently differentiated. In that region, making information public improves the independent seller's response to demand conditions, and the platform can internalize part of the resulting gain through its commission revenue. As products become closer substitutes, however, the balance shifts. Disclosure then mainly intensifies price competition and reduces the platform's ability to steer demand toward its integrated arm, so secrecy becomes privately optimal. The threshold γ_P captures exactly this switch between the coordination value of disclosure and the strategic value of secrecy.

The consumer-surplus logic is best understood by separating the public-information margin from the exclusive-information margin. Public precision improves the third-party seller's response to market conditions and, therefore, strengthens the competitive discipline faced by the platform. When products are sufficiently differentiated, this effect is weak and the main consequence of disclosure is sharper state-contingent pricing, which hurts consumers. Once substitutability becomes sufficiently high, however, better-informed rival pricing constrains the platform more effectively, so public precision begins to raise consumer surplus. This is why the effect of disclosure changes sign at γ_D. Exclusive precision works through a different channel. It improves the platform's private adaptation to demand conditions, including its choice of common effort, but it also strengthens the platform's private steering motive. When products are not very substitutable, the steering effect dominates and exclusive information hurts consumers. When products are sufficiently close substitutes, however, the platform's better-informed adjustment of price and effort becomes valuable enough that exclusive precision raises consumer surplus. This explains the higher threshold $\gamma_E > \gamma_D$.

Putting the two margins together yields the three-region consumer-surplus ranking. For low γ, both public and exclusive precision are harmful, so consumers prefer no acquisition. For intermediate values of γ, public precision is beneficial whereas exclusive precision is still harmful, so consumers prefer acquisition with full disclosure, which preserves the beneficial public component while eliminating the privately held informational advantage. For high γ, both margins become beneficial, but exclusive precision becomes more valuable than public precision. In that region, consumers prefer acquisition with full garbling, because keeping information private allows the platform to adjust more effectively without simultaneously strengthening the rival's price responsiveness too much. The threshold γ_C therefore marks the point at which the preferred consumer policy shifts from transparent acquisition to opaque acquisition.

Total welfare maximization is as in the baseline model and is maximized at $(\eta_P, \eta_I) = (0, 0)$ because information acquisition is costly and the gains from better state-contingent adjustment are not large enough to offset those costs. Conditional on a positive stock of data, however, the welfare comparison between disclosure and secrecy is no longer monotone. For intermediate values of γ, secrecy dominates because disclosure mainly amplifies pricing distortions without generating sufficient allocative gains. By contrast, when products are either sufficiently differentiated or sufficiently close substitutes, disclosure becomes optimal because the informational gains from better coordinated market responses outweigh the additional distortions. The two thresholds $\left(\gamma_W^-, \gamma_W^+\right)$ therefore identify the intermediate region in which secrecy is second-best optimal, even though full transparency is preferred outside that range.

5.4 Seller participation and marketplace variety

What if the third-party seller must incur a fixed participation cost to remain on the platform? This feature adds an extensive margin to the information problem. Under pure intermediation, the platform already discloses what it learns, so the participation margin mainly reinforces the platform's own tendency to share information: more public precision raises the seller's expected profitability and therefore supports entry and variety. Relative to the platform's private choice, consumer surplus and total welfare then become more favorable to transparency, because information affects not only price responsiveness but also the number of active sellers.

Under hybrid organization the participation channel introduces an additional welfare trade-off. Opacity does not only protect the platform's informational advantage; it also reduces the rival seller's expected profitability and can deter entry. Yet, when variety is valued by consumers and by total welfare, both welfare standards tilt further toward transparency than the platform's private choice. Relative to the baseline model, participation costs therefore strengthen the case for disclosure duties and can make them desirable.

5.5 Seller-specific services

Suppose the platform can direct some service effort specifically to its own retail arm rather than only through a marketplace-wide public good. This scenario is most naturally tied to hybrid organizations. Exclusive information now becomes useful in two distinct ways: it can refine the platform's own price and can also make targeted services more effective. Relative to the baseline, the platform's private value of hidden information is therefore even larger.

Because the extra service dimension introduces efficiencies, the welfare comparison becomes more complex. To the extent that targeted services create significant efficiency gains, consumer surplus and total welfare may value internally held precision more than in the benchmark model. At the same time, the platform's private bias toward secrecy becomes stronger because targeted services are another channel through which hidden information can be used to favor the integrated seller. Thus the basic policy ranking remains: hybrid organization raises both the private value of data acquisition and the private value of secrecy, but the efficient policy response depends on whether the regulator can act on collection, disclosure, or the downstream use of data.

5.6 Double marginalization

A final remark concerns double marginalization. In the baseline, the platform is compensated through revenue sharing and therefore does not impose an additional upstream markup on top of sellers' retail prices. Suppose instead that the third-party seller faced a wholesale fee or a per-unit transaction charge, so that the standard vertical-pricing distortion arose. Under full market coverage, however, this does not alter the welfare logic developed in the paper. The reason is that aggregate demand is fixed: all consumers buy in equilibrium, and prices only affect the division of demand between sellers. A higher retail price by the independent seller distorts market shares, but it does not exclude consumers from the market. As a result, double marginalization does not generate an additional contraction in total output, and the welfare effects of information disclosure are qualitatively unchanged.

In particular, under pure intermediation the platform still has no business-stealing motive to keep information private, so its disclosure incentive is unaffected. Under hybrid organization, secrecy still matters because it distorts the rival seller's pricing and weakens the competitive constraint faced by the integrated arm. But under full coverage this effect operates only through the allocation of demand across sellers, not through the size of the market. For that reason, the basic welfare trade-off remains the same as in the baseline: disclosure may still raise or reduce welfare depending on whether the allocative gains from better demand adjustment outweigh the stronger extraction of consumer surplus, but double marginalization does not add a new output margin.

By contrast, the analysis changes once aggregate demand becomes price-responsive. This happens, for example, in the uncovered-market extension or when one introduces an intensive margin. In those environments, higher prices reduce total consumption rather than merely reallocate demand across sellers. Double marginalization then creates the standard under-consumption distortion, and better information becomes more valuable from a welfare perspective because it improves price and service decisions in a setting where output is already inefficiently low. In such cases, disclosure can mitigate not only informational distortions but also part of the welfare loss associated with excessive pricing. This is why transparency requirements become more attractive once the model allows prices to affect aggregate demand.

6 Conclusion

The paper studies information management under two platform business models: pure intermediation and hybrid organization. Its central claim is that data policy has two interconnected dimensions —

how much information the platform collects and how much of that information it discloses — and that the distortion on each dimension depends on the platform's organizational form. In a pure-intermediation marketplace, the platform discloses all the information it chooses to acquire, because exclusive precision has no private downstream value. In a hybrid marketplace, by contrast, the same information serves two purposes: it improves common platform services and it can be used to shift downstream competition in favor of the platform's own retail arm. As a result, the platform values private information more than public information and has an incentive to garble as much as possible.

These results have broader implications for information design. First, information policies are inherently multi-dimensional: data acquisition, processing, and disclosure are distinct but interdependent choices. Second, the value of information depends on the organizational environment in which it is used. Hybrid platforms face different incentives than pure intermediaries because information can be deployed both to improve coordination and to shape competition. Finally, information design interacts with real decisions — in this case, service provision — so that the optimal use of information depends on how it affects both allocative efficiency and strategic behavior.

