Oligopolistic Information Markets

Peter Achim & Roland Strausz*

November 24, 2025

Abstract

In modern information markets, buyers routinely combine signals from multiple sellers. We develop a model of "portfolio competition" to analyze this distinctive feature. We show that the combinability of information overturns standard oligopoly intuition. Unlike traditional markets, competitive pressure does not necessarily protect buyers: when signals are complements, sellers can leverage the buyer's desire for the joint portfolio to extract the full social surplus, regardless of the number of competitors. We characterize the precise conditions for rent extraction, which reduce to a simple geometric test for symmetric sellers. Furthermore, we find that the canonical logic of market entry fails. Entry is never socially excessive because efficient portfolio choices eliminate business-stealing effects. Paradoxically, entry can reduce competitive pressure: when entrants provide strong complementarities, they shift the buyer's threat point, allowing all sellers to extract higher rents.

*Achim: Department of Economics and Related Studies, University of York (email: peter.achim@york.ac.uk); Strausz: School of Business and Economics, Humboldt-Universität zu Berlin (email: strauszr@hu-berlin.de). We are very grateful to Helmut Bester, Hector Chade, Amparo Urbano, and Kun Zhang. Roland Strausz acknowledges financial support from the European Union through the ERC-grant PRIVDIMA (project number 101096682) and the Deutsche Forschungsgemeinschaft through CRC-TRR 190 (project number 280092119).

1 Introduction

As modern economies become increasingly information-driven, information has become a tradable commodity. Financial analysts sell earnings forecasts to investors, data brokers sell consumer profiles to marketers, and rating agencies sell credit assessments to lenders. A defining institutional feature of these information markets is that buyers routinely combine information from multiple sources. Institutional investors assemble forecasts from distinct financial data providers; e-commerce firms aggregate fraud scores from multiple risk intelligence vendors; pharmaceutical companies integrate clinical trial data from different research organizations.

To reflect this empirical reality, we develop the first model of oligopolistic pricing in information markets that explicitly captures the combinability of signals. Modeling price competition in the traditional sense of Bertrand, we show that this combinability creates a distinctive competitive structure: sellers price-compete not just against individual rivals but against all possible combinations of rivals that buyers might assemble. We term this competitive structure *portfolio competition*.

Analyzing portfolio competition reveals that the economics of information markets overturns standard oligopoly intuition. In traditional markets, competition protects buyers: as the number of sellers increases, prices fall toward marginal cost, and buyers capture the surplus. We show that in information markets, this logic frequently breaks down.

Our first main result establishes that competitive pressure does not necessarily protect buyers from full rent extraction. The intuition relies on the distinction between substitutes and complements in information. When signals are substitutes (e.g., overlapping forecasts), sellers compete intensely to be included in the portfolio, driving prices down. However, when signals are complements (e.g., distinct data points that reveal a pattern only when combined), the buyer's value of the whole portfolio exceeds the sum of the parts. In this scenario, sellers can collectively leverage the buyer's fear of losing the entire portfolio to extract the full social surplus, leaving the buyer with nothing. We characterize the precise boundary between these outcomes using a "balancedness" condition, which formally captures the degree of complementarity in the market and generalizes the complements-versus-substitutes dichotomy to arbitrary market sizes.

Our second set of results challenges the canonical view of market entry. In standard models of free entry with substitute differentiated products, entrants impose a business-stealing externality, so equilibrium entry is socially excessive. We prove that in information markets, excessive entry never occurs. Because buyers combine signals to form efficient portfolios, there is no "business stealing" in equilibrium—every active seller contributes unique value.

Paradoxically, we find that entry can actually reduce competitive pressure. In standard markets, a new entrant improves the buyer's outside option. In information markets with strong complementarities, a new entrant can increase the value of the "grand coalition" so significantly that it renders the buyer's threat to exclude individual sellers non-credible. Consequently, the entry of a new firm can shift the equilibrium from one where the buyer retains surplus to one where sellers extract everything.

These findings suggest that traditional antitrust heuristics are ill-suited for the data economy. Because market failures arise from insufficient entry (coordination failures) rather than excessive entry, and because competition does not guarantee consumer surplus, regulators must look beyond concentration indices like HHI. The structure of information complementarity, not the number of firms, determines welfare.

The remainder of the paper is organized as follows. Section 2 discusses related literature. Section 3 develops our portfolio competition framework that captures the combinability of competitive information signals. In Section 4, we first study the duopoly case of two firms, obtaining the result that the complementarity of signals determines the equilibrium degree of rent extraction. Section 5 extends this analysis to the general oligopoly model, obtaining the insight that in this more general framework the correct defining notion is a balancedness condition. Section 6 shows that with symmetric sellers, this balancedness condition has a simple geometric interpretation. Section 7 characterizes the free-entry equilibrium, showing that it never exhibits any excess entry. Section 8 concludes.

2 Related Literature

The theoretical literature on information markets focuses on monopolistic settings. Bergemann et al. (2018); Ali et al. (2022); Yang (2022) analyze how a monopolist designs optimal information-selling mechanisms. Certification models like Lizzeri (1999); Stahl and Strausz (2017); DeMarzo et al. (2019) analyze information provision in monopolistic certification markets. While Gentzkow and Kamenica (2017a,b); Li and Norman (2021); Wu (2023) study Bayesian persuasion with competing senders, their framework excludes transferable utility and thus price competition.

These frameworks abstract from oligopolistic price competition where buyers combine signals from multiple sellers. We fill this gap by developing the first model of oligopolistic information provision with portfolio competition.

Our portfolio competition model connects to the multi-product pricing literature of Tauman et al. (1997) and Arribas and Urbano (2005), who characterize equilibrium prices when buyers assemble portfolios. Information markets constitute a special case of their framework: they naturally satisfy strict monotonicity, where each seller provides value not fully replicated by others.

Leveraging strict monotonicity allows us to achieve a full characterization of the conditions under which competitive pressure protects buyers from rent extraction— weak and strong balancedness. These conditions generalize the duopoly complements-versus-substitutes dichotomy to oligopoly with arbitrary numbers of sellers. Strict monotonicity ensures (i) buyers purchase from all efficient sellers in equilibrium, and (ii) the cooperative game's core is non-empty, enabling our balancedness analysis.

This extends Arribas and Urbano (2005, Section 5), who analyze concave and convex (but not strictly monotone) value functions. Their analysis shows that without strict monotonicity, the core may be empty and total revenue varies across equilibria. Our strict monotonicity assumption guarantees core non-emptiness via the Bondareva-Shapley theorem and unique total revenue across all equilibria, though individual price divisions remain indeterminate. This enables our complete characterization of buyer surplus: weak balancedness determines when full extraction is achievable (Proposition 4), while strong balancedness (equivalently, supermodularity) ensures every equilibrium extracts full surplus (Propositions 6-7). For solution concepts like the Shapley value or nucleolus, our strict monotonicity ensures both always exist and lie in the core, whereas in Arribas et al.'s non-monotone setting, these concepts may fall outside the (possibly empty) core.

Our analysis also reveals a fundamental distinction between entry dynamics in markets for divisible versus indivisible goods. The classical oligopoly literature on differentiated products (Singh and Vives, 1984; Dixit and Stiglitz, 1977) analyzes competition where firms produce divisible goods and buyers choose continuous quantities from each seller. In such settings, the entry externalities literature (Mankiw and Whinston, 1986; Spence, 1976) establishes that substitutes generate business-stealing effects, where each entrant captures demand from incumbents, leading to socially excessive entry. Conversely, with complements, entry creates business-enhancing effects that entrants do not fully internalize, potentially leading to insufficient entry.

By contrast, our framework with indivisible information signals exhibits neither effect. Entry is never excessive because the buyer purchases from all firms from whom it is efficient to buy, and this holds in any equilibrium regardless of entry. This eliminates business-stealing and business-enhancing effects entirely. When insufficient entry occurs, it arises from coordination failures in pricing rather than uninternalized business-enhancing externalities. This demonstrates that the canonical entry externalities from divisible goods markets do not extend to unit-demand settings, where the indivisibility of goods fundamentally alters competitive dynamics.

Focusing on a buyer without any private information, we abstract from issues studied in monopolistic models of information selling such as Bergemann et al. (2018); Yang (2022). Incorporating such private information in a competitive model is non-trivial as it would require a model of competing mechanism design as pioneered in McAfee (1993).

Modeling the demand side by a representative buyer allows us to abstract from information externalities between buyers, which is the focus of, for example, Admati and Pfleiderer (1986); Choi et al. (2019); Acemoglu et al. (2022); Bergemann et al. (2022).

Our model of portfolio competition is orthogonal to the literature on information sharing between competing firms (Vives, 1984, 1988; Raith, 1996), where the focus is on firms exchanging information with each other.

3 Model

Following Raiffa and Schlaifer (1961), we study a market where information helps a representative buyer make decisions under uncertainty. In particular, the buyer faces an unknown state of the world ω, drawn from a finite set

$$\omega \in \Omega = \{\omega_1, \ldots, \omega_m\},$$

according to a commonly known prior distribution $p_0 \in \Delta(\Omega)$. The buyer must decide which information sources to purchase before making a decision that will determine both the quality of her eventual choice and the total cost of information acquisition.

Information is provided by n independent sellers, indexed by $N = \{1, ..., n\}$, each possessing private signals correlated with the true state ω. This creates the portfolio competition environment identified in the introduction: the buyer can purchase information from any subset, $S \subseteq N$, of the N sellers and combine their signals to form better posterior beliefs. The key economic tension emerges because each seller's value to the buyer depends not only on the quality of that seller's own signal but critically on which other sellers the buyer also includes in her information portfolio.

Each seller i observes a private signal σ_i drawn from a finite space Σ_i, with the joint signal profile $\sigma = (\sigma_1, \ldots, \sigma_n) \in \Sigma = \Pi_i \Sigma_i$ distributed according to a Blackwell experiment $P(\sigma|\omega)$ conditional on the state. While sellers share common knowledge of this joint distribution, each observes only her own signal realization, ensuring that different sellers may provide genuinely different, and potentially complementary, information about the underlying state.

The buyer's willingness to pay for information reflects how improved posterior beliefs translate into higher payoffs in her operational environment. We capture this through the buyer's belief-based value function

$$v : \Delta(\Omega) \longrightarrow \mathbb{R},$$

which assigns to each posterior belief $p \in \Delta(\Omega)$ the maximum expected payoff the buyer can achieve when making decisions based on that belief. This belief-based value function is central to understanding portfolio competition: when the buyer considers purchasing from multiple sellers, she evaluates not just the cost of each information source, but how different combinations of signals affect her posterior beliefs and thus her decision-making value.

We assume v is strictly convex, reflecting the property that information has decision-making value. It implies that information becomes more valuable when combined with other information sources, a key driver of the complementarity patterns discussed in the introduction. The convexity assumption ensures that the marginal value of any seller's signal depends on which other signals the buyer also acquires, formalizing the economic intuition that sellers face portfolio competition rather than simple head-to-head rivalry.

Portfolio competition arises because buyers can combine signals from multiple sellers while sellers cannot observe competitors' information. Each seller's value to the buyer depends not only on signal quality, but on how that signal interacts with others in the buyer's information portfolio. Whether sellers compete as substitutes or complements depends on the buyer's value function v. When combining information sources yields diminishing returns, sellers' information is substitutable; when information sources reinforce each other, sellers compete as complements and can sustain higher prices.

Timing The timing of the portfolio selection game is as follows:

1. *Simultaneous Pricing.* Each seller simultaneously sets a nonnegative price $t_i \geq 0$ for access to their private signal. These prices represent binding commitments; sellers cannot condition their pricing on the buyer's eventual portfolio choices or on competitors' pricing decisions.

2. *Portfolio Selection.* The buyer observes all posted prices $t = (t_1, \dots, t_n)$ and selects which subset $S \subseteq N$ of sellers to buy from, paying the total cost $\sum_{i \in S} t_i$. This decision is made ex ante, before any information content is revealed, so the buyer must evaluate the expected value of different information portfolios based on the known signal structure and her value function v.

3. *Information Realization and Payoffs.* Nature draws the true state $\omega \sim p_0$ and the corresponding signal profile $\sigma \sim P(.|\omega)$. Each purchased seller $i \in S$ observes their signal realization σ_i and discloses it to the buyer. The buyer then updates her beliefs based on the observed signal profile and realizes her final payoff $v(p_m)$, where p_m represents her posterior belief after observing signals from her chosen portfolio S.

The buyer's portfolio selection problem requires evaluating the expected value of different information combinations. For any subset $S \subseteq N$ of sellers, the buyer must assess how signals from that specific portfolio will improve her decision-making. This valuation process lies at the heart of portfolio competition: sellers compete not just on individual signal quality, but on their marginal contribution to the buyer's preferred information portfolio. When the buyer purchases information from sellers in set S, she receives signal realizations that allow Bayesian updating from her prior p_0 to posterior beliefs p_{σ_S}.