The distinction between these information margins matters also directly for policy. The DMA and related data-governance initiatives implicitly address three separate questions: whether gatekeepers should be allowed to collect marketplace data, whether they should be allowed to use those data for their own downstream activities, and whether some of the resulting information should instead be shared with business users. The model directly studies the first and third questions and helps interpret the second. It shows that these instruments are not equivalent. When regulation can act directly on data collection, the first-best benchmark points to limits on acquisition itself. When such controls are infeasible, however, or once data have already been collected, disclosure mandates become second-best remedies. Their bite is strongest in hybrid platforms, precisely because hybrid organization creates a private use for exclusive information that is absent under pure intermediation. The broader implication is that platform regulation should address information management through the joint lens of business model and market environment. Rules that are innocuous for a pure intermediary can materially affect conduct in a hybrid ecosystem. And even within the same hybrid platform, the appropriate remedy may differ across market settings: what is justified in thick covered markets need not coincide with what is desirable in thinner environments where information affects participation as well as allocation.

Appendix: proofs and supporting derivations

This appendix reports all proofs and supporting calculations in the order in which the corresponding objects appear in the text. Throughout, I write

$$\mu_P(s_P) \equiv \mathrm{E}\left[\sigma\theta \mid s_P\right] = \sigma\sqrt{\eta_P}\, s_P, \qquad \mu_I(s_I) \equiv \mathrm{E}\left[\sigma\theta \mid s_I\right] = \sigma\sqrt{\eta_I}\, s_I.$$

The key identities used repeatedly below are

$$\mathrm{E}\left[\mu_P(s_P)^2\right] = \eta_P\sigma^2, \qquad \mathrm{E}\left[\mu_I(s_I)^2\right] = \eta_I\sigma^2,$$

$$\mathrm{E}\left[\mu_I(s_I) \mid s_P\right] = \mu_P(s_P), \qquad \mathrm{E}\left[(\mu_I(s_I) - \mu_P(s_P))^2\right] = (\eta_I - \eta_P)\sigma^2,$$

and

$$\mathrm{E}\left[\mathrm{Var}\left(\sigma\theta \mid s_I\right)\right] = (1-\eta_I)\sigma^2.$$

These equalities let the whole analysis be written directly in terms of the two signals (s_I, s_P).

Proof of Proposition 1. Under pure intermediation, after observing s_P the platform chooses service effort to maximize

$$\tau q_1(p_1, p_2 \mid s_P)(p_1 - (c-a)) + \tau q_2(p_1, p_2 \mid s_P)(p_2 - (c-a)) - \frac{a^2}{2b},$$

where, assuming full marker coverage,

$$q_1(p_1, p_2 \mid s_P) = \frac{1}{2} + \frac{\mu_P(s_P)}{t} - \frac{p_1 - p_2}{2t}, \qquad q_2(p_1, p_2 \mid s_P) = 1 - q_1(p_1, p_2 \mid s_P).$$

The first-order condition with respect to a is simply

$$\tau - \frac{a}{b} = 0,$$

which yields

$$a^{\mathrm{PI}} = b\tau.$$

Given a^{PI}, seller $i \in \{1, 2\}$ solves

$$\max_{p_i \in \mathbb{R}} q_i(p_i, p_j \mid s_P)\big(p_i - (c - b\tau)\big).$$

The first-order conditions are

$$q_1(p_1, p_2 \mid s_P) - \frac{p_1 - c + b\tau}{2t} = 0, \qquad q_2(p_1, p_2 \mid s_P) - \frac{p_2 - c + b\tau}{2t} = 0.$$

Substituting the Hotelling demand system, we have the linear system

$$2p_1 = t + 2\mu_P(s_P) + p_2 + c - b\tau, \qquad 2p_2 = t - 2\mu_P(s_P) + p_1 + c - b\tau,$$

whose solution yields

$$p_1^{\text{PI}}(s_P) = c - b\tau + t + \frac{2}{3}\mu_P(s_P), \qquad p_2^{\text{PI}}(s_P) = c - b\tau + t - \frac{2}{3}\mu_P(s_P).$$

Since neither a^{PI} nor the sellers' pricing system depends on s_I once s_P is fixed, the platform's additional internal signal has no payoff relevance in this benchmark. ■

Expected platform profit under pure intermediation. Using the equilibrium prices, operating profits are

$$\pi_1^{\text{PI}}(s_P) = \left(t + \frac{2\mu_P(s_P)}{3}\right)\left(\frac{1}{2} + \frac{\mu_P(s_P)}{3t}\right),$$

$$\pi_2^{\text{PI}}(s_P) = \left(t - \frac{2\mu_P(s_P)}{3}\right)\left(\frac{1}{2} - \frac{\mu_P(s_P)}{3t}\right).$$

Hence

$$\pi_1^{\text{PI}}(s_P) + \pi_2^{\text{PI}}(s_P) = t + \frac{4\mu_P(s_P)^2}{9t}.$$

The platform therefore obtains

$$\Pi_P^{\text{PI}}(s_P) = \tau\left(t + \frac{4\mu_P(s_P)^2}{9t}\right) - \frac{b\tau^2}{2}.$$

Taking expectations and using $\mathrm{E}\left[\mu_P(s_P)^2\right] = \eta_P\sigma^2$ yields

$$\Pi_P^{\text{PI}}(\eta_I, \eta_P) = \Pi_0^{\text{PI}} + \frac{4\tau\sigma^2}{9t}\eta_P - K(\eta_I),$$

where $\Pi_0^{\text{PI}} \equiv \tau t - b\tau^2/2$.

Proof of Proposition 2. Fix η_I. Equation (1) is strictly increasing in η_P, and feasibility requires $\eta_P \leq \eta_I$. Hence any optimum must satisfy

$$\eta_P^{\text{PI}} = \eta_I^{\text{PI}}.$$

Substituting $\eta_P = \eta_I \equiv \eta$ into (1) gives the one-dimensional problem

$$\max_{\eta\in[0,1)} \left\{\frac{4\tau\sigma^2}{9t}\eta - k\frac{\eta}{1-\eta}\right\}.$$

The derivative is

$$\frac{4\tau\sigma^2}{9t} - \frac{k}{(1-\eta)^2}.$$

An interior solution therefore exists if and only if

$$k < \frac{4\tau\sigma^2}{9t} \equiv \bar{k}^{\mathrm{PI}}(t,\tau,\sigma).$$

Whenever this condition holds, the first-order condition yields

$$1-\eta = \sqrt{\frac{k}{\bar{k}^{\mathrm{PI}}(t,\tau,\sigma)}},$$

which proves the formula for $\eta_I^{\mathrm{PI}} = \eta_P^{\mathrm{PI}}$. ■

Expected consumer surplus and total welfare under pure intermediation. Given equilibrium prices, the indifferent consumer is located at

$$x^{\mathrm{PI}}(\theta, s_P) = \frac{1}{2} + \frac{\theta}{t} - \frac{2\mu_P(s_P)}{3t}.$$