Strategic Payoffs To formalize this evaluation process, we let

$$V_{\emptyset} = \sum_{\omega \in \Omega} p_0(\omega) v(p_0) = v(p_0),$$

represent the buyer's outside option value of making decisions based solely on the prior distribution. We define the *portfolio value* V_S as the buyer's expected payoff from purchasing information from subset S net of $V_{\emptyset}$:

$$V_S = \sum_{\omega \in \Omega} \sum_{\sigma_S \in \Sigma_S} p_0(\omega) P(\sigma_S | \omega) v(p_{\sigma_S}) - V_{\emptyset},$$

where p_{σ_S} denotes the buyer's posterior and $P(\sigma_S \mid \omega)$ the marginal probability of observing the signal profile when the state is ω. The portfolio value V_S represents the buyer's willingness to pay for information portfolio S, which directly determines each seller's market value and competitive position; it captures the strategic essence of portfolio competition. Crucially, a seller's individual worth depends not only on their signal quality, but on how their information complements or substitutes for others in the buyer's optimal portfolio choice.

To avoid confusion, we distinguish terminology: $v(p)$ is the buyer's *value function* mapping beliefs to decision payoffs, while V_S denotes the *portfolio value*—the expected value (measured via v) from purchasing signals from coalition S rather than none.

The buyer's net benefit from purchasing information portfolio S at prices t is her information value minus the total cost:

$$V_S - \sum_{i \in S} t_i.$$

Assumption 1 (Strict Monotonicity)**.** *Each seller provides strictly positive incremental informational value: for all coalitions $S \subseteq N$ and all sellers $i \in N \setminus S$,*

$$V_{S \cup \{i\}} > V_S.$$

This assumption reflects markets where information providers differentiate to avoid direct redundancy. While incremental value may be arbitrarily small (capturing near-redundant signals), we rule out perfect redundancy. As we discuss in the related literature, strict monotonicity ensures efficient equilibrium portfolios, active participation by all sellers when signals are costless, and unique total revenue across equilibria.

Sellers face strategic interdependence in portfolio competition: each earns their posted price t_i

only if the buyer includes them in her chosen portfolio, and zero otherwise, yielding payoff

$$\Pi_i(t,S) = \begin{cases} t_i, & \text{if } i \in S, \\ 0, & \text{otherwise.} \end{cases}$$

This creates the strategic pressure that drives our analysis: sellers must price competitively enough to gain inclusion, yet cannot directly observe the competitive threats they face from alternative seller combinations.

Equilibrium Concept We analyze pure-strategy subgame perfect Nash equilibria of the portfolio selection game. An equilibrium consists of prices $t^* = (t_1^*, \dots, t_n^*)$ and a buyer portfolio choice function $S^*(t)$, mapping prices t into a portfolio selection $S \subseteq N$, such that:

(i) (*Buyer Optimization*) Given posted prices t, the buyer selects the portfolio that maximizes her net benefit:

$$S^*(t) \in \arg\max_{S \subseteq N} \left\{ V_S - \sum_{i \in S} t_i \right\}.$$

(ii) (*Seller Optimization*) No seller can improve their payoff by unilaterally changing their price, taking the buyer's optimal response as given:

$$\Pi_i(t^*, S^*(t^*)) \geq \Pi_i((t_i, t_{-i}^*), S^*(t_i, t_{-i}^*)) \quad \forall t_i \geq 0; \forall i \in N.$$

This equilibrium concept captures the essence of portfolio competition: sellers must anticipate not just direct rivalry, but the buyer's ability to substitute entire portfolios of competing information sources.

Running example We introduce a running example that illustrates our results throughout the paper. The example microfounds the value function $v(p)$ and yields a tractable portfolio value V_S.

Example. *Our buyer is a monopolist who produces quantity $q \geq 0$ in a downstream market with a cost function $C(q)$. Suppose the monopolist's inverse demand function, $P(q|p)$, depends on the posterior beliefs $p = \mathbb{P}(\omega = h)$ about a binary state $\omega \in \Omega = \{l, h\}$ with prior $p_0 = 1/2$. Given a posterior belief p, the monopolist maximizes profits, resulting in the monopolist's value function*

$$v(p) = \max_{q \geq 0} \left[P(q|p) \cdot q - C(q) \right].$$

For our running example, we specify linear demand $P(q|p) = 2p - q$ and no costs of production $C(q) = 0$. The first-order condition yields optimal quantity $q^* = p$, giving:

$$v(p) = (2p - q^*) \cdot q^* = (2p - p) \cdot p = p^2.$$

The quadratic value function has an important implication: the expected value from signals equals the variance of the induced posterior distribution. From $p_0 = 1/2$ and $v(p) = p^2$, it follows that $V_\emptyset = v(p_0) = 1/4$, which together with the law of iterated expectations, $\mathbb{E}[\hat{p}] = p_0$, yields

$$V_S = \mathbb{E}[\hat{p}^2] - V_\emptyset = \mathbb{E}[\hat{p}^2] - (\mathbb{E}[\hat{p}])^2 = \mathrm{Var}(\hat{p}). \tag{1}$$

Concerning the informative signals of the sellers in S, we assume that each seller i observes an independently drawn binary signal $s_i \in \{l, h\}$ with identical signal accuracy α, that is $P(s_i = \omega|\omega) = \alpha \in (1/2, 1)$ for all i. This framework of independent and symmetric signals is a natural benchmark and yields a distribution of posteriors, $\hat{p}$, that is tractable.

The resulting example is canonical because it yields closed-form solutions while capturing the essence of information markets: each seller provides an independent but imperfect signal about a common underlying state. We use this example to illustrate our main results, beginning with duopoly in Section 4.

4 Duopoly

We begin by considering the case of two sellers, $n = 2$. For simplicity, we write $V_i = V_{\{i\}}$ and $V_n = V_{\{1,2\}}$. We also focus on the case that each experiment is individually valuable. We therefore assume that V is strictly increasing in the following sense: $V_i > 0$, and $V_n > \max\{V_1, V_2\}$. The latter means that, conditional on each state, the sellers' experiments are at most partially correlated.

Informational Complements and Substitutes. The sellers are said to be *informational complements* (see Börgers et al., 2013), if the joint value exceeds the sum of the parts:

$$V_n \geq V_1 + V_2.$$

This means that each seller provides unique insights that are not replaceable by the other. Conversely, the sellers are *informational substitutes* if:

$$V_n \leq V_1 + V_2.$$

Here, the information from each seller overlaps, and their combined contribution is less than the sum of their individual values.

Using the common convention that $-i \in \{1,2\} \backslash \{i\}$, we define i's marginal contribution, ΔV_i, to V_n as

$$\Delta V_i = V_n - V_{-i}.$$

Our monotonicity assumption on V implies that the marginal contributions are strictly positive: $\Delta V_i = V_n - V_{-i} > \max\{V_1, V_2\} - V_{-i} \geq V_{-i} - V_{-i} = 0$.

Demand for Information. After each seller $i = 1,2$ sets their price t_i, the buyer chooses among four options: obtaining no signal and paying 0; obtaining a signal only from seller 1 and paying t_1; obtaining a single signal from seller 2 and paying t_2; or obtaining a signal from both sellers, paying $t_1 + t_2$.

1. *Two signals:* The buyer optimally buys from both sellers if and only if it holds

$$t_1 + t_2 \leq V_n \ \wedge \ t_1 \leq V_n - V_2 = \Delta V_1 \ \wedge \ t_2 \leq V_n - V_1 = \Delta V_2. \tag{2}$$

   The first inequality ensures that the joint surplus from two signals exceeds the cost. The second and third inequalities ensure that the buyer does not strictly prefer purchasing from just one seller. If either seller charges too much relative to the other's value, the buyer would deviate to buying only one signal. Hence, all three constraints must be satisfied for two signals to be optimal.

2. *A single signal from seller $i = 1,2$:* The buyer optimally purchases only from seller i and not from $-i$ if and only if it holds

$$t_i \leq V_i \ \wedge \ t_1 + t_2 > V_n \ \wedge \ V_{-i} - t_{-i} \leq V_i - t_i.$$

   The first condition ensures that seller i alone provides positive net value. The second set of conditions guarantees that the buyer does not prefer two signals or a single signal of the opponent.

3. *No signal:* The buyer optimally declines to purchase any signal if and only if it holds

$$t_1 \geq V_1 \ \wedge \ t_2 \geq V_2 \ \wedge \ t_1 + t_2 \geq V_n.$$

   Each condition eliminates one of the possible choices. If buying either single signal does not yield a surplus, and purchasing both signals together is also too costly, then the buyer is better off relying on the prior and not buying any signal.

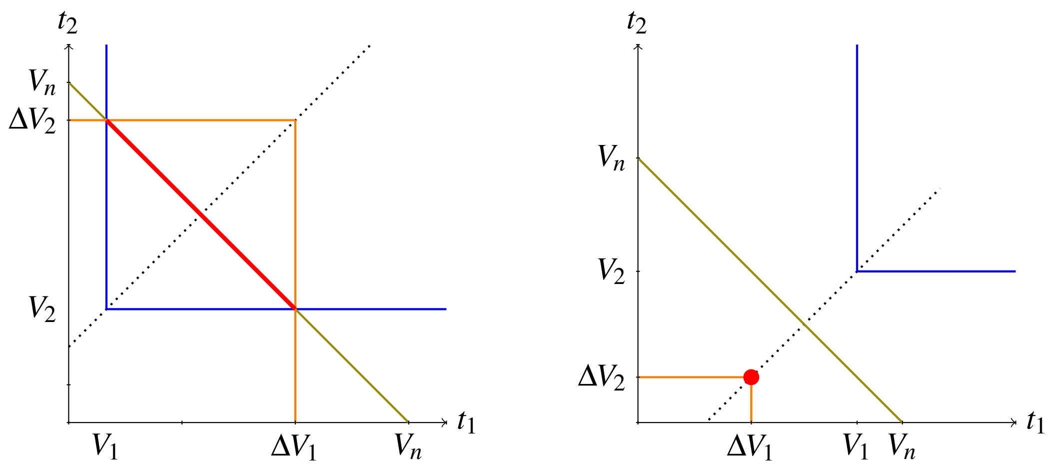

Figure 1: The left panel of the diagram illustrates the complementary case $V_1 + V_2 < V_n$, the right panel the alternative, substitute case. The orange line depicts prices such that the buyer is indifferent between buying two or one signal; in the area below it the buyer strictly prefers the option to buy both signals. The olive line depicts pairs such that the buyer is indifferent between two signals and no signal; in the area below it the buyer strictly prefers the two signals. The blue line depicts prices so that the buyer is indifferent between a single signal and no signal; in the area below it the buyer strictly prefers the option to buy one signal. The red line (left) and the red dot (right) show the set of equilibrium prices. In equilibrium, the buyer buys both signals, and each seller extracts at most its marginal contribution.

These inequalities define the buyer's demand correspondence over the price space. In particular, the regions partition the (t_1, t_2) plane into areas where different combinations of signals are optimal. The key takeaway is that the buyer's decision hinges on comparing net surpluses—the improvement in market value induced by the purchased information, minus the price charged.

The following proposition characterizes the equilibrium when the signals are incomplete — that is, even after purchasing all available signals, the buyer does not learn the state with certainty.

Proposition 1. *Suppose $V_i > 0$ and $V_n > \max\{V_1, V_2\}$. In every equilibrium, the buyer purchases both signals. The equilibrium prices are as follows.*

(i) If the signals are substitutes, each seller sets a price equal to its marginal contribution, $t_i = \Delta V_i$.

(ii) If the signals are complements, then the set of equilibrium prices is given by the budget line $t_1 + t_2 = V_n$, subject to the individual surplus constraints $t_i \leq \Delta V_i$.

The proposition's proof which we provide in the appendix is straightforward. It directly formalizes the intuitive idea that any excluded seller can profitably deviate by lowering her price to capture positive demand, forcing purchasing from all sellers in equilibrium. For signals that are substitutes it then follows that each seller faces direct competitive pressure limiting prices to individual marginal contributions, whereas for signals that are complements, equilibrium prices lie on the buyer's budget constraint since neither seller can be easily replaced by the other.