Conditional consumer surplus is therefore

$$\begin{aligned} CS^{\mathrm{PI}}(\theta, s_P) &= \int_0^{x^{\mathrm{PI}}(\theta,s_P)} \left(v + \theta - p_1^{\mathrm{PI}}(s_P) - tz^2\right) dz \\ &\quad + \int_{x^{\mathrm{PI}}(\theta,s_P)}^1 \left(v - \theta - p_2^{\mathrm{PI}}(s_P) - t(1-z)^2\right) dz \\ &= CS_0^{\mathrm{PI}} + \frac{\theta^2}{t} - \frac{4\mu_P(s_P)\theta}{3t} + \frac{4\mu_P(s_P)^2}{9t}, \end{aligned}$$

where $CS_0^{\mathrm{PI}} \equiv v - c + b\tau - \frac{13t}{12}$. Taking expectations conditional on s_P and using $\mathrm{E}\left[\theta \mid s_P\right] = \mu_P(s_P)$ and $\mathrm{E}\left[\theta^2 \mid s_P\right] = \mu_P(s_P)^2 + (1-\eta_P)\sigma^2$ yields

$$\mathrm{E}\left[CS^{\mathrm{PI}}(\theta, s_P) \mid s_P\right] = CS_0^{\mathrm{PI}} + \frac{(1-\eta_P)\sigma^2}{t} + \frac{\mu_P(s_P)^2}{9t}.$$

Finally, using $\mathrm{E}\left[\mu_P(s_P)^2\right] = \eta_P\sigma^2$ yields

$$CS^{\mathrm{PI}}(\eta_I, \eta_P) = CS_0^{\mathrm{PI}} + \left(1 - \frac{8}{9}\eta_P\right)\frac{\sigma^2}{t}.$$

Total welfare equals consumer surplus plus total operating profits, net of service and information costs. From the equilibrium calculations,

$$\pi_1^{\mathrm{PI}}(s_P) + \pi_2^{\mathrm{PI}}(s_P) = t + \frac{4\mu_P(s_P)^2}{9t}, \qquad \psi(a^{\mathrm{PI}}) = \frac{b\tau^2}{2}.$$

Hence

$$W^{\mathrm{PI}}(\eta_I, \eta_P) = W_0^{\mathrm{PI}} + \left(1 - \frac{4}{9}\eta_P\right)\frac{\sigma^2}{t} - K(\eta_I),$$

where $W_0^{\mathrm{PI}} \equiv CS_0^{\mathrm{PI}} + t - b\tau^2/2$.

Proof of Corollary 1. The consumer-surplus formula (2) is strictly decreasing in η_P and does not depend on η_I. Hence consumers choose

$$\eta_P^{\mathrm{PI},CS} = 0$$

and are indifferent over η_I.

The welfare formula (3) is strictly decreasing in η_P. Conditional on $\eta_P = 0$, it is also decreasing in η_I because $K(\eta_I)$ is increasing. Therefore welfare chooses

$$\eta_P^{\mathrm{PI},W} = 0, \qquad \eta_I^{\mathrm{PI},W} = 0.$$

The platform instead chooses $\eta_P^{\mathrm{PI}} = \eta_I^{\mathrm{PI}} > 0$ whenever $k < \bar{k}^{\mathrm{PI}}(t,\tau)$. The over-disclosure and over-acquisition claims follow immediately. ■

Proof of Proposition 3. I proceed in four steps.

Step 1. Linear strategy conjecture. Conjecture linear strategies of the form

$$p_1^{\mathrm{H}}(s_I, s_P) = \alpha_1 + \beta_1\mu_P(s_P) + \gamma_1\big(\mu_I(s_I) - \mu_P(s_P)\big),$$

$$p_2^{\mathrm{H}}(s_P) = \alpha_2 + \beta_2\mu_P(s_P), \qquad a^{\mathrm{H}}(s_I, s_P) = \alpha_a + \beta_a\mu_P(s_P) + \gamma_a\big(\mu_I(s_I) - \mu_P(s_P)\big).$$

Because $\mathrm{E}\left[\mu_I(s_I) - \mu_P(s_P) \mid s_P\right] = 0$, the forecasting objects used by seller S_2 are

$$\widetilde{p}_1^{\mathrm{H}}(s_P) = \alpha_1 + \beta_1\mu_P(s_P), \qquad \widetilde{a}^{\mathrm{H}}(s_P) = \alpha_a + \beta_a\mu_P(s_P).$$

Step 2. Seller S_2's problem. Conditional on s_P, seller S_2 chooses p_2 to maximize

$$\hat{\pi}_2(p_2 \mid s_P) = \mathrm{E}\left[q_2(p_1^{\mathrm{H}}(s_I, s_P), p_2 \mid s_I)\big(p_2 - c + a^{\mathrm{H}}(s_I, s_P)\big) \mid s_P\right].$$

Differentiating under the expectation sign gives

$$0 = \mathrm{E}\left[q_2(p_1^{\mathrm{H}}(s_I, s_P), p_2 \mid s_I) - \frac{p_2 - c + a^{\mathrm{H}}(s_I, s_P)}{2t} \mid s_P\right].$$

Using $\mathrm{E}\left[\mu_I(s_I) \mid s_P\right] = \mu_P(s_P)$ and the linear conjecture, this reduces to

$$\frac{1}{2} - \frac{\mu_P(s_P)}{t} - \frac{2p_2 - \widetilde{p}_1^{\mathrm{H}}(s_P) - c + \widetilde{a}^{\mathrm{H}}(s_P)}{2t} = 0.$$

Matching the constant and $\mu_P(s_P)$ coefficients yields

$$2\alpha_2 = t + \alpha_1 + c - \alpha_a, \qquad 2\beta_2 = -2 + \beta_1 - \beta_a. \tag{A1}$$

Step 3. The platform's problem. The platform chooses (p_1, a) to maximize

$$q_1(p_1, p_2 \mid s_I)(p_1 - (c - a)) + \tau q_2(p_1, p_2 \mid s_I)(p_2 - (c - a)) - \frac{a^2}{2b}.$$

Substituting the linear conjecture into the first-order conditions (4)-(5) and collecting coefficients yields

$$2\alpha_1 = (1 + \tau)\alpha_2 + (1 - \tau)c - (1 - \tau)\alpha_a + t, \tag{A2}$$

$$2\beta_1 = (1 + \tau)\beta_2 + 2 - (1 - \tau)\beta_a, \tag{A3}$$

$$2\gamma_1 = 2 - (1 - \tau)\gamma_a, \tag{A4}$$

and

$$\alpha_a = b\left[\frac{1 + \tau}{2} + \frac{1 - \tau}{2t}(\alpha_2 - \alpha_1)\right], \tag{A5}$$

$$\beta_a = b\left[\frac{1 - \tau}{t} + \frac{1 - \tau}{2t}(\beta_2 - \beta_1)\right], \tag{A6}$$

$$\gamma_a = b\left[\frac{1 - \tau}{t} - \frac{1 - \tau}{2t}\gamma_1\right]. \tag{A7}$$

Step 4. Solving the coefficient system. The signal-gap block, (A4)–(A7), yields

$$\gamma_1 = \frac{2\big(2t - b(1 - \tau)^2\big)}{4t - b(1 - \tau)^2}, \qquad \gamma_a = \frac{2b(1 - \tau)}{4t - b(1 - \tau)^2}.$$

The public-signal block, (A1), (A3), and (A6), yields

$$\beta_1 = \frac{(1 - \tau)(2t - b)}{t(3 - \tau)}, \qquad \beta_2 = -\frac{2t + b(1 - \tau)}{t(3 - \tau)}, \qquad \beta_a = \frac{b(1 - \tau)}{t(3 - \tau)}.$$

Finally, the constant block, (A1), (A2), and (A5), yields

$$\alpha_1 = c + \frac{(3 + \tau)\left(t - \frac{b}{2}\right)}{3 - \tau}, \qquad \alpha_2 = c + \frac{3t - \frac{b}{2}(3 + \tau)}{3 - \tau}, \qquad \alpha_a = \frac{b(3 + \tau)}{2(3 - \tau)}.$$

Substituting these coefficients back into the conjectured strategies gives the equilibrium stated in the proposition. ■

Expected platform profit under hybrid organization. Using the equilibrium strategies, realized quantities are

$$q_1^{\mathrm{H}}(s_I, s_P) = \frac{3 - 2\tau}{2(3 - \tau)} + \frac{\mu_P(s_P)}{t(3 - \tau)} + \frac{2\big(\mu_I(s_I) - \mu_P(s_P)\big)}{4t - b(1 - \tau)^2},$$

$$q_2^{\mathrm{H}}(s_I, s_P) = 1 - q_1^{\mathrm{H}}(s_I, s_P).$$

Substituting these quantities together with the equilibrium prices and service effort into the platform's objective gives

$$\Pi^{\mathrm{H}}(s_I,s_P)=q_1^{\mathrm{H}}(s_I,s_P)\big(p_1^{\mathrm{H}}(s_I,s_P)-\big(c-a^{\mathrm{H}}(s_I,s_P)\big)\big)$$
$$+\tau q_2^{\mathrm{H}}(s_I,s_P)\big(p_2^{\mathrm{H}}(s_P)-\big(c-a^{\mathrm{H}}(s_I,s_P)\big)\big)-\frac{a^{\mathrm{H}}(s_I,s_P)^2}{2b}.$$

After simplification,

$$\Pi^{\mathrm{H}}(s_I,s_P)=\Pi_0^{\mathrm{H}}+\frac{4t-b(1-\tau)^2}{2t^2(3-\tau)^2}\mu_P(s_P)^2+\frac{2}{4t-b(1-\tau)^2}\big(\mu_I(s_I)-\mu_P(s_P)\big)^2+R_\Pi(s_I,s_P),$$

where $R_\Pi(s_I,s_P)$ is linear or bilinear in the two signal-based objects and therefore has zero expectation. Taking expectations and using

$$\mathrm{E}\left[\mu_P(s_P)^2\right]=\eta_P\sigma^2,\qquad \mathrm{E}\left[\big(\mu_I(s_I)-\mu_P(s_P)\big)^2\right]=(\eta_I-\eta_P)\sigma^2,$$

yields equation (7).

Proof of Proposition 4. For fixed η_I, equation (7) is affine in η_P. The coefficient on η_P can be written as

$$\frac{4t-b(1-\tau)^2}{2t^2(3-\tau)^2}-\frac{2}{4t-b(1-\tau)^2}=-\frac{\big(b(1-\tau)^2-2t(1-\tau)\big)^2}{2t^2(3-\tau)^2\big(4t-b(1-\tau)^2\big)}\le 0,$$

with strict inequality on the admissible set. Hence the platform chooses maximal garbling:

$$\eta_P^{\mathrm{H}}=0.$$

Substituting $\eta_P=0$ into (7) gives

$$\Pi_P^{\mathrm{H}}(\eta_I,0)=\Pi_0^{\mathrm{H}}+\frac{2\sigma^2}{4t-b(1-\tau)^2}\eta_I-k\frac{\eta_I}{1-\eta_I}.$$

An interior solution exists if and only if

$$k<\frac{2\sigma^2}{4t-b(1-\tau)^2}\equiv\bar{k}_P^{\mathrm{H}}(b,t,\tau,\sigma),$$

and then

$$\eta_I^{\mathrm{H}}=1-\sqrt{\frac{k}{\bar{k}_P^{\mathrm{H}}(b,t,\tau,\sigma)}}.$$

The comparative statics follow by direct differentiation. ■

Expected consumer surplus under hybrid organization. Let $x^{\mathrm{H}}(\theta,s_I,s_P)$ denote the indifferent consumer under the hybrid equilibrium. Substituting the equilibrium strategies into the Hotelling expressions and simplifying gives

$$\mathrm{E}\left[CS^{\mathrm{H}}(\theta, s_I, s_P) \mid s_I, s_P\right] = CS_0^{\mathrm{H}} + \frac{\mu_P(s_P)^2}{t(3-\tau)^2} + \frac{4t}{\left(4t - b(1-\tau)^2\right)^2}\left(\mu_I(s_I) - \mu_P(s_P)\right)^2 + \frac{\mathrm{Var}(\theta \mid s_I)}{t} + R_{CS}(s_I, s_P),$$

where $R_{CS}(s_I, s_P)$ has zero expectation. Taking expectations and using the signal identities at the beginning of the appendix yields equation (8):

$$CS^{\mathrm{H}}(\eta_I, \eta_P) = CS_0^{\mathrm{H}} + \left[\frac{1}{t(3-\tau)^2}\eta_P + \frac{4t}{\left(4t - b(1-\tau)^2\right)^2}(\eta_I - \eta_P) + \frac{1-\eta_I}{t}\right]\sigma^2.$$

Proof of Proposition 5. For fixed η_I, consumer surplus is affine in η_P. The coefficient on η_P is

$$\frac{1}{t(3-\tau)^2} - \frac{4t}{\left(4t - b(1-\tau)^2\right)^2} = -\frac{\left(b(1-\tau)^2 - 2t(2-\tau)\right)\left(b(1-\tau)^2 - 2t\tau\right)}{t(3-\tau)^2\left(4t - b(1-\tau)^2\right)^2} < 0$$

on the admissible set, so consumers always choose $\eta_P = 0$.

Conditional on $\eta_P = 0$, the coefficient on η_I in $CS^{\mathrm{H}}(\eta_I, 0)$ is

$$\frac{4t}{\left(4t - b(1-\tau)^2\right)^2} - \frac{1}{t}.$$

This term is positive if and only if $4t - b(1-\tau)^2 < 2t$, namely if and only if

$$b > \frac{2t}{(1-\tau)^2} \equiv b_C^{\mathrm{H}}(t, \tau).$$

Hence consumers choose $\eta_I = 1$ when $b > b_C^{\mathrm{H}}(t, \tau)$ and $\eta_I = 0$ otherwise. The final comparison with the platform's choice follows immediately from Proposition 4. ∎

Expected total welfare under hybrid organization. Using the equilibrium allocation and integrating out the residual uncertainty conditional on (s_I, s_P) yields

$$\mathrm{E}\left[W^{\mathrm{H}}(\theta, s_I, s_P) \mid s_I, s_P\right] = W_0^{\mathrm{H}} + \frac{10t - 4\tau t - b(1-\tau)^2}{2t^2(3-\tau)^2}\mu_P(s_P)^2 + \frac{6\left(2t - b(1-\tau)^2\right)}{\left(4t - b(1-\tau)^2\right)^2}\left(\mu_I(s_I) - \mu_P(s_P)\right)^2 + \frac{\mathrm{Var}(\theta \mid s_I)}{t} + R_W(s_I, s_P),$$

where $R_W(s_I, s_P)$ has zero expectation. Taking expectations gives equation (9):

$$W^{\mathrm{H}}(\eta_I, \eta_P) = W_0^{\mathrm{H}} + \frac{10t - 4\tau t - b(1-\tau)^2}{2t^2(3-\tau)^2}\eta_P\sigma^2 + \left[\frac{1-\eta_I}{t} + \frac{6\big(2t - b(1-\tau)^2\big)}{\big(4t - b(1-\tau)^2\big)^2}(\eta_I - \eta_P)\right]\sigma^2 - K(\eta_I).$$