This result illustrates how strategic pricing in a duopolistic market of information depends on the informational landscape. When sellers provide complementary signals, they can extract the full value they jointly create, leaving no rents to the buyer. The exact division of rents between the two sellers is indeterminate: any sharing rule satisfying the buyer's participation and incentive constraints can be sustained in equilibrium. By contrast, when the sellers' signals are substitutes, the interaction becomes more competitive. In the substitute case, each seller's marginal contribution ΔV_i is constrained by competitive pressure from the alternative seller, creating a binding constraint that prevents full surplus extraction and ensures the buyer retains positive surplus. This competitive pressure disciplines pricing and ensures that no seller extracts more than its marginal contribution to total value. As a result, the buyer obtains a strictly positive rent from consuming the signals.

Example. *We analyze the duopoly case ($n = 2$) for our running example. When the buyer purchases from only one seller, Bayes' rule yields posterior beliefs:*

$$\hat{p}(h) = \mathbb{P}(\omega = h \mid s_1 = h) = \frac{p_0 \cdot \alpha}{p_0 \cdot \alpha + (1 - p_0)(1 - \alpha)} = \alpha$$

By symmetry, $\hat{p}(l) = \mathbb{P}(\omega = h \mid s_1 = l) = 1 - \alpha$. Since each posterior occurs with probability $1/2$, the portfolio value is

$$V_1 = V_2 = \left(\alpha - \frac{1}{2}\right)^2.$$

Portfolio value increases quadratically with signal accuracy α.

With two independent signals, let m denote the number of high signals. The distribution of m is $\text{Binomial}(2, \alpha)$ *when $\omega = h$ and* $\text{Binomial}(2, 1 - \alpha)$ *when $\omega = l$. Applying Bayes' rule for each realization yields posterior beliefs:*

$$\hat{p}_0 \equiv \mathbb{P}(h|0) = \frac{(1-\alpha)^2}{(1-\alpha)^2 + \alpha^2}; \quad \hat{p}_1 \equiv \mathbb{P}(h|1) = \frac{1}{2}; \quad \hat{p}_2 \equiv \mathbb{P}(h|2) = \frac{\alpha^2}{\alpha^2 + (1-\alpha)^2}.$$

Note that contradictory signals ($m = 1$) yield no information beyond the prior, while agreement reinforces belief in either direction. The probabilities are:

$$\mathbb{P}(0) = \mathbb{P}(2) = \frac{\alpha^2 + (1-\alpha)^2}{2}, \quad \mathbb{P}(1) = 2\alpha(1-\alpha).$$

The expected posterior squared equals:

$$\mathbb{E}[\hat{p}^2] = \mathbb{P}(0)\hat{p}_0^2 + \mathbb{P}(1)\hat{p}_1^2 + \mathbb{P}(2)\hat{p}_2^2 = \frac{3\alpha^2 - 3\alpha + 1}{4\alpha^2 - 4\alpha + 2}.$$

The value of the complete portfolio of both signals is therefore:

$$V_n = \mathbb{E}[\hat{p}^2] - V_\emptyset = \frac{3\alpha^2 - 3\alpha + 1}{4\alpha^2 - 4\alpha + 2} - \frac{1}{4}.$$

Direct calculation shows signals are substitutes:

$$V_1 + V_2 - V_n = \frac{(1-2\alpha)^4}{4 - 8(1-\alpha)\alpha} > 0.$$

The signals provide overlapping information about the same state, so acquiring both yields less than twice the value of a single signal. By Proposition 1, each seller charges $t_i = \Delta V_i = V_n - V_{-i}$ in the unique equilibrium, and the buyer retains positive surplus $V_n - (2V_n - V_1 - V_2) > 0$.

5 Oligopoly

The duopoly analysis revealed that buyers always purchase the socially efficient portfolio, i.e., from both sellers, and that whether the buyer obtains a positive surplus depends on whether signals are complements or substitutes. A natural question arises: which of these insights extend to markets with $n > 2$ sellers, and where do new complexities emerge?

The structural difference is stark. In duopoly, the buyer has only two effective deviations: exclude seller 1 or exclude seller 2. With n sellers, however, the buyer can potentially switch to any of $2^n - 1$ alternative coalitions. These additional alternatives fundamentally change the competitive problem.

A simple complements-versus-substitutes dichotomy does not suffice to determine surplus division. With two sellers, each faced competition only from the other individual seller. With n sellers, seller i competes not just against individual rivals, but against all possible coalitions of rivals that the buyer might assemble. This portfolio competition creates a new analytical challenge: characterizing when coalition-based competitive pressure is sufficient to protect buyers from rent extraction.

Despite this added complexity, some fundamental insights from duopoly extend naturally. Most notably, the buyer's portfolio choice remains socially efficient in any equilibrium: the competitive pressure that forced inclusion of both sellers continues to operate with any number of sellers. The underlying logic remains the same: any excluded seller has a strong incentive to lower her price to reenter the bundle. Since sellers set prices while their cost of providing information is zero, sellers can ensure inclusion by setting sufficiently low but positive prices.

To address the surplus division question, we develop a new characterization based on the concept of exposure. A seller is exposed when the buyer can profitably switch to some alternative coalition, creating competitive pressure on that seller's price. Unlike duopoly, where exposure to individual competitors fully determined outcomes, oligopoly exposure to coalitions requires new analytical tools from linear programming and cooperative game theory.

Definition 1 (Exposure)**.** *Seller i is said to be* exposed *to a subset $S_{-i} \subseteq N \setminus \{i\}$ at price vector $t = (t_1, \ldots, t_n)$ if*

$$\sum_{j \in N \setminus S_{-i}} t_j = V_N - V_{S_{-i}}.$$

Intuitively, a price vector t exposes seller i if the buyer would refuse to purchase from seller i when i raises her price above t_i. In this sense, exposure formalizes the competitive pressure that disciplines seller i's price.

Note that in duopoly, every seller with positive price was necessarily exposed to the single alternative of the other seller. The exposure condition thus generalizes this competitive pressure to encompass the much richer set of coalition-based threats that emerge with multiple sellers. This more general condition captures the idea that seller i is exposed when the buyer is exactly indifferent between the current arrangement and switching to subset S_{-i}, excluding seller i (and possibly others not in S_{-i}).

Exposure is a local condition: it depends on the configuration of prices and values that make a specific alternative bundle S_{-i} the buyer's next-best option. Intuitively, a seller is exposed if she must keep her price pinned just below a threshold, beyond which the buyer would defect.

Having defined exposure, we can now provide a complete characterization of an oligopoly equilibrium. Unlike duopoly where equilibrium conditions were relatively simple, oligopoly requires two types of constraints: buyer incentive compatibility (which generalizes from duopoly) and seller exposure (which captures the new oligopoly complexity)

Proposition 2 (Equilibrium prices via buyer-optimality and exposure)**.** *(S^*, t^*) is a pure-strategy subgame perfect Nash equilibrium if and only if:*

1. *(S^*, t^*) exhibits "portfolio-efficiency": The buyer prefers to buy from all sellers over any strict subset:*

$$S^* = N \text{ and } \sum_{i \in N \setminus S} t_i^* \leq V_N - V_S, \quad \forall S \subseteq N. \tag{3}$$

2. *t^* exhibits "seller-exposure": For each seller $i \in N$, there exists a subset $S_{-i} \subseteq N \setminus \{i\}$ such that*

$$\sum_{j \in N \setminus S_{-i}} t_j = V_N - V_{S_{-i}}. \tag{4}$$

Note that when $n = 2$, condition (3) reduces to the familiar constraints $t_1 + t_2 \leq V_N$ and $t_i \leq \Delta V_i$

in (2), while condition (4) becomes the exposure to individual competitors that we saw in the duopoly analysis.

Our baseline analysis assumes zero production costs for expositional clarity. With per-sale costs $c_i \geq 0$, the framework extends naturally by identifying the set of *efficient* sellers. Define the efficient portfolio

$$N^* \in \arg\max_{S \subseteq N} \left\{ V_S - \sum_{i \in S} c_i \right\}.$$

If for all $S \subseteq N^*$ and all $i \in N^* \setminus S$, seller i's incremental informational value exceeds its cost,

$$V_{S \cup \{i\}} - V_S > c_i,$$

then the portfolio value V_S remains strictly monotone on N^*. In this case, our complete analysis applies to the game among the $|N^*|$ viable sellers: all our equilibrium and surplus characterization results hold with N replaced by N^* and prices t_i interpreted as gross prices (buyer payments) from which sellers earn profits $\pi_i = t_i - c_i$. Sellers in $N \setminus N^*$ optimally remain inactive as their costs exceed their incremental value in any portfolio.

This framework accommodates heterogeneous costs while preserving the portfolio competition structure. The market endogenously selects which sellers participate, and among efficient sellers, competition operates exactly as characterized in our zero-cost baseline.

5.1 Cooperative Benchmark and Equilibrium Existence

In duopoly, equilibrium analysis was relatively straightforward because we could directly characterize prices through individual marginal contributions and simple budget constraints. With oligopoly, however, the exponential growth in buyer deviation possibilities (from 2 to $2^n - 1$ alternative coalitions) makes direct equilibrium characterization unwieldy. This complexity forces us to employ more sophisticated analytical tools that were unnecessary in the two-seller case.

To better understand the structure of oligopoly equilibrium prices, we now consider a centralized, cooperative version of the pricing problem. This representation of the problem was already established in Arribas and Urbano (2005), who derived it in the broader context of a general multiproduct price competition. Applying this result to our context serves three purposes: (1) as a benchmark for evaluating competitive outcomes, (2) as a technical tool for characterizing the complete set of equilibria and connecting to our later analysis of when balancedness conditions determine surplus division, and (3) as a bridge connecting our information market analysis to the broader multiproduct price competition literature in industrial organization.

Oligopoly sellers must defend against an exponentially large set of buyer deviations. The cooperative framework allows us to capture this full competitive landscape through a systematic linear programming approach. Suppose that instead of setting prices independently, the sellers coordinate to choose prices jointly in order to maximize their total revenue. They take the buyer's preferences as given, and design prices so that the buyer voluntarily purchases the full bundle N, while extracting as much value from her as possible.

The buyer still chooses her bundle by maximizing net utility, so the sellers must ensure that the buyer prefers the full bundle over any strict subset. That is, the buyer's net value from purchasing from N must weakly exceed the net value of any $S \subseteq N$. These constraints define the feasible set of prices. Formally, the sellers solve the following LP-problem:

$$\max_{t \in \mathbb{R}^N_{\geq 0}} \quad \sum_{i \in N} t_i \quad \text{s.t.} \quad \sum_{i \in N \setminus S} t_i \leq V_N - V_S \quad \text{for all } S \subset N. \tag{5}$$

We refer to this as the *Primal (Revenue) LP*. Its objective captures the sellers' joint revenue. The constraints ensure that the buyer is willing to purchase from all sellers. The feasible region defined by this LP is a polytope in $\mathbb{R}^n$, bounded by one linear inequality for each strict subset $S \subseteq N$. Since there are $2^n - 1$ such subsets, the polytope is defined by exponentially many halfspaces, a complexity that duopoly entirely avoided with its simple triangular feasible region.

Each inequality reflects the condition that the buyer must not prefer to exclude the sellers in $N \setminus S$. Geometrically, each constraint slices off a part of the price space where the buyer would defect to a smaller bundle. This geometric complexity explains why the duopoly's simple marginal contribution bounds become insufficient: with multiple sellers, the buyer's deviation threats create a much richer constraint structure that requires linear programming techniques to analyze systematically.

This has an important implication that generalizes our duopoly exposure insight: for any seller i who receives a strictly positive price in the optimal solution, there must be some subset $S_i \subseteq N \setminus \{i\}$ such that the constraint corresponding to S_i binds with equality:

$$\sum_{j \in N \setminus S_i} t_j = V_N - V_{S_i}.$$

This binding constraint condition is precisely what we termed "exposure" in our equilibrium characterization: seller i is exposed to coalition S_i when the buyer's threat to switch to S_i becomes binding. For duopoly ($n = 2$), this reduces to each seller being exposed to their individual rival, recovering the familiar competitive pressure we analyzed earlier.

The buyer's incentive constraints are acting as resource constraints from the sellers' point of view. The solution must stretch those resources to their limits, and each seller who extracts revenue must be "pushed back" against one of the buyer's constraints. The seller cannot raise her price further, because she is already exposed to a binding deviation.

Most remarkably, and in line with Arribas and Urbano (2005), this cooperative revenue maximization problem identifies a set of prices that are exactly the same price profiles that arise in equilibrium in the strategic game. Even though sellers do not coordinate their prices in the actual game, competition leads them to an outcome that mirrors the outcome of this centralized optimization. This equivalence between cooperative and competitive outcomes, while intuitive in duopoly, becomes a powerful and non-obvious result in oligopoly markets.