Proof of Proposition 6. Rewrite (9) as

$$W^{\mathrm{H}}(\eta_I, \eta_P) = W_0^{\mathrm{H}} + \frac{\sigma^2}{t} + \sigma^2\left(\frac{6\big(2t - b(1-\tau)^2\big)}{\big(4t - b(1-\tau)^2\big)^2} - \frac{1}{t}\right)\eta_I$$

$$+\sigma^2\left(\frac{10t - 4\tau t - b(1-\tau)^2}{2t^2(3-\tau)^2} - \frac{6\big(2t - b(1-\tau)^2\big)}{\big(4t - b(1-\tau)^2\big)^2}\right)\eta_P - K(\eta_I).$$

Consider first acquisition. The coefficient on η_I can be written as

$$\frac{-b^2(1-\tau)^4 + 2bt(1-\tau)^2 - 4t^2}{t\big(4t - b(1-\tau)^2\big)^2},$$

which is strictly negative because its numerator equals $-\big(b^2(1-\tau)^4 - 2bt(1-\tau)^2 + 4t^2\big)$ and the term in parentheses is strictly positive. Hence, even before the acquisition cost is subtracted, welfare is decreasing in η_I. Therefore the first-best planner sets

$$\eta_I^{\mathrm{H},W} = 0.$$

Now hold fixed a positive stock of internally acquired data. Welfare prefers disclosure if and only if the coefficient on η_P is positive, that is, if and only if

$$\frac{10t - 4\tau t - b(1-\tau)^2}{2t^2(3-\tau)^2} > \frac{6\big(2t - b(1-\tau)^2\big)}{\big(4t - b(1-\tau)^2\big)^2}.$$

After rearranging, this condition is equivalent to

$$\big(b(1-\tau)^2 - 2t(1-\tau)\big)\left[b(1-\tau)^2 - \left(10 - 3\tau - \sqrt{72 - 48\tau + 9\tau^2}\right)t\right]$$

$$\times\left[b(1-\tau)^2 - \left(10 - 3\tau + \sqrt{72 - 48\tau + 9\tau^2}\right)t\right] < 0.$$

Only the middle root is relevant in the admissible region. Hence welfare prefers disclosure if and only if

$$b > \frac{10 - 3\tau - \sqrt{72 - 48\tau + 9\tau^2}}{(1-\tau)^2}\,t \equiv b_W^{\mathrm{H}}(t, \tau).$$

The threshold is linear in t, so it is increasing in t. A direct differentiation shows that it is also increasing in τ. ■

Proof of Corollary 9. Under a full-disclosure mandate, $\eta_P = \eta_I \equiv \eta$, so (7) becomes

$$\Pi_P^{\mathrm{H},D}(\eta) = \Pi_0^{\mathrm{H}} + \sigma^2 \frac{4t - b(1-\tau)^2}{2t^2(3-\tau)^2}\eta - k\frac{\eta}{1-\eta}.$$

The platform therefore acquires data if and only if

$$k < \frac{(4t - b(1-\tau)^2)\sigma^2}{2t^2(3-\tau)^2} \equiv \bar{k}_D^{\mathrm{H}}(b,t,\tau,\sigma).$$

Finally,

$$\bar{k}_D^{\mathrm{H}}(b,t,\tau,\sigma) < \bar{k}_P^{\mathrm{H}}(b,t,\tau,\sigma) \quad \iff \quad \left(4t - b(1-\tau)^2\right)^2 < 4t^2(3-\tau)^2,$$

which holds because $4t - b(1-\tau)^2 < 2t(3-\tau)$ for every $\tau < 1$. Hence mandatory disclosure weakly reduces acquisition. ■

Proof of Corollary 10. Differentiate the two thresholds directly. Since

$$\bar{k}_P^{\mathrm{H}}(b,t,\tau,\sigma) = \frac{2\sigma^2}{4t - b(1-\tau)^2},$$

one obtains

$$\frac{\partial \bar{k}_P^{\mathrm{H}}(b,t,\tau,\sigma)}{\partial \tau} = -\frac{4b(1-\tau)\sigma^2}{\left(4t - b(1-\tau)^2\right)^2} < 0.$$

Likewise,

$$\bar{k}_D^{\mathrm{H}}(b,t,\tau,\sigma) = \frac{(4t - b(1-\tau)^2)\sigma^2}{2t^2(3-\tau)^2},$$

so

$$\frac{\partial \bar{k}_D^{\mathrm{H}}(b,t,\tau,\sigma)}{\partial \tau} = \frac{2\left(b(1-\tau) + 2t\right)\sigma^2}{t^2(3-\tau)^3} > 0.$$

The sign conclusions follow immediately. ■

Equilibrium characterization under Singh-Vives specification. Let

$$\mu_P(s_P) \equiv \mathrm{E}\left[\sigma\theta \mid s_P\right], \qquad \mu_I(s_I) \equiv \mathrm{E}\left[\sigma\theta \mid s_I\right], \qquad g(s_I, s_P) \equiv \mu_I(s_I) - \mu_P(s_P),$$

Consider first the pure intermediation mode. Both sellers observe only s_P. Seller $i \in \{1,2\}$ solves

$$\max_{p_i \in \mathbb{R}_+} \; q_i\left(\mu_I(s_I), p_i, p_j\right)(p_i - (c-a)),$$

where

$$q_1\left(\mu_I(s_I), p_1, p_2\right) = \frac{(1+\mu_P(s_P))(1-\gamma) - p_1 + \gamma p_2}{1-\gamma^2}, \qquad q_2\left(\mu_I(s_I), p_2, p_1\right) = \frac{(1-\mu_P(s_P))(1-\gamma) - p_2 + \gamma p_1}{1-\gamma^2}.$$

The first-order conditions are

$$2p_1 = (1-\gamma)(1+\mu_P(s_P)) + \gamma p_2 + c - a,$$

$$2p_2 = (1-\gamma)(1-\mu_P(s_P)) + \gamma p_1 + c - a.$$

Solving yields

$$p_1^{\mathrm{PI}}(s_P) = \frac{1-\gamma+c-a}{2-\gamma} + \frac{1-\gamma}{2+\gamma}\mu_P(s_P),$$

$$p_2^{\mathrm{PI}}(s_P) = \frac{1-\gamma+c-a}{2-\gamma} - \frac{1-\gamma}{2+\gamma}\mu_P(s_P).$$

Since

$$q_1\left(\cdot\right) + q_2\left(\cdot\right) = \frac{2-p_1-p_2}{1+\gamma},$$

the platform chooses a to solve

$$\max_{a\in\mathbb{R}} \ \frac{1}{2}\left[(p_1-(c-a))q_1\left(\cdot\right) + (p_2-(c-a))q_2\left(\cdot\right)\right] - \frac{a^2}{2},$$

whose first-order condition is

$$a = \frac{q_1\left(\cdot\right)+q_2\left(\cdot\right)}{2} = \frac{2-p_1-p_2}{2(1+\gamma)}.$$

Substituting the price system into this equation gives

$$a^{\mathrm{PI}} = \frac{1-c}{1+\gamma\left(1-\gamma\right)}.$$

Replacing this expression in the pricing formulas yields

$$p_1^{\mathrm{PI}}(s_P) = p_0^{\mathrm{PI}} + \frac{1-\gamma}{2+\gamma}\mu_P(s_P), \qquad p_2^{\mathrm{PI}}(s_P) = p_0^{\mathrm{PI}} - \frac{1-\gamma}{2+\gamma}\mu_P(s_P),$$

where

$$p_0^{\mathrm{PI}} = \frac{c(1+\gamma)-\gamma^2}{1+\gamma\left(1-\gamma\right)}.$$