Proposition 3. *The cooperative revenue maximization problem* (5) *has a solution set $T^* \subseteq \mathbb{R}^n$ which is non-empty, where T^* denotes the set of optimal price vectors that solve the revenue maximization LP (5). Every solution in T^* constitutes a subgame perfect Nash equilibrium. Moreover, any price profile in T^* minimizes the buyer's surplus across all subgame perfect Nash equilibria.*

The claims provided by Proposition 3 are a central finding in Arribas and Urbano (2005) and are contained within the combination of their Propositions 2 and 4 and, most pointedly, Corollary 2. We therefore refer to Arribas and Urbano (2005) for a formal proof of our proposition. We only remark that to see existence, it suffices to note that the feasible region is always non-empty (e.g., $t = 0$ is trivially feasible), and bounded above by the total value V_N. Therefore, an optimal solution always exists.

5.2 Buyer Surplus

Having characterized oligopoly equilibria through the LP framework, we now address a fundamental question: when do buyers retain surplus in oligopoly markets? In duopoly, this question had a clean answer through the complements-vs.-substitutes dichotomy. Does this simple logic extend to markets with $n > 2$ sellers, or do new complexities emerge?

In duopoly, competitive pressure came from exactly two sources, making surplus division depend solely on whether the sum of marginals exceeds the total value. With n sellers, however, competitive pressure can emerge from any of 2^{n-1} alternative coalitions, fundamentally changing how we must analyze surplus division. This exponential growth in competitive threats forces us beyond the elementary tools that sufficed for duopoly, requiring more sophisticated techniques from linear programming and cooperative game theory, particularly the theory of balanced collections and core existence.

We now investigate how much surplus the buyer retains in a competitive equilibrium. The buyer surplus U^* is the difference between joint value of information V_N and the sum of equilibrium prices:

$$U^* = V_N - \sum_{i=1}^{n} t_i^*.$$

Since all equilibrium outcomes are portfolio-efficient (the buyer purchases from every seller), this question then reduces to identifying the maximum revenue the sellers can charge in equilibrium.

Insufficiency of the Marginal Contributions Test. The duopoly analysis provided an elegant test for buyer surplus: compare the sum of individual marginal contributions, $\Delta V_1 + \Delta V_2$, to the total value V_N. When $\sum_i \Delta V_i < V_N$, competitive pressure from individual rivals ensures buyers retain surplus.

As Proposition 1 showed, this reduces to the simple complements-vs.-substitutes dichotomy:

$$U^* > 0 \quad \text{if and only if} \quad V_1 + V_2 > V_N.$$

The economic logic was transparent. Each seller's price was constrained by their individual marginal contribution—the loss in buyer value if that seller were excluded. If individual marginal contributions summed to less than total value, then competitive pressure from individual exclusion threats forced total prices below the buyer's value, preserving surplus.

The two-seller case suggests a natural generalization. Following the duopoly template exactly, we would test whether:

$$\sum_{i=1}^{n} \left(V_N - V_{N \setminus \{i\}} \right) \; < \; V_N. \tag{6}$$

This duopoly-inspired test asks: can oligopoly sellers extract more than their individual contributions to total value, just as in the two-seller case?

It is indeed easy to see that condition (6) is a sufficient condition for positive buyer surplus. Since the buyer's value is monotonic in coalition size, each seller's individual marginal contribution $V_N - V_{N \setminus \{i\}}$ provides the loosest possible constraint on their pricing. Any exposure to a smaller subset $S_i \subset N \setminus \{i\}$ would impose a tighter bound $V_N - V_{S_i} > V_N - V_{N \setminus \{i\}}$.

This duopoly-inspired test captures one important source of competitive pressure. But oligopoly reveals a fundamental difference: with $n \geq 3$, sellers face not just individual rivals but entire coalitions of alternatives. While duopoly sellers could only be threatened with individual exclusion, oligopoly sellers must defend against buyers switching to any subset of competitors; a form of competitive pressure that simply cannot exist when $n = 2$. However, this duopoly-inspired test has a critical limitation. For $n \geq 3$, sellers may be exposed to smaller sets of competitors, not just the grand coalition minus themselves. This means that even when the simple sum-of-marginals test fails, the buyer may still retain surplus due to coalition-based competitive pressure. The following example makes this concrete.

Example (Coalition-based threats)**.** *Consider three sellers and a buyer whose portfolio value is* *given by:*

$$V_{\{1,2,3\}} = 1,\ V_{\{2,3\}} = V_{\{1,3\}} = V_{\{1,2\}} = 0.5,\ V_{\{1\}} = V_{\{2\}} = V_{\{3\}} = 0.4,\ V_{\emptyset} = 0.$$

The individual marginal contributions to the full bundle sum to

$$(1-0.5)+(1-0.5)+(1-0.5) = 1.5 > 1.$$

In duopoly, marginal contributions summing to $1.5 > 1$ *would imply zero buyer surplus. However, the buyer-optimality (exposure) constraints for the oligopoly case* $n = 3$ *are:*

$$\begin{aligned} t_2 + t_3 &\leq V_{\{1,2,3\}} - V_{\{1\}} = 0.6, \\ t_1 + t_3 &\leq V_{\{1,2,3\}} - V_{\{2\}} = 0.6, \\ t_1 + t_2 &\leq V_{\{1,2,3\}} - V_{\{3\}} = 0.6, \\ t_1 + t_2 + t_3 &\leq V_{\{1,2,3\}} \ = \ 1. \end{aligned}$$

Summing the first three constraints gives

$$(t_1+t_3)+(t_1+t_2)+(t_2+t_3) = 2(t_1+t_2+t_3) \leq 1.8,$$

implying $t_1+t_2+t_3 \leq 0.9$. *Thus total seller revenue cannot exceed* 0.9 *in equilibrium. Since equilibrium portfolio-efficiency implies* $t_1+t_2+t_3 \leq V_N = 1$*, the buyer retains a surplus of at least* 0.1.

The duopoly sum-of-marginals test fails because oligopoly introduces coalition-based competitive pressure. The constraint $t_1+t_2 \leq 0.6$ *represents the buyer's threat to exclude seller 3 and purchase the portfolio* $\{1,2\}$. *This form of competition is absent in duopoly where only individual-exclusion threats exist.*

The example clarifies that the elementary tools that sufficed for duopoly become inadequate. Duopoly required checking just one inequality; oligopoly requires systematically analyzing exponentially many coalition-based threats. We now develop this more sophisticated condition through the mathematical theory of balancedness; a framework unnecessary in duopoly but essential for oligopoly surplus division.

Balancedness and Buyer Surplus. The failure of duopoly's simple marginal-contributions test reveals why oligopoly requires fundamentally different analytical tools. While duopoly needed to check only one inequality (V_1+V_2 vs V_N), oligopoly must systematically analyze exponentially many coalition-based threats. This complexity leads naturally to the theory of balanced collections.

The key insight is that competitive pressure from different coalitions must be systematically weighted to determine whether full surplus extraction is possible. A balanced collection assigns weights to each potential deviation coalition such that every seller is 'covered' by these threats in balanced proportion—no seller is over-threatened or under-threatened relative to others.

Definition 2 (Balanced collection)**.** *A collection of nonnegative weights $\{\gamma_S\}$ indexed by all non-empty, proper subsets $S \subset N$ is called a* balanced collection *if for each seller $i \in N$,*

$$\sum_{\substack{S \subset N \\ i \in S}} \gamma_S = 1.$$

When $n = 2$, this condition becomes trivial: with only subsets $\{1\}$ and $\{2\}$, any balanced collection must assign $\gamma_{\{1\}} = \gamma_{\{2\}} = 1$, recovering exactly the duopoly setup where each seller faces one exclusion threat. For $n > 2$, however, balanced collections capture the exponentially richer structure of coalition-based competitive pressure that emerges only in oligopoly.

With this framework, we can precisely characterize when full surplus extraction is achievable:

Definition 3 (Weak balancedness)**.** *The portfolio function V is* ***weakly balanced*** *for N if for every balanced collection $\{\gamma_S\}_{S \subsetneq N}$, the following holds:*

$$\sum_{S \subsetneq N} \gamma_S \left(V_N - V_{N \setminus S}\right) \geq V_N.$$

We now show that weak balancedness is both necessary and sufficient for the existence of a zero-surplus equilibrium.

Proposition 4. *The following are equivalent:*

(i) *There exists an equilibrium in which the buyer's surplus is zero.*

(ii) *The portfolio function V is weakly balanced.*

Proposition 4 characterizes when full extraction is achievable, but says nothing about whether it occurs in every equilibrium. The following example shows that when weak balancedness holds but additional structure is absent, multiple equilibria can coexist with different surplus levels.

Example (Multiplicity of Equilibria)**.** *Consider three sellers and a buyer whose portfolio value is given by:*

$$V_{\{1,2,3\}} = 1,\ V_{\{2,3\}} = V_{\{1,3\}} = V_{\{1,2\}} = 0.4,\ V_{\{1\}} = V_{\{2\}} = V_{\{3\}} = 0.3,\ V_{\emptyset} = 0.$$

The symmetric price vector $(1/3, 1/3, 1/3)$ *supports an equilibrium with zero buyer surplus. However, the asymmetric price vector* $(0.1, 0.1, 0.6)$ *also supports an equilibrium, yielding buyer surplus of* 0.2*. Sellers 1 and 2 compete intensely (each exposed by the other), while seller 3 exploits* *their rivalry to charge a high price. This demonstrates that weak balancedness alone does not guarantee universal extraction.*

Strong Balancedness and Universal Extraction. The previous example illustrates why weak balancedness is insufficient for universal full extraction. Weak balancedness restricts competitiveness only at the *aggregate* level: it guarantees that the grand coalition N can defend total revenue V_N against any proper subcoalition. What it does *not* regulate is the structure of competition *within* subcoalitions. If some subcoalition is itself highly competitive relative to its own sub-subcoalitions, sellers inside it may be forced to set prices strictly below their balancedness levels even though the grand coalition remains protected. This is exactly what happened in the example: the grand coalition was weakly balanced, but the subcoalitions imposed strong internal competitive pressure that allowed a positive-surplus equilibrium.

To ensure full extraction in every equilibrium, competitive pressure must be uniformly weak across all subcoalitions, not just at the top level. This leads to the following strengthening of balancedness:

Definition 4 (Strong balancedness)**.** *The portfolio function V is **strongly balanced** if it is weakly balanced for every subset $N' \subseteq N$.*

Strong balancedness requires that the balancedness inequality holds not just for the full set of sellers, but for every possible sub-market. This seemingly technical strengthening has profound implications: it turns out to be equivalent to a familiar economic property.

Proposition 5. *The following are equivalent:*

(i) *V is strongly balanced,*

(ii) *V is supermodular (i.e., $V_{S\cup T} + V_{S\cap T} \geq V_S + V_T$ for all $S, T \subseteq N$),*

(iii) *The marginal contribution function $G(S) := V_N - V_{N\setminus S}$ is submodular.*

This equivalence reveals that strong balancedness captures gross complementarity at all levels. When information sources are supermodular (gross complements everywhere), no coalition-based competitive pressure can emerge to sustain positive buyer surplus. Submodularity of marginal contributions means that larger coalitions provide weaker competitive threats, preventing the competitive stalemate observed in our three-seller example.

With this equivalence established, we can characterize universal extraction:

Proposition 6. *Suppose V is strongly balanced (or equivalently, supermodular). Then in every subgame perfect Nash equilibrium, the buyer has zero surplus.*

Our result recovers, in our information-market environment, the full-extraction result for convex games in Tauman et al. (1997) (and its reformulation in Arribas and Urbano (2005)). Their convexity assumption corresponds to supermodularity of V, and hence to strong balancedness in our terminology. Our argument is slightly different, and exploits a key property of supermodular value functions: if two coalitions bind exposure constraints (are "tight"), their union must also bind. This closure property forces all equilibria to achieve the same total revenue through a telescoping argument along chains of tight coalitions.

Combining our results, we obtain a complete characterization of surplus extraction in oligopoly:

- **Existence:** Full extraction is achievable (in some equilibrium) if and only if V is weakly balanced (Proposition 4).
- **Universality:** Full extraction occurs in every equilibrium if V is strongly balanced/supermodular (Propositions 5 and 6).

Whether the converse holds—zero surplus in all equilibria implies supermodularity—remains an open question. Our counterexample shows that weak balancedness without supermodularity allows multiplicity, but does not rule out the possibility that universal extraction always implies supermodularity.

For duopoly ($n = 2$), weak and strong balancedness coincide: both reduce to the complementarity condition $V_N \geq V_1 + V_2$. This explains why our duopoly analysis showed unique surplus division across all equilibria. For $n \geq 3$, the conditions diverge, with weak balancedness characterizing achievability and strong balancedness ensuring inevitability.

The Core and the Bondareva–Shapley Theorem. The balancedness characterization reveals an even deeper connection that highlights oligopoly's analytical complexity. While duopoly surplus division required only elementary analysis, oligopoly connects directly to fundamental results in cooperative game theory—a connection that simply cannot arise when $n = 2$.

This bridge between competitive pricing and cooperative stability demonstrates another dimension of complexity that duopoly concealed. We now explore this connection through the core and the Bondareva-Shapley Theorem.

A *transferable utility (TU) game* is defined by a portfolio value $v : 2^N \to \mathbb{R}$, where $v(S)$ gives the value that coalition $S \subseteq N$ can generate independently. The key insight is to define such a game where each coalition's value represents its marginal contribution to the buyer's information value:

$$v(S) = V_N - V_{N \setminus S}.$$

This transforms our competitive pricing problem into a question about whether sellers can form a 'stable' allocation of the total value they create. This construction differs from Arribas and Urbano (2005), who use the buyer's value function directly as the primitive. Our marginal-contribution formulation represents a change of scale that, combined with strict monotonicity, ensures the Bondareva-Shapley theorem always applies: the core is never empty, and balancedness conditions fully characterize buyer surplus.

Indeed, the *core* of this game is the set of allocations $(\phi_i)_{i\in N}$ satisfying:

- *Efficiency:* $\sum_{i\in N}\phi_i = v(N)$,
- *Coalitional rationality:* $\sum_{i\in S}\phi_i \geq v(S)$ for all $S\subseteq N$.

An allocation in the core is stable in the sense that no coalition has an incentive to deviate.

This cooperative framework becomes relevant because oligopoly's exponential constraint structure has the same mathematical form as core existence problems. In duopoly, such connections never arose because the constraint system was too simple to exhibit the geometric complexity that characterizes core theory. The *Bondareva–Shapley Theorem* (Bondareva, 1963; Shapley, 1967) provides a complete characterization of core-nonemptiness:

Theorem 1 (Bondareva–Shapley Theorem)**.** *The core of the transferable utility (TU) game* (N,v) *is nonempty if and only if for every* balanced collection $\lambda = (\lambda_S)$*,*

$$\sum_{S\subseteq N}\lambda_S v(S) \geq v(N).$$

This connection between competitive equilibrium and cooperative stability has no duopoly analogue. With $n=2$, the constraint system was elementary enough that core-theoretic concepts were unnecessary.

Combining Proposition 4 with the Bondareva-Shapley Theorem, we find that the buyer earns zero surplus if and only if the core of this TU-game is nonempty. This reframes buyer surplus as a signal of instability in an associated cooperative problem: when no balanced mixture of marginal contributions can replicate the grand coalition's value, the competitive constraint system lacks the geometric consistency needed for full surplus extraction. This connection illustrates how oligopoly analysis requires tools from entirely different areas of game theory that were unnecessary in duopoly. The simple complements-vs.-substitutes test gave way to balancedness conditions, which in turn connect to core existence—a hierarchy of increasing sophistication that reflects oligopoly's fundamental complexity.

For duopoly ($n=2$), weak and strong balancedness coincide: both reduce to the complementarity condition $V_N \geq V_1+V_2$. This explains why our duopoly analysis showed unique surplus division across all equilibria. For $n\geq 3$, the conditions diverge, with weak balancedness characterizing achievability and strong balancedness ensuring inevitability.

6 Symmetric Sellers

While the general oligopoly analysis captures the full complexity of information markets, it obscures several key economic insights about how portfolio competition differs from traditional product competition. In particular, we show that when all sellers are symmetric, oligopolistic competition may still completely fail to protect buyers against full rent extraction. The symmetric case, where sellers provide statistically identical information sources, not only yields sharper analytical results but also reveals fundamental economic forces that are harder to detect in the asymmetric setting. This case is empirically relevant for markets like financial data providers, credit rating agencies, or consulting firms within the same tier, where sellers offer similar-quality but independent information sources.

Moreover, the symmetric framework allows us to address two critical questions that the general analysis leaves open: (1) How does the endogenous market structure compare to the social optimum? and (2) How do information market dynamics differ from conventional oligopoly predictions?

The symmetric case is a setting in which the buyer's information value depends only on the number of sellers from whom she purchases signals, not on their specific composition. That is, for any seller set S, the value to the buyer is determined by the number of active sellers. Formally, the portfolio value V_S is symmetric if there is a function $\hat{V} : \{0,1,\dots,n\} \to \mathbb{R}$ such that

$$V_S = \hat{V}_{|S|}, \qquad \hat{V}_0 = 0, \quad \hat{V}_s \text{ is nondecreasing in } s \text{ and bounded.}$$

This symmetry assumption captures markets where sellers have achieved similar technological capabilities or market positions, but each provides independent draws from the same information structure. Unlike traditional symmetric oligopoly where firms sell identical products, information symmetry preserves the fundamental portfolio competition dynamic: buyers still benefit from purchasing multiple signals because independent sources reduce uncertainty even when they have identical statistical properties.

Our first insight is that, in a symmetric setting, the covering dual condition boils down to a simple and geometric interpretation. Recall the dual condition characterizing zero buyer surplus in equilibrium:

$$\sum_{S\subseteq N} \gamma_S \left(V_N - V_{N\setminus S}\right) \ \geq \ V_N \quad \text{for every balanced collection } \gamma.$$

In the symmetric case, a balanced collection assigns equal weight to all subsets of the same size. That is, for each $k = 1,\dots,n-1$, let γ_k denote the weight assigned to each subset of size k. Then, since there are $\binom{n}{k}$ such subsets, the total contribution of subsets of size k to the dual sum is $\binom{n}{k}\gamma_k(\hat{V}_n - \hat{V}_{n-k}) = b_k(\hat{V}_n - \hat{V}_{n-k})$, where $b_k = \binom{n}{k}\gamma_k$.

The balancedness condition imposes that the total weight falling on coalitions that include any fixed seller equals 1. In this symmetric case, this leads to the constraint:

$$\sum_{k=1}^{n-1} \binom{n-1}{k-1} \gamma_k = 1,$$

which, in terms of the b_k, becomes

$$\sum_{k=1}^{n-1} k\, b_k = n.$$

This has a natural interpretation: the total "mass" of weights distributed across coalitions must sum to n, when measured by how many sellers each coalition includes.

Now consider the objective of the dual problem:

$$\sum_{k=1}^{n-1} b_k \, (\hat{V}_n - \hat{V}_{n-k}).$$

Let us reparametrize the sum by substituting $j = n - k$, and define $c_j := b_{n-j}$. Then the objective becomes:

$$\sum_{j=1}^{n-1} c_j \, (\hat{V}_n - \hat{V}_j), \quad \text{where} \quad \sum_{j=1}^{n-1} (n-j) c_j = n.$$

That is, we are forming a convex combination of marginal contributions $\hat{V}_n - \hat{V}_j$, weighted by c_j, with the weights summing (in a shifted sense) to n. Since $\sum_j (n-j) c_j = n$, you can think of distributing "mass" n across the indices j, and then the weighted sum is minimized by pouring all the mass onto the j that makes $\left(\hat{V}_n - \hat{V}_j\right)/(n-j)$ smallest. Therefore:

$$\sum_{j=1}^{n-1} c_j (\hat{V}_n - \hat{V}_j) \;\geq\; \hat{V}_n \quad \text{for all such } c_j \qquad \Longleftrightarrow \qquad \frac{\hat{V}_n - \hat{V}_j}{n-j} \;\geq\; \frac{\hat{V}_n}{n} \quad \forall\, j,$$

which is equivalent to:

$$\hat{V}_j \;\leq\; \frac{j}{n} \hat{V}_n \quad \text{for all } j = 1, \ldots, n-1.$$

This is a purely geometric condition: the function $k \mapsto \hat{V}_k$ must lie on or below the straight line connecting $(0,0)$ and $(n, \hat{V}_n)$. In economic terms, this captures when early portfolio sizes provide relatively little value compared to the full portfolio, eliminating credible threats that would create competitive pressure. This pattern arises when portfolio value exhibits increasing returns (convexity) or mixed curvature with initially low marginal value, as illustrated in the convex and mixed cases (Panels 2 and 3 of Figure 2).

Hence, in the symmetric setting, the complex condition involving all balanced weights simplifies to a single geometric check: the buyer receives no surplus in any equilibrium if and only if the function $k \mapsto \hat{V}_k$ lies below the line segment connecting the origin to $(n, \hat{V}_n)$.

This geometric insight immediately suggests how to characterize equilibrium pricing. Since the buyer's optimal deviation determines competitive pressure, we can identify equilibrium prices by finding the coalition size that provides the buyer's best outside option.

Defining

$$k_n^* \equiv \arg\min_{k \in \{0,\ldots,n-1\}} \frac{\hat{V}_n - \hat{V}_k}{n-k} \text{ and } \bar{t}_n \equiv \frac{\hat{V}_n - \hat{V}_{k_n^*}}{n - k_n^*}$$

we obtain the following result.

Proposition 7. *Consider an information market with n symmetric firms. Then, the following holds:*

(i) *In a symmetric equilibrium, the unique fee per seller equals* $\bar{t}_n$.

(ii) *Sellers extract full surplus in the symmetric equilibrium if* $k_n^* = 0$.

(iii) *The sellers extract full surplus in every equilibrium if* $\hat{V}_k$ *is strictly convex.*

For part (i), the proof uses contradiction to show that any price below the critical slope allows sellers to profitably raise prices since the buyer still prefers the full portfolio, while any price above this slope makes the buyer want to switch to a smaller, more cost-effective portfolio size. The equilibrium price must therefore equal the smallest rate of value loss per excluded seller, which represents the buyer's most credible threat to reduce her information portfolio when facing symmetric pricing.

Part (ii) is a corollary: if $k_n^* = 0$, then the buyer's most attractive deviation from the full portfolio is to drop *all* sellers. In that case, the critical slope $\bar{t}_n$ coincides with the average value per seller in the grand coalition, so the symmetric equilibrium price satisfies $n\bar{t}_n = V_N$. Thus the buyer is exactly indifferent between buying from all sellers and walking away entirely, and the symmetric equilibrium leaves her with zero surplus while sellers extract the full value V_N.

For part (iii), note that with symmetric firms, supermodularity of V is equivalent to strict convexity of the sequence $(\hat{V}_k)_{k=0}^n$; that is, strict gross complementarity among signals corresponds exactly to $\hat{V}_{k+1} - \hat{V}_k$ being strictly increasing in k. Proposition 6 shows that supermodularity of V implies universal extraction: in every equilibrium, the buyer's surplus is zero and sellers capture the entire value V_N. Combining these observations, strict convexity of $(\hat{V}_k)$ is equivalent to supermodularity in the symmetric environment, and therefore implies that sellers extract full surplus in *every* equilibrium, not just in the symmetric one.

The proposition shows that, in the symmetric equilibrium, each seller gets $\bar{t}$ which corresponds to the minimum slope of any line segment connecting a point $(k, \hat{V}_k)$ to $(n, \hat{V}_n)$. Equivalently, the

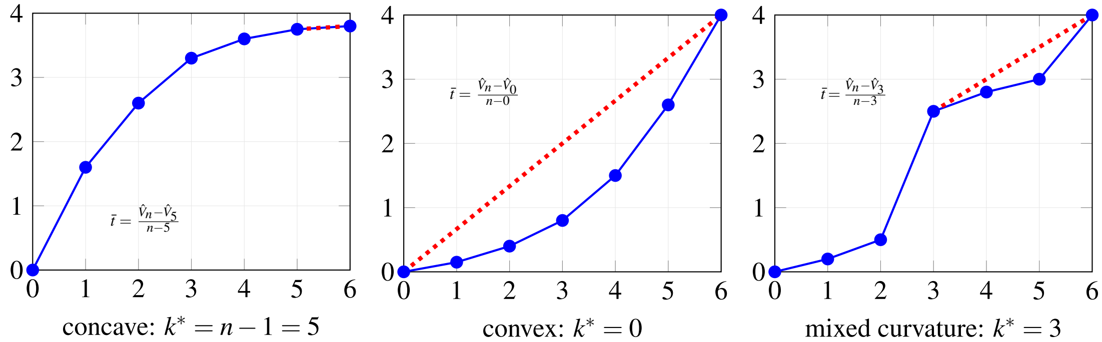

Figure 2: Symmetric equilibrium pricing when $n = 6$. The equilibrium price $\bar{t}_n$ equals the smallest slope from any point $(k, \hat{V}_k)$ to $(n, \hat{V}_n)$. The first panel illustrates the concave case, where this slope is smallest for $k^* = n - 1 = 5$. The second panel illustrates the convex case, where this slope is smallest for $k^* = 0$. The third panel illustrates the mixed case, where the smallest slope is in the interior.

buyer's threat is to drop from n to k^*, the point that gives the best average value per seller left out:

$$k_n^* = \arg\min_{k<n} \frac{\hat{V}_n - \hat{V}_k}{n-k}.$$

Figure 2 highlights the geometric representation, illustrating how k_n^* and $\bar{t}_n$ depend on the curvature of V_i and leads to the following corollary.