Consider now a hybrid platform. Let

$$\Omega(\gamma) \equiv 4+4\gamma-9\gamma^2, \qquad \Delta(\gamma) \equiv 2+6\gamma-2\gamma^2-3\gamma^3,$$

and assume $\Omega(\gamma) > 0$. Conjecture linear strategies of the form

$$p_1^{\mathrm{H}}(s_I, s_P) = p_{10}^{\mathrm{H}} + p_{1P}^{\mathrm{H}}\mu_P(s_P) + p_{1I}^{\mathrm{H}}g(s_I, s_P),$$

$$p_2^{\mathrm{H}}(s_P) = p_{20}^{\mathrm{H}} + p_{2P}^{\mathrm{H}}\mu_P(s_P), \qquad a^{\mathrm{H}}(s_I, s_P) = a_0^{\mathrm{H}} + a_P^{\mathrm{H}}\mu_P(s_P) + a_I^{\mathrm{H}}g(s_I, s_P).$$

Conditional on s_P, seller S_2 forecasts

$$p_1^{\mathrm{H}}(s_P) = p_{10}^{\mathrm{H}} + p_{1P}^{\mathrm{H}}\mu_P(s_P), \qquad a^{\mathrm{H}}(s_P) = a_0^{\mathrm{H}} + a_P^{\mathrm{H}}\mu_P(s_P),$$

because $\mathrm{E}[g(s_I, s_P) \mid s_P] = 0$. Seller S_2 solves

$$\max_{p_2 \in \mathbb{R}_+} \left(p_2 - \left(c - a^{\mathrm{H}}(s_P)\right)\right) \mathrm{E}[q_2(\cdot) \mid s_P],$$

which yields

$$2p_2 = (1-\gamma)(1 - \mu_P(s_P)) + \gamma p_1^{\mathrm{H}}(s_P) + c - a^{\mathrm{H}}(s_P).$$

Given p_2, the platform chooses (p_1, a) to maximize

$$(p_1 - (c - a))q_1(\cdot) + \frac{1}{2}(p_2 - (c - a))q_2(\cdot) - \frac{a^2}{2}.$$

The first-order condition with respect to p_1 is

$$2p_1 = (1-\gamma)(1 + \mu_I(s_I)) + \frac{3\gamma}{2}p_2 + \left(1 - \frac{\gamma}{2}\right)(c - a),$$

while the first-order condition with respect to a is

$$a = q_1(\cdot) + \frac{1}{2}q_2(\cdot).$$

Using the demand system, this can be rewritten as

$$(1-\gamma^2)a = \frac{3}{2}(1-\gamma) + \frac{1}{2}(1-\gamma)\mu_I(s_I) + \left(\frac{\gamma}{2} - 1\right)p_1 + \left(\gamma - \frac{1}{2}\right)p_2.$$

Matching constants, public-signal terms, and gap terms yields

$$2p_{20}^{\mathrm{H}} = 1 - \gamma + \gamma p_{10}^{\mathrm{H}} + c - a_0^{\mathrm{H}}, \qquad 2p_{2P}^{\mathrm{H}} = -(1-\gamma) + \gamma p_{1P}^{\mathrm{H}} - a_P^{\mathrm{H}},$$

$$2p_{10}^{\mathrm{H}} = 1 - \gamma + \frac{3\gamma}{2}p_{20}^{\mathrm{H}} + \left(1 - \frac{\gamma}{2}\right)(c - a_0^{\mathrm{H}}),$$

$$2p_{1P}^{\mathrm{H}} = 1 - \gamma + \frac{3\gamma}{2}p_{2P}^{\mathrm{H}} - \left(1 - \frac{\gamma}{2}\right)a_P^{\mathrm{H}}, \qquad 2p_{1I}^{\mathrm{H}} = 1 - \gamma - \left(1 - \frac{\gamma}{2}\right)a_I^{\mathrm{H}},$$

$$(1-\gamma^2)a_0^{\mathrm{H}} = \frac{3}{2}(1-\gamma) + \left(\frac{\gamma}{2}-1\right)p_{10}^{\mathrm{H}} + \left(\gamma - \frac{1}{2}\right)p_{20}^{\mathrm{H}},$$

$$(1-\gamma^2)a_P^{\mathrm{H}} = \frac{1}{2}(1-\gamma) + \left(\frac{\gamma}{2}-1\right)p_{1P}^{\mathrm{H}} + \left(\gamma - \frac{1}{2}\right)p_{2P}^{\mathrm{H}},$$

$$(1-\gamma^2)a_I^{\mathrm{H}} = \frac{1}{2}(1-\gamma) + \left(\frac{\gamma}{2}-1\right)p_{1I}^{\mathrm{H}}.$$

The seller's objective is strictly concave in p_2. The platform's Hessian with respect to (p_1, a) is

$$H = \begin{pmatrix} -\frac{2}{1-\gamma^2} & -\frac{2-\gamma}{2(1-\gamma^2)} \\ -\frac{2-\gamma}{2(1-\gamma^2)} & -1 \end{pmatrix},$$

whose determinant is

$$\det(H) = \frac{4+4\gamma-9\gamma^2}{4(1-\gamma^2)^2}.$$

Hence $\Omega(\gamma) > 0$ guarantees a unique linear equilibrium.

Solving the coefficient system yields

$$p_{10}^{\mathrm{H}} = \frac{c(\gamma^2+5\gamma+4) - 3\gamma^3 - 3\gamma^2 + \gamma - 2}{\Delta(\gamma)},$$

$$p_{20}^{\mathrm{H}} = \frac{-c\gamma^3 + c\gamma^2 + 6c\gamma + 4c - 2\gamma^3 - 3\gamma^2 - 2}{\Delta(\gamma)},$$

$$a_0^{\mathrm{H}} = \frac{(1-c)(6+2\gamma-\gamma^2)}{\Delta(\gamma)},$$

$$p_{1P}^{\mathrm{H}} = \frac{\gamma(3\gamma^2-5\gamma+1)}{\Delta(\gamma)}, \qquad p_{2P}^{\mathrm{H}} = -\frac{2\gamma^3-5\gamma^2+2\gamma+2}{\Delta(\gamma)}, \qquad a_P^{\mathrm{H}} = \frac{2-\gamma^2}{\Delta(\gamma)},$$

and

$$p_{1I}^{\mathrm{H}} = \frac{(1-\gamma)(2+\gamma-4\gamma^2)}{\Omega(\gamma)}, \qquad a_I^{\mathrm{H}} = \frac{2\gamma(1-\gamma)}{\Omega(\gamma)},$$

which completes the equilibrium characterization.