Corollary 1. *Suppose $\hat{V}_s$ is concave in s. Then, in any oligopolistic market outcome, the buyer receives positive rents. Suppose $\hat{V}_s$ is convex in s. Then, in any oligopolistic market outcome, the buyer does not receive any rents.*

The fact that the oligopolistic market outcome fully extracts the buyer's surplus reveals a fundamental difference between information and traditional product markets. In conventional oligopoly, symmetry typically leads to uniform marginal costs determining prices, leaving all rents to the buyer rather than the sellers. Here, the cost of providing information is zero, but the value depends on the buyer's entire information portfolio. The geometric condition captures when competitive pressure from alternative portfolio configurations can sustain full surplus extraction.

Example. *We return to our running example, which considers symmetric sellers. When the buyer purchases from $k \leq n$ sellers, let m denote the number of high signals, $s_i = h$, she observes. The posterior belief is:*

$$\hat{p}_m = \mathbb{P}(\omega = h \mid m) = \frac{1}{1 + \left(\frac{1-\alpha}{\alpha}\right)^{2m-k}}$$

The probability of observing m equals:

$$\mathbb{P}(m) = \frac{1}{2}\binom{k}{m}[\alpha^m(1-\alpha)^{k-m} + (1-\alpha)^m\alpha^{k-m}]$$

so that the information value from k sellers is:

$$\hat{V}_k = \sum_{m=0}^{k} \mathbb{P}(m)\cdot \hat{p}_m^2 - \frac{1}{4}$$

The portfolio value $\hat{V}_k$ is concave in k, exhibiting the substitute property $k\hat{V}_1 > \hat{V}_k$ for all $k \geq 2$. This generalizes the duopoly result where we verified $2V_1 > V_2$. The law of large numbers implies:

$$\lim_{k\to\infty} \hat{V}_k = \frac{1}{4},$$

showing that the buyer achieves the full information outcome, as the number of sellers approaches infinity.

By Proposition 7, concavity ensures $k_n^* = n-1$ for all n, and the buyer retains positive surplus. This illustrates how portfolio value curvature determines surplus division: concave values protect buyers through competitive pressure, while convex structures enable complete rent extraction.

7 Endogenous Entry

The analysis thus far has taken the number of information sellers as given. A complete theory of oligopolistic information markets requires endogenizing market structure: What determines the equilibrium number of firms? How does the combinability of information provision affect entry incentives?

To address these questions, we extend our model of oligopolistic information markets by endogenous entry. The subsequent model demonstrates that information markets depart from standard oligopoly theory with respect to entry. First, entry can reduce competitive pressure rather than intensify it, with the relationship between market structure and efficiency depending critically on the curvature of the portfolio value. Second, contrary to oligopolistic models which typically display excessive entry due to business stealing externalities, information markets exhibit no excessive entry—there is no business stealing, since the buyer's equilibrium portfolio choice is always socially efficient. While efficient entry is always an equilibrium outcome, additional equilibria with insufficient entry may exist due to coordination failures at the pricing stage, suggesting that regulatory interventions such as price regulation and entry subsidies could improve market efficiency.

Before developing the general framework, we illustrate with an extreme but straightforward example the result that entry in information markets can display a strong anti-competitive effect—in this example, entry results in the buyer losing all consumption rents.

Example (Anti-Competitive Entry). *Consider a market with two symmetric sellers where* $\hat{V}_1 = 10$*,* $\hat{V}_2 = 15$*. By Proposition 5, the equilibrium price is:* $t_2 = \min\{15/2, (15-10)/1\} = 5$*. At this price, the buyer purchases both signals, obtaining surplus* $15 - 2\cdot 5 = 5$*.*

Now suppose a third seller enters with $\hat{V}_3 = 36$*. The new equilibrium price becomes:* $t_3 = \min\{36/3, (36-10)/2, (36-15)/1\} = 12$*. At this price* t_3*, the buyer now obtains surplus* $36 - 3\cdot 12 = 0$*, despite having access to strictly more information.*

The example illustrates that entry can eliminate the buyer's surplus entirely while benefiting all sellers. This contradicts the standard intuition that entry intensifies competition and benefits buyers. The driving mechanism is complementarities in information provision: the entrant provides such strong complementarities that the buyer's best threat point shifts from excluding one seller (yielding $\hat{V}_2 = 15$) to excluding all sellers (yielding $\hat{V}_0 = 0$). As a result, the entrant weakens rather than strengthens competitive pressures. Empirically, such strong complementarities could arise when an entrant introduces novel information technology that amplifies the value of existing signals—for instance, an AI-powered analytics firm that not only provides its own data but enables buyers to extract far more value from previously purchased information through advanced pattern recognition and synthesis.

This phenomenon cannot arise in traditional product markets where goods are rivalrous and entry typically intensifies competition. It emerges precisely because of the combinability of information—buyers combine signals from multiple sources, creating portfolio competition where entry can paradoxically weaken competitive constraints. Understanding when and why such effects arise requires a systematic analysis of endogenous market structure.

7.1 The Entry Model

To analyze entry systematically, we consider a market with unlimited potential entrants. The natural framework extends our symmetric seller analysis by allowing the number of active sellers to be determined endogenously through entry decisions.

We maintain the symmetric structure of Section 6 where all sellers provide statistically identical but independent signals. The buyer's value from s active sellers is $\hat{V}_s$, which is strictly increasing and bounded:

$$\hat{V}_0 = 0 < \hat{V}_1 < \hat{V}_2 < \ldots < \hat{V}_\infty \equiv \lim_{s\to\infty} \hat{V}_s \in \mathbb{R}.$$

Each potential entrant must pay fixed cost $F > 0$ to enter the market. This cost captures the infrastructure required for data collection, processing, and distribution. The boundedness of V_∞ ensures that for sufficiently large s, the marginal value of additional sellers becomes negligible.

Extending earlier definitions, we define two key concepts. First, the marginal contribution

$$\Delta\hat{V}_s \equiv \hat{V}_s - \hat{V}_{s-1},$$

which measures the value added by the s-th seller. By the strict monotonicity of $\hat{V}_s$, marginal contributions are strictly positive. Second, the social surplus

$$\hat{S}_s \equiv \hat{V}_s - sF,$$

captures total welfare with a market entry of s sellers. The socially optimal number of sellers, s^* maximizes total surplus:

$$s^* = \arg\max_{s\geq 0} \hat{S}_s = \arg\max_{s\geq 0}\{\hat{V}_s - sF\}.$$

A necessary condition for s^* is the (discrete) first-order condition

$$\Delta\hat{V}_{s^*+1} \leq F \leq \Delta\hat{V}_{s^*}. \tag{7}$$

For a concave $\hat{V}_s$, this condition is also sufficient for s^* because of the differences $\Delta\hat{V}_s$ being decreasing. For this case, the usual optimality condition marginal value equals marginal cost determines s^*.

Finally, we extend the timing as presented in Section 3 with a prior stage in which sellers first decide whether to pay F to enter the information market:

1. *Entry.* Sellers simultaneously decide whether to pay F and enter
2. *Pricing.* Active sellers observe the number of entrants s and set prices
3. *Portfolio Selection.* The buyer observes prices and selects her information portfolio
4. *Information Realization and Payoffs.* Signals are realized and payoffs determined

This resulting framework models the following three aspects of entry. First, it captures competitive entry pressure through free entry with identical potential entrants. Second, it endogenizes market structure as an equilibrium outcome rather than an exogenous parameter. Third, it enables clean welfare comparisons between market outcomes and social optima by explicitly modeling the resource cost of entry.

7.2 Equilibrium Entry under Symmetric Pricing

We first analyze entry following Section 6 that with symmetric firms, the unique symmetric equilibrium price is:

$$\bar{t}_s = \min_{k\in\{0,\dots,s-1\}} \frac{\hat{V}_s - \hat{V}_k}{s-k}.$$

It directly follows that s^m represent an equilibrium number of sellers s^m if and only if it satisfies:

$$\bar{t}_{s^m+1} \leq F \leq \bar{t}_{s^m}. \tag{8}$$

The first inequality ensures that if an additional firm enters, it would not recoup its entry costs. The second inequality ensures that each entering firm recoups its entry costs, making entry a best response when expecting exactly $s^m - 1$ other firms to enter.

We first establish that the symmetric equilibrium price when entry is efficient, $\bar{t}_{s^*}$, exceeds the entry costs F. To see this, note first that the efficiency of s^* implies that for all s, the following two (equivalent) inequalities hold

$$\hat{V}_{s^*} - s^*F \geq \hat{V}_s - sF \Leftrightarrow \frac{\hat{V}_{s^*} - \hat{V}_s}{s^* - s} \geq F.$$

In particular, the latter inequality holds for all $k < s^*$, so that it also holds for the minimum among all $k < s^*$. We therefore obtain

$$\bar{t}_{s^*} = \min_k \frac{\hat{V}_{s^*} - \hat{V}_k}{s^* - k} \geq F.$$

Moreover, note that taking $s = s^* + 1$, the two inequalities reduce to

$$\hat{V}_{s^*+1} - \hat{V}_{s^*} \leq F.$$

In the words of the previous section, the slope of the line-segment from $(s^*, \hat{V}_{s^*})$ to $(s^*+1, \hat{V}_{s^*+1})$ lies below F. By contrast, all slopes of the line-segments from $(k, \hat{V}_k)$ to $(s^*, \hat{V}_{s^*})$ exceed F for all $k < s^*$, as already established. Taken together, this implies that of all the slopes of the line-segment from $(k, \hat{V}_k)$ to $(s^*+1, \hat{V}_{s^*+1})$, the smallest slope obtains for $k = s^*$. Hence,

$$k^*_{s^*+1} = s^* \text{ and } \bar{t}_{s^*+1} = \frac{\hat{V}_{s^*+1} - \hat{V}_{s^*}}{s^* + 1 - s^*} = \hat{V}_{s^*+1} - \hat{V}_{s^*} \leq F.$$

Thus we obtain

$$\bar{t}_{s^*+1} \leq F \leq \bar{t}_{s^*},$$

which by (8) implies that with symmetric pricing efficient entry is, independent of the curvature of

$\hat{V}_s$, an equilibrium outcome.

Now suppose $\hat{s} > s^*$ is not efficient, then this implies

$$\frac{\hat{V}_{\hat{s}} - \hat{V}_{s^*}}{\hat{s} - s^*} < F,$$

which leads to

$$\bar{t}_{\hat{s}} = \min_{k<\hat{s}} \frac{\hat{V}_{\hat{s}} - \hat{V}_k}{\hat{s} - k} \leq \frac{\hat{V}_{\hat{s}} - \hat{V}_{s^*}}{\hat{s} - s^*} < F.$$

So that if $\hat{s}$ would enter the market, the equilibrium price $\bar{t}_{\hat{s}}$ is too low for the firms to recoup their entry costs. As a result $\hat{s}$ cannot be part of an equilibrium with endogenous entry.

We thus have proven the following proposition

Proposition 8. *Suppose firms anticipate that when n symmetric firms enter, the symmetric equilibrium $\bar{t}_n$ obtains. Then efficient entry $(s^m = s^*)$ is an equilibrium, whereas excessive entry $(s^m > s^*)$ is not.*

Contrary to standard oligopoly models where business-stealing creates excessive entry, entry in information markets is therefore never excessive.

The proposition leaves open the possibility of insufficient entry as an equilibrium. We next argue that this possibility depends on the curvature of $\hat{V}_i$. Following our analysis in Section 5, the shape of $\hat{V}_s$ determines how equilibrium prices evolve with market size. If $\hat{V}_s$ is concave, then $\bar{t}_s = \Delta\hat{V}_s$ is decreasing in s. As a result, s^m is unique and follows the intuitive interpretation that sellers enter until the marginal entrant cannot recover the entry cost. By contrast, if $\hat{V}_s$ exhibits increasing returns initially, $\bar{t}_s$ may increase over some range. This may lead to non-monotonicity in pricing, generating multiple locally stable market structures.

Proposition 9. *Under symmetric pricing, the curvature of $\hat{V}_s$ determines equilibrium entry as follows: i) If $\hat{V}_s$ is strictly concave then entry is efficient: $s^m = s^*$. ii) If $\hat{V}_s$ is strictly convex, then no entry $(s^m = 0)$ is an equilibrium for $F > \hat{V}_1/1$, despite potentially large social value from entry. iii) If $\hat{V}_s$ has non-constant curvature, then entry may be insufficient ($s^m < s^*$) but not excessive $(s^m > s^*)$.*

7.3 Entry under Alternative Pricing Equilibria

The symmetric pricing assumption, while being based on a straightforward extension of Section 6 and yielding unique predictions, is not the unique subgame perfect equilibrium outcome of the entry game. We clarify this by examining how relaxing this assumption affects our results.

The entry game admits multiple subgame perfect equilibria through coordination on different pricing expectations. Consider any market size $\hat{s}$ satisfying:

$$\bar{t}_{\hat{s}} \geq F > \bar{t}_{\hat{s}+1}.$$

We can sustain $\hat{s}$ as an equilibrium by designating a set $\hat{S}$ with $\hat{s} = |\hat{S}|$ sellers, and focusing on the following pricing equilibrium in the subgames where a set S of seller enters: For any of these pricing subgames, the entering firms play the symmetric pricing equilibrium $\bar{t}_{|S|}$ except for the subgames with both $|S| = \hat{s}+1$ and $\hat{S} \subset S$. For these specific subgames, the $\hat{s}+1$ entering sellers play the asymmetric pricing equilibrium in which seller $s' \in S \backslash \hat{S}$ charges a price $t_{s'}$ below F. Under such equilibrium behavior, it is a subgame perfect equilibrium outcome of the overall free-entry game that exactly the sellers from set $\hat{S}$ and no other sellers $s \notin \hat{S}$ enter; a seller in $\hat{S}$ expects a non-negative payoff from entering, so that not entering, leading to a payoff of zero, is not a profitable deviation; a seller not in $\hat{S}$ receives zero from not entering, while expecting a negative payoff from the deviation to enter.

This multiplicity is not merely a technical curiosity. By contrast, it reflects genuine coordination challenges in information markets. Unlike manufacturing where marginal costs pin down a unique competitive price, information's zero marginal cost creates a coordination game with multiple Nash equilibria in the pricing subgame.

Hence, the consideration of alternative pricing equilibrium outcomes exacerbates inefficient entry.

We however next argue that alternative pricing equilibrium outcomes does not affect the impossiblity of excessive entry. To see this, consider any equilibrium with $\hat{s}$ active firms and asymmetric prices. Order the prices in descending order: $t_{\hat{s}}^1 \geq t_{\hat{s}}^2 \geq ... \geq t_{\hat{s}}^{\hat{s}}$. For the buyer to prefer purchasing from all $\hat{s}$ firms rather than excluding the r most expensive ones, it must hold that:

$$\sum_{i=1}^{r} t_{\hat{s}}^i \leq \hat{V}_{\hat{s}} - \hat{V}_{\hat{s}-r} \quad \text{for all } r \in \{1,...,\hat{s}\}$$

In equilibrium, at least one of these inequalities must bind (otherwise some seller could profitably raise their price). This implies that each of the top r prices satisfies:

$$\frac{1}{r}\sum_{i=1}^{r} t_{\hat{s}}^i \leq \frac{\hat{V}_{\hat{s}} - \hat{V}_{\hat{s}-r}}{r}$$

Now suppose $\hat{s} > s^*$, where s^* maximizes social surplus. From the efficiency of s^*, we have:

$$\hat{V}_{\hat{s}} - \hat{s}F < \hat{V}_{s^*} - s^*F$$

which can be rewritten as:

$$\frac{\hat{V}_{\hat{s}} - \hat{V}_{s^*}}{\hat{s} - s^*} < F$$

Setting $r = \hat{s} - s^*$ in the buyer's constraint yields:

$$\frac{1}{\hat{s} - s^*} \sum_{i=1}^{\hat{s}-s^*} t_{\hat{s}}^i \leq \frac{\hat{V}_{\hat{s}} - \hat{V}_{s^*}}{\hat{s} - s^*} < F$$

Therefore, the average of the top $\hat{s} - s^*$ prices is strictly less than F, which implies that at least one of these prices must be below F. But cost recovery requires $t_{\hat{s}}^i \geq F$ for all active firms i. This contradiction shows that $\hat{s} > s^*$ cannot be sustained as an equilibrium under any pricing arrangement. As a result $\hat{s} > s^*$ is also not sustainable as an entry equilibrium with asymmetric equilibrium prices.

In addition to this robustness result, we mention two further such results. First, the buyer purchases from all active sellers. This result follows directly from Proposition 2, which establishes that any equilibrium is portfolio-efficient. Second, the possibility of anti-competitive entry as illustrated in our example above is also robust. Whenever $\hat{V}_{s+1}/(s+1) > (\hat{V}_s - \hat{V}_k)/(s-k)$ for the relevant k, entry reduces buyer surplus under any equilibrium pricing.

By contrast, we mention that the following results are specific to symmetric pricing. First, efficiency under concavity. Alternative pricing equilibria can destroy the alignment between private and social incentives even with concave $\hat{V}_s$. Second, the monotone comparative statics in the entry cost F. With asymmetric pricing, increases in F might not monotonically reduce entry if sellers coordinate on different equilibria. Third, uniqueness of market structure. As argued, the multiplicity of locally stable points relies on the specific functional form of $\bar{t}_s$.

Finally, we emphasize that the multiplicity of equilibria has important welfare consequences, justifying regulatory intervention. Total surplus $\hat{V}_s - sF$ is maximized at s^*, but the market may coordinate on inefficient equilibria. A social planner could potentially improve outcomes through the following two means. First, the planner may use entry subsidies: When $\bar{t}_{s^*} < F < \Delta\hat{V}_{s^*}$, subsidies of $F - \bar{t}_{s^*}$ ensure efficient entry. Second, the planner may use pricing regulation. By mandating symmetric pricing, the planner eliminates asymmetric equilibria that deter efficient entry. However, a practical implementation of these tools requires the regulator to observe $\hat{V}_s$, which may be informationally demanding. Whenever regulators cannot observe such information directly, a proper analysis requires explicitly modeling firms possessing private information about the value of their signal vis-à-vis a regulator, which is beyond the scope of the current paper.

8 Conclusion

This paper develops a comprehensive framework for oligopolistic price competition in information markets, revealing fundamental differences from traditional product markets. The combinability of information sources transforms competition from classical Bertrand rivalry into portfolio competition, where sellers compete against all possible coalitions of rivals.

Contributions. Our analysis yields three core insights. First, we provide a complete characterization of when sellers can extract full surplus through weak and strong balancedness conditions, which generalizes the duopoly complements-versus-substitutes dichotomy to arbitrary numbers of sellers. Weak balancedness determines when full extraction is achievable in some equilibrium, while strong balancedness (equivalent to supermodularity) ensures it occurs in every equilibrium. When weak but not strong balancedness holds, multiple equilibria coexist with different surplus levels. Second, we demonstrate that information markets exhibit fundamentally different entry dynamics than traditional oligopoly: whereas standard models predict excessive entry due to business-stealing, information markets never exhibit excessive entry under any pricing equilibrium. Portfolio competition eliminates business-stealing, as buyers purchase from all (efficient) sellers in equilibrium. Third, for symmetric sellers, we derive a geometric characterization showing that universal surplus extraction occurs when portfolio value lies below the line connecting the origin to the full-market value, a condition arising when information sources exhibit convexity or mixed curvature patterns.

The endogenous entry analysis reveals striking departures from standard oligopoly theory. While efficient entry is always an equilibrium outcome, markets may coordinate on inefficient equilibria with insufficient entry. Moreover, entry can paradoxically reduce competitive pressure: when new sellers provide strong complementarities, they shift the buyer's threat point in ways that allow all sellers—incumbents and entrants alike—to extract higher prices. This anti-competitive effect of entry cannot arise in traditional markets with rivalrous goods.

Policy Implications. Our findings carry clear implications for policy and regulation. Traditional antitrust approaches focusing on market concentration may be misguided when applied to information markets. The number of competitors matters less than the structure of information complementarities. Many sellers may still fail to protect buyers when strong complementarities exist. Regulatory intervention should prioritize addressing coordination failures that lead to insufficient entry, with targeted tools such as entry subsidies or pricing regulation. Such implementation requires, however, that regulators can observe the portfolio value.

We also stress the importance of maintaining open access in information markets. Exclusivity clauses, where sellers condition access to their signals on buyers not purchasing from others, are especially harmful. Such restrictions block the formation of informative portfolios, undermine competition, and erode buyer surplus—nullifying the protective mechanisms portfolio competition can otherwise provide. Regulators should prohibit these practices to safeguard efficient information aggregation and market performance.

Scope of Applicability. While our framework is tailored to information markets, the portfolio competition structure extends to any market where buyers combine differentiated products from multiple sellers. The underlying mathematical apparatus—linear programming, cooperative core conditions, balancedness—applies whenever portfolio values are defined over subsets.

However, our universal surplus characterization relies critically on strict monotonicity of portfolio values, which naturally arises in information markets but not generally elsewhere. In markets for heterogeneous goods or platform services, non-monotonic values and perfect substitutes require the more general framework of Arribas and Urbano (2005), which introduces empty cores and revenue multiplicity and precludes complete surplus characterization.

Nonetheless, our analysis applies directly to examples such as non-overlapping software modules or independent consulting services, provided each seller's value is non-replicable. The economic structure, rather than generality, enables comprehensive surplus characterization: just as monopoly theory presupposes market power, our results presuppose differentiated sources.

Extensions. Our results open several research directions. Extending the model to heterogeneous buyers would clarify how diversity shapes information provision. Dynamic considerations—information obsolescence, learning effects, reputation—could reveal further dimensions of portfolio competition. The geometric surplus division predictions invite empirical tests in financial information services, credit rating agencies, and related markets. Relaxing strict monotonicity to allow redundant signals would require new equilibrium characterizations, broadening the relevance of the framework.

In summary, portfolio competition provides a robust foundation for analyzing modern markets where buyers combine multiple products, including platform markets, data services, and bundled goods. The interplay between competitive pricing and cooperative stability, governed by balancedness, is central in information economies. Explicitly prohibiting exclusivity clauses is essential to preserve these benefits. As information markets proliferate, understanding and regulating portfolio competition will be vital for both theory and policy.

Online Appendix – Not for Publication

Proof of Proposition 1: We split the argument into two parts. First, we show that no equilibrium can involve the buyer purchasing nothing or only one signal, so in any equilibrium (S^*,t^*) it holds that $S^* = N$. Second, we derive the necessary and sufficient conditions on prices t^* depending on whether signals are substitutes or complements.

1. *Ruling out "no signal" and "single signal" outcomes.*

   Suppose, by contradiction, that in some candidate equilibrium the buyer does *not* buy both signals.

   (a) *Case A: No signal.* Then both sellers earn zero profit at (t_1,t_2). But each seller i can profitably deviate by setting a sufficiently low price, say $\tilde{t}_i = V_i/2$. Because $V_i > 0$, this guarantees $V_i - \tilde{t}_i > 0$, so the buyer strictly prefers buying signal i alone to remaining with no signal. Hence, seller i's deviation in its price t_i yields strictly positive profit, contradicting the seller optimization condition in our equilibrium definition.

   (b) *Case B: A single signal from seller i.* Then seller $-i$ makes zero profit but can deviate to a price
   $$\tilde{t}_{-i} = \Delta V_{-i}/2 > 0.$$
   At $(t_i,\tilde{t}_{-i})$, the incremental surplus from adding signal $-i$ is
   $$(V_n - t_i - \tilde{t}_{-i}) - (V_i - t_i) = V_n - V_i - \tilde{t}_{-i} = \Delta V_{-i} - \tilde{t}_{-i} = \Delta V_{-i}/2 > 0,$$
   so the buyer strictly prefers purchasing both. Seller $-i$ thus secures a strictly positive profit of $\Delta V_{-i}/2$, contradicting the seller optimization condition for t_{-i} in our equilibrium definition.

   Hence, in any equilibrium (t^*,S^*), the buyer purchases *both* signals, $S^* = N = \{1,2\}$, and equilibrium prices $t^* = (t_1^*,t_2^*)$ satisfy (2).

2. *Equilibrium price characterization.* Having established that in any equilibrium (t^*,S^*), the equilibrium prices t^* satisfy
   $$t_1 + t_2 \leq V_n \ \wedge\ t_1 \leq \Delta V_1 \ \wedge\ t_2 \leq \Delta V_2, \tag{9}$$
   we now show that, in any equilibrium, the latter two inequalities must hold with equality under substitutes, whereas under complements, the first inequality must hold with equality while at most one of latter two inequalities holds with equality.