Proof of Proposition 12. Under pure intermediation, both sellers observe only the public signal. The equilibrium is linear, with symmetric pricing slopes and a constant marketplace effort. Direct substitution yields the reduced-form ex ante payoff and welfare expressions

$$\Pi^{\mathrm{PI}}(\eta_I, \eta_P) = \Pi_0^{\mathrm{PI}} + \underbrace{\frac{1-\gamma}{(1+\gamma)(2+\gamma)^2}}_{\equiv \bar{k}^{\mathrm{PI}}(\gamma)} \eta_P \sigma^2 - K(\eta_I),$$

$$CS^{\mathrm{PI}}(\eta_I, \eta_P) = CS_0^{\mathrm{PI}} - \frac{(1-\gamma)(3+\gamma)}{(1+\gamma)(2+\gamma)^2}\,\eta_P \sigma^2,$$

and

$$W^{\mathrm{PI}}(\eta_I, \eta_P) = W_0^{\mathrm{PI}} - \frac{1-\gamma}{(2+\gamma)^2}\,\eta_P \sigma^2 - K(\eta_I).$$

The disclosure result follows immediately because $\Pi^{\mathrm{PI}}(\cdot)$ depends on η_P but not on $\eta_I - \eta_P$. Hence any acquired precision is fully disclosed. Given $\eta_P = \eta_I$, the platform solves

$$\max_{\eta_I \in [0,1)} \left\{ \bar{k}^{\mathrm{PI}}(\gamma)\eta_I \sigma^2 - K(\eta_I) \right\},$$

which yields immediately the result. Since

$$\frac{d}{d\gamma}\bar{k}^{\mathrm{PI}}(\gamma) = -\frac{\gamma^2 + 2\gamma + 4}{(1+\gamma)^2(2+\gamma)^3} < 0,$$

the incentive to acquire information decreases with product substitutability. The coefficient on η_P in $CS^{\mathrm{PI}}(\cdot)$ is strictly negative, while undisclosed precision does not enter at all. The same is true for total welfare, which additionally bears the information cost $K(\eta_I)$. Hence consumer surplus is maximized at zero public precision, whereas total welfare is maximized by no acquisition. ■

Proof of Proposition 13. Direct substitution of equilibrium values into the main objectives yields the reduced-form ex ante payoff and welfare expressionse

$$\Pi^{\mathrm{H}}(\eta_I, \eta_P) = \Pi_0^{\mathrm{H}} + A_P(\gamma)\eta_P \sigma^2 + B_P(\gamma)(\eta_I - \eta_P)\sigma^2 - K(\eta_I),$$

$$CS^{\mathrm{H}}(\eta_I, \eta_P) = CS_0^{\mathrm{H}} + A_{CS}(\gamma)\eta_P \sigma^2 + B_{CS}(\gamma)(\eta_I - \eta_P)\sigma^2,$$

and

$$W^{\mathrm{H}}(\eta_I, \eta_P) = W_0^{\mathrm{H}} + A_W(\gamma)\eta_P \sigma^2 + B_W(\gamma)(\eta_I - \eta_P)\sigma^2 - K(\eta_I),$$

where

$$A_P(\gamma) \equiv \frac{4 + 12\gamma - 18\gamma^2 - 40\gamma^3 + 51\gamma^4 - 5\gamma^5 - 6\gamma^6}{2(1+\gamma)(2 + 6\gamma - 2\gamma^2 - 3\gamma^3)^2},$$

$$B_P(\gamma) \equiv \frac{(1-\gamma)^2(1+2\gamma)}{(1+\gamma)(4 + 4\gamma - 9\gamma^2)},$$

$$A_{CS}(\gamma) \equiv \frac{18\gamma^6 - 15\gamma^5 - 73\gamma^4 + 87\gamma^3 + 13\gamma^2 - 24\gamma - 4}{2(1+\gamma)(2 + 6\gamma - 2\gamma^2 - 3\gamma^3)^2},$$

$$B_{CS}(\gamma) \equiv \frac{(\gamma - 1)(4\gamma^2 - \gamma - 2)(14\gamma^2 - 7\gamma - 6)}{2(1+\gamma)(4 + 4\gamma - 9\gamma^2)^2},$$

$$A_W(\gamma) \equiv \frac{\gamma(\gamma - 1)(12\gamma^4 - 12\gamma^3 - 22\gamma^2 + 13\gamma + 12)}{2(1+\gamma)(2 + 6\gamma - 2\gamma^2 - 3\gamma^3)^2},$$

and

$$B_W(\gamma) \equiv \frac{(\gamma - 1)(28\gamma^4 - 28\gamma^3 - 15\gamma^2 + 12\gamma + 4)}{2(1+\gamma)(4 + 4\gamma - 9\gamma^2)^2}.$$

Since

$$\Pi^{\mathrm{H}}(\eta_I, \eta_P) = \Pi_0^{\mathrm{H}} + B_P(\gamma)\eta_I\sigma^2 + \big(A_P(\gamma) - B_P(\gamma)\big)\eta_P\sigma^2 - K(\eta_I),$$

disclosure is a boundary choice. On the admissible region,

$$B_P(\gamma) \equiv \frac{(1-\gamma)^2(1+2\gamma)}{(1+\gamma)(4+4\gamma-9\gamma^2)} > 0.$$

Moreover,

$$A_P(\gamma) - B_P(\gamma) = -\frac{(2\gamma^3 - 5\gamma^2 + 2\gamma + 2)\Psi_P(\gamma)}{2(1+\gamma)(4+4\gamma-9\gamma^2)(2+6\gamma-2\gamma^2-3\gamma^3)^2},$$

where

$$\Psi_P(\gamma) \equiv 18\gamma^6 + 15\gamma^5 - 91\gamma^4 + 24\gamma^3 + 40\gamma^2 - 4\gamma - 4.$$

Now $2\gamma^3 - 5\gamma^2 + 2\gamma + 2 > 0$ on $[0,1)$, while $\Psi_P(\gamma)$ has a unique root in $(0, \gamma_{\max})$, namely $\gamma_P \approx 0.387867$. Hence $A_P(\gamma) - B_P(\gamma)$ is positive for $\gamma < \gamma_P$ and negative for $\gamma > \gamma_P$, which proves the disclosure rule. Since $B_P(\gamma) > 0$, the platform always acquires information when k is sufficiently small.

Consider now consumer surplus. The sign of $A_{CS}(\gamma)$ is the sign of

$$\Psi_D(\gamma) \equiv 18\gamma^6 - 15\gamma^5 - 73\gamma^4 + 87\gamma^3 + 13\gamma^2 - 24\gamma - 4,$$

which has a unique root in $(0, \gamma_{\max})$, equal to $\gamma_D \approx 0.823566$. Thus public precision is consumer-improving if and only if $\gamma > \gamma_D$.

For exclusive precision,

$$B_{CS}(\gamma) = \frac{(\gamma-1)(4\gamma^2-\gamma-2)(14\gamma^2-7\gamma-6)}{2(1+\gamma)(4+4\gamma-9\gamma^2)^2}.$$

Since $\gamma < \gamma_{\max} < \frac{7+\sqrt{385}}{28}$, the factor $14\gamma^2 - 7\gamma - 6$ is strictly negative on the whole admissible region. Because $\gamma < 1$ as well, the sign of $B_{CS}(\gamma)$ is the sign of $4\gamma^2 - \gamma - 2$. Hence, exclusive precision raises consumer surplus if and only if $\gamma > \frac{1+\sqrt{33}}{8}$.

To obtain the overall policy ranking, compare $A_{CS}(\gamma)$, $B_{CS}(\gamma)$, and 0. The difference between disclosure and secrecy is

$$A_{CS}(\gamma) - B_{CS}(\gamma) = -\frac{(2\gamma^3 - 5\gamma^2 + 2\gamma + 2)\Psi_C(\gamma)}{2(1+\gamma)(4+4\gamma-9\gamma^2)^2(2+6\gamma-2\gamma^2-3\gamma^3)^2},$$

where

$$\Psi_C(\gamma) \equiv 252\gamma^8 - 204\gamma^7 - 1004\gamma^6 + 1197\gamma^5 + 291\gamma^4 - 596\gamma^3 - 34\gamma^2 + 88\gamma + 8.$$

On $(0, \gamma_{\max})$, the relevant sign change occurs at the larger admissible root $\gamma_C \approx 0.863205$. Since $\gamma_D < \gamma_E < \gamma_C$, the overall ranking is exactly the one stated in the proposition: no acquisition for $\gamma \leq \gamma_D$, disclosure for $\gamma_D < \gamma \leq \gamma_C$, and secrecy for $\gamma_C < \gamma < \gamma_{\max}$.

Finally, consider total welfare. The sign of $A_W(\gamma)$ is the sign of

$$\gamma(\gamma-1)\big(12\gamma^4-12\gamma^3-22\gamma^2+13\gamma+12\big).$$

The quartic term has no roots in $[0,1)$, so it is strictly positive there. Hence $A_W(\gamma)\le 0$, with equality only at $\gamma=0$. Likewise,

$$B_W(\gamma)=\frac{(\gamma-1)(28\gamma^4-28\gamma^3-15\gamma^2+12\gamma+4)}{2(1+\gamma)(4+4\gamma-9\gamma^2)^2},$$

and the quartic factor is again strictly positive on $[0,1)$, implying $B_W(\gamma)<0$ throughout the admissible set. Therefore, both public and exclusive precision are ex ante welfare-reducing, so total welfare is maximized by no acquisition.

Conditional on a positive stock of data, disclosure is preferred to secrecy whenever $A_W(\gamma)\ge B_W(\gamma)$. Their difference is

$$A_W(\gamma)-B_W(\gamma)=\frac{(1-\gamma)(2\gamma^3-5\gamma^2+2\gamma+2)\Psi_W(\gamma)}{2(1+\gamma)(4+4\gamma-9\gamma^2)^2(2+6\gamma-2\gamma^2-3\gamma^3)^2},$$

where

$$\Psi_W(\gamma)=126\gamma^7-129\gamma^6-214\gamma^5+171\gamma^4+150\gamma^3-74\gamma^2-32\gamma+8.$$

This polynomial has exactly two roots in $(0,\gamma_{\max})$, namely $\gamma_W^-\approx 0.200962$ and $\gamma_W^+\approx 0.670305$. Hence disclosure dominates secrecy below γ_W^- and above γ_W^+, whereas secrecy dominates disclosure in the intermediate region. ■

References

[1] Abhishek, Vibhanshu, Kinshuk Jerath, and Z. John Zhang. 2016. "Agency Selling or Reselling? Channel Structures in Electronic Retailing." *Management Science* 62(8): 2259–2280.

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

[3] Bergemann, Dirk, Alessandro Bonatti, and Alex Smolin. 2018. "The Design and Price of Information." *American Economic Review* 108(1): 1–48.

[4] Bergemann, Dirk, and Alessandro Bonatti. 2019. "Markets for Information: An Introduction." *Annual Review of Economics* 11: 85–107.

[5] Bergemann, Dirk, and Alessandro Bonatti. 2024. "Data, Competition, and Digital Platforms." *American Economic Review* 114(8): 2553–2595.

[6] Chemla, Gilles. 2003. "Downstream competition, foreclosure, and vertical integration." *Journal of Economics & Management Strategy*, 12(2), 261-289.

[7] Cheng, Hsing Kenneth, Kyung Sung Jung, Young Kwark, and Jingchuan Pu. 2023. "Impact of Own Brand Product Introduction on Optimal Pricing Models for Platform and Incumbent Sellers." *Information Systems Research* 34(3): 1131–1147.

[8] de Cornière, Alexandre, and Greg Taylor. 2019. "A Model of Biased Intermediation." *RAND Journal of Economics* 50(4): 854–882.

[9] Dewan, Sanjeev, Jooho Kim, and Tingting Nian. 2023. "Economic Impacts of Platform-Endorsed Quality Certification: Evidence from Airbnb." *MIS Quarterly* 47(3): 1353–1368.

[10] Engert, Martin, Andreas Hein, Likoebe M. Maruping, Jason Bennett Thatcher, and Helmut Krcmar. 2025. "Self-Organization and Governance in Digital Platform Ecosystems: An Information Ecology Approach." *MIS Quarterly* 49(1): 91–122.

[11] Gal-Or, Esther. 1985. "Information Sharing in Oligopoly." *Econometrica* 53(2): 329–343."

[12] Hagiu, Andrei, and Julian Wright. 2015. "Marketplace or Reseller?" *Management Science* 61(1): 184–203.

[13] Hagiu, Andrei, Tat-How Teh, and Julian Wright. 2022. "Should Platforms Be Allowed to Sell on Their Own Marketplaces?" *RAND Journal of Economics* 53(2): 297–327.

[14] He, Shu, Jing Peng, Jianbin Li, and Liping Xu. 2020. "Impact of Platform Owner's Entry on Third-Party Stores." *Information Systems Research* 31(4): 1467–1484.

[15] Huber, Thomas L., Thomas Kude, and Jens Dibbern. 2017. "Governance Practices in Platform Ecosystems: Navigating Tensions Between Cocreated Value and Governance Costs." *Information Systems Research* 28(3): 563–584.

[16] Krämer, Jan, and Shiva Shekhar. 2025. "Regulating Digital Platform Ecosystems Through Data Sharing and Data Siloing: Consequences for Innovation and Welfare." *MIS Quarterly* 49(1): 123–154.

[17] Long, Fei, and Wilfred Amaldoss. 2024. "Self-Preferencing in E-Commerce Marketplaces: The Role of Sponsored Advertising and Private Labels." *Marketing Science* 43(5): 925–952.

[18] Mao, Shengjun, Sanjeev Dewan, and Yi-Jen (Ian) Ho. 2023. "Personalized Ranking at a Mobile App Distribution Platform." *Information Systems Research* 34(3): 811–827.

[19] Mikalef, Patrick, Maria Boura, George Lekakos, and John Krogstie. 2020. "The Role of Information Governance in Big Data Analytics Driven Innovation." *Information & Management* 57(7): 103361.

[20] Nocke, Volker, and Rey, Patrick. 2018. "Exclusive dealing and vertical integration in interlocking relationships." *Journal of Economic Theory*, 177, 183-221.

[21] Raith, Michael. 1996. "A General Model of Information Sharing in Oligopoly." *Journal of Economic Theory* 71(1): 260–288.

[22] Shapiro, Carl. 1986. "Exchange of Cost Information in Oligopoly." *Review of Economic Studies* 53(3): 433–446.

[23] Singh, Nirvikar, and Xavier Vives. 1984. "Price and quantity competition in a differentiated duopoly." *The Rand journal of economics*: 546-554.

[24] Vives, Xavier. 1984. "Duopoly Information Equilibrium: Cournot and Bertrand." *Journal of Economic Theory* 34(1): 71–94.

[25] Vives, Xavier. 2010. *Information and Learning in Markets: The Impact of Market Microstructure*. Princeton, NJ: Princeton University Press.

[26] von Scherenberg, Franziska, Malte Hellmeier, and Boris Otto. 2024. "Data Sovereignty in Information Systems." *Electronic Markets* 34: Article 15.

[27] Wang, Ruitong, and Ye Qiu. 2024. "Dual Role and Product Featuring Strategy of Digital Platform." *Marketing Science* 43(6): 1168–1187.

[28] Zha, Yong, Quan Li, Tingliang Huang, and Yugang Yu. 2023. "Strategic Information Sharing of Online Platforms as Resellers or Marketplaces." *Marketing Science* 42(4): 659–678.

[29] Zhu, Feng, and Qihong Liu. 2018. "Competing with Complementors: An Empirical Look at Amazon.com." *Strategic Management Journal* 39(10): 2618–2642.