(a) *Substitutes:* $V_n \leq V_1 + V_2$. In this case, the latter two inequalities in (9) imply the first inequality in (9):

$$t_1 + t_2 \leq \Delta V_1 + \Delta V_2 = (V_n - V_2) + (V_n - V_1) = V_n + (V_n - V_1 - V_2) \leq V_n.$$

It follows that any price vector $t = (t_1, t_2)$ with one of the two inequalities in (9) slack, contradicts the seller optimization condition in our equilibrium definition, because the seller i for whom it holds $t_i < \Delta V_i$ has the profitable deviation to raise its price by $\varepsilon \in (0, \Delta V_i - t_i)$ so that his profits are raised strictly. Hence, under substitutes the equilibrium (S^*, t^*) is unique:

$$(S^*, t_1^*, t_2^*) = (N, \Delta V_1, \Delta V_2).$$

(b) *Complements:* $V_n \geq V_1 + V_2$. We first show that in this case, the latter two inequalities in (9) cannot both hold with equality, because this would violate the first inequality:

$$t_1 + t_2 = \Delta V_1 + \Delta V_2 = (V_n - V_1) + (V_n - V_2) = (V_n - V_1 - V_2) + V_n \geq V_n.$$

In other words, complements means that the sum of the marginal contributions exceeds the buyers value from purchasing both signals. Hence, in any equilibrium (S^*, t^*) we must have some seller i for whom it holds $t_i < \Delta V_i$. Fix this seller i. We next show that we then must have that the first inequality in (9) binds (holds with equality). For suppose not, then seller i can raise its price by $\varepsilon \in (0, \Delta V_i - t_i^*)$ so that (2) remains to holds, implying that the buyer buys both signals. This strictly raises seller i's profits, which would contradict seller i's optimization condition in our equilibrium definition.

Hence, any equilibrium (S^*, t^*) must exhibit $t_1 + t_2 = V_n$, $t_1 \leq \Delta V_1$, $t_2 \leq \Delta V_2$. To see that any such combination is indeed an equilibrium, note that given that seller i sets a price $t_i \leq \Delta V_i$, seller $-i$ best response is setting $t_{-i} = V_n - t_i$, which satisfies $t_{-i} \leq \Delta V_{-i}$ due to complementarity condition $V_n \geq V_1 + V_2$. Indeed, increasing t_{-i} beyond $V_n - t_i$ leads the buyer not to buy from seller $-i$, lowering the price lowers seller $-i$'s profits.

Hence, with complements, any equilibrium (S^*, t^*) satisfies $S* = N$ and $t_1^* + t_2^* = V_n$, and $t_1^* \leq \Delta V_1$ and $t_2^* \leq \Delta V_2$ with one of the latter two inequalities being slack. This implies that the set of equilibrium prices is a line segment on the budget line $t_1 + t_2 = V_n$, bounded by individual surplus constraints $t_i \leq \Delta V_i$.

□

Proof of Proposition 2: (*Only if*) We first show that if (S^*, t^*) is a subgame perfect Nash equilibrium, then (S^*, t^*) satisfies portfolio-efficiency ($S^* = N$) and t^* exhibits seller-exposure.

(i) Suppose to the contrary that $S^* \subsetneq N$. Then there exists at least one seller $j \notin S^*$ who has been excluded. By optimality of S^*, we then must have:

$$V_{S^* \cup \{j\}} - \sum_{i \in S^* \cup \{j\}} t_i^* \leq V_{S^*} - \sum_{i \in S^*} t_i^*,$$

otherwise the buyer would have preferred $S^* \cup \{j\}$. This inequality simplifies to:

$$t_j^* \geq V_{S^* \cup \{j\}} - V_{S^*} > 0.$$

Now suppose seller j lowers her price to:

$$\tilde{t}_j = (V_{S^* \cup \{j\}} - V_{S^*})/2 > 0.$$

With all other prices unchanged, the buyer now evaluates $S^* \cup \{j\}$ at:

$$V_{S^* \cup \{j\}} - \left(\tilde{t}_j + \sum_{i \in S^*} t_i^* \right) = V_{S^* \cup \{j\}} - (V_{S^* \cup \{j\}} - V_{S^*})/2 - \sum_{i \in S^*} t_i^* > V_{S^*} - \sum_{i \in S^*} t_i^*.$$

Thus the buyer strictly prefers $S^* \cup \{j\}$, and seller j earns $\tilde{t}_j > 0$. This is a profitable deviation, contradicting equilibrium. Hence, $S^* = N$.

(ii) Fix $i \in N$. Suppose, by contradiction, that t^* is such that no subset $S \subseteq N \setminus \{i\}$ satisfies the binding condition $\sum_{j \in N \setminus S} t_j^* = V_N - V_S$. The assumption implies that

$$\varepsilon \equiv \min_{S \subseteq N \setminus \{i\}} \left(V_N - V_S - \sum_{j \in N \setminus S} t_j^* \right),$$

is strictly positive.

Fix some $\delta \in (0, \varepsilon)$ and consider seller i deviating to price $\tilde{t}_i = t_i^* + \delta$.

For any subset $S \subseteq N \setminus \{i\}$, the buyer obtains $V_N - \sum_{j \in N \setminus \{i\}} t_j^* - (t_i^* + \delta)$ from purchasing N, whereas she obtains $V_S - \sum_{j \in S} t_j^*$ from purchasing S. Hence, the buyer prefers N to S if and only if:

$$V_N - \sum_{j \in N \setminus \{i\}} t_j^* - (t_i^* + \delta) \geq V_S - \sum_{j \in S} t_j^*,$$

which by rearranging is equivalent to

$$\delta \leq V_N - V_S - \sum_{j \in N \setminus S} t_j^*.$$

Since $\delta < \varepsilon$, and ε is the minimum over all such S, this inequality holds for all $S \subseteq N \setminus \{i\}$.

For any subset T containing seller i, the deviation increases the cost of both N and T by δ, leaving their relative attractiveness unchanged. Hence, the buyer continues to prefer N over T.

It follows that the deviation increases seller i's profit from t_i^* to $t_i^* + \delta > t_i^*$, while the buyer still chooses N. This contradicts the assumption that t_i^* was optimal for seller i.

Therefore, there must exist some $S_{-i} \subseteq N \setminus \{i\}$ such that the exposure condition binds.

(*If*) Conversely, suppose $(S^*, t^*) = (N, t^*)$ and t^* satisfies both full-bundle-optimality (3) and seller-exposure (4). Then (3) ensures the buyer weakly prefers to buy from all n sellers so that $S^* = N$ satisfies the equilibrium requirement of buyer optimization, while (4) ensures no seller can profitably raise prices so that t^* satisfies the equilibrium requirement of seller optimization. To see the latter, note that for any seller i, the binding constraint with S_{-i} ensures that raising t_i slightly would lead the buyer to no longer buy from i. Therefore, no profitable deviation exists, as all sellers sell their signals (as $S^* = N$) and deviating from their price t_i^* strictly lowers their profit. Hence, $(S^*, t^*) = (N, t^*)$ is a subgame perfect Nash equilibrium. □

Proof of Proposition 4: We prove the equivalence using LP duality.

(i) ⇒ (ii): Suppose there exists an equilibrium with zero buyer surplus. By Proposition 3, the revenue-maximizing LP achieves optimal value V_N:

$$\max_{t \geq 0} \sum_{i \in N} t_i \quad \text{subject to} \quad \sum_{i \in N \setminus S} t_i \leq V_N - V_S \quad \forall S \subsetneq N.$$

The dual LP is:

$$\min_{\gamma \geq 0} \sum_{S \subsetneq N} \gamma_S (V_N - V_S) \quad \text{subject to} \quad \sum_{S : i \notin S} \gamma_S \geq 1 \quad \forall i \in N.$$

By strong duality, the dual optimal value also equals V_N. Through the change of variables $\lambda_T = \gamma_{N \setminus T}$, the dual constraints become $\sum_{T : i \in T} \lambda_T \geq 1$. By complementary slackness, since all primal variables are strictly positive (due to strict monotonicity of V), the dual constraints bind at optimum: $\sum_{T : i \in T} \lambda_T = 1$, defining a balanced collection that achieves the dual minimum V_N. Since any balanced collection is feasible for the dual (as $=1$ implies ≥ 1), and the dual minimum is V_N, every balanced collection must satisfy $\sum_S \gamma_S(V_N - V_{N\setminus S}) \geq V_N$, establishing weak balancedness.

(ii) ⇒ (i): Suppose V is weakly balanced. Since all dual objective coefficients $V_N - V_{N\setminus T}$ are strictly positive (by strict monotonicity), any optimal dual solution can be taken to have all constraints binding, i.e., to be a balanced collection. Therefore, the dual LP minimum can be computed over balanced collections with $\sum_{T:i\in T} \lambda_T = 1$. Since every balanced collection satisfies

$$\sum_{S\subsetneq N} \gamma_S(V_N - V_{N\setminus S}) \geq V_N,$$

the dual optimal value is at least V_N. Since the primal LP has constraint $\sum_{i\in N} t_i \leq V_N$ (from $S = \emptyset$), its optimal value is at most V_N. By strong duality, both optima are equal, hence both equal V_N.

Therefore, there exists a feasible price vector t^* with $\sum_i t_i^* = V_N$ satisfying all buyer optimality constraints. By Proposition 3, this constitutes an equilibrium with buyer surplus $V_N - V_N = 0$. □

Proof of Proposition 5: We establish the equivalence through submodularity of complement-difference functions. For each $H \subseteq N$, define $G^H_S := V_H - V_{H\setminus S}$ for all $S \subseteq H$.

Step 1: V is supermodular $\iff$ G^H is submodular for every $H \subseteq N$.

Fix any $H \subseteq N$ and $T_1, T_2 \subseteq H$. Submodularity of G^H requires:

$$[V_H - V_{H\setminus T_1}] + [V_H - V_{H\setminus T_2}] \geq [V_H - V_{H\setminus(T_1\cup T_2)}] + [V_H - V_{H\setminus(T_1\cap T_2)}].$$

Simplifying:

$$2V_H - V_{H\setminus T_1} - V_{H\setminus T_2} \geq 2V_H - V_{H\setminus(T_1\cup T_2)} - V_{H\setminus(T_1\cap T_2)}.$$

Using set identities $H\setminus(T_1\cup T_2) = (H\setminus T_1)\cap(H\setminus T_2)$ and $H\setminus(T_1\cap T_2) = (H\setminus T_1)\cup(H\setminus T_2)$, this reduces to:

$$V_{H\setminus T_1} + V_{H\setminus T_2} \leq V_{(H\setminus T_1)\cap(H\setminus T_2)} + V_{(H\setminus T_1)\cup(H\setminus T_2)}.$$

Setting $A := H\setminus T_1$ and $B := H\setminus T_2$:

$$V_A + V_B \leq V_{A\cap B} + V_{A\cup B},$$

which is exactly supermodularity of V. Since H, T_1, T_2 were arbitrary, the equivalence holds.

Step 2: G^H submodular for all H $\implies$ V is strongly balanced.

Fix $H \subseteq N$. By assumption G^H is normalized ($G^H_\emptyset = 0$), nondecreasing, and submodular.

Consider the linear program:

$$\max\{\, w^\top x \,\Big|\, x \geq 0,\ x(U) \leq G^H(U)\ \forall U \subseteq H\,\}$$

and its dual:

$$\min\Big\{\sum_{T\subseteq H} y(T)\, g(T) \Big|\ y \geq 0,\ \sum_{T\ni i} y(T)\ \geq\ w_i\ \forall i \in H\Big\}.$$

Let $x^*(w)$ be the solution to the primal *LP*. Set $w = \mathbf{1}^H = (1,\ldots,1)^\top$. and fix any chain $\emptyset = U_0 \subset S_1 \subset \cdots \subset S_{|H|} = H$ with $|S_m| = m$. Since G^H is submodular, it follows from Theorem 44.3 in Schrijver (2003), that

$$x^*(\mathbf{1}^H) = G^H_{S_m} - G^H_{S_{m-1}} \quad (m = 1,\ldots,|H|).$$

and the value of the primal LP is equal to the value of the dual LP. From the primal LP, we further obtain that

$$(\mathbf{1}^H)^\top x^*(\mathbf{1}^H) = \sum_{m=1}^{|H|} \big(G^H_{S_m} - G^H_{S_{m-1}}\big) = G^H_H - G^H_\emptyset = G^H_H,$$

which says that the primal LP equals $G^H(H)$ and hence the dual minimum (over fractional covers of H) equals $g(H)$:

$$\min\Big\{\sum_{T\subseteq H} y(T)\, g(T) :\ y \geq 0,\ \sum_{T\ni i} y(T)\ \geq 1\ \forall i \in H\Big\}\ =\ g(H).$$

This is precisely the balancedness inequality on H.

Step 3: V is strongly balanced $\Longrightarrow$ G^H is submodular on all H.

Fix H, and define the dual minimum

$$D^H(r)\ :=\ \min_{\gamma\geq 0}\Big\{\sum_{T\subseteq N} \gamma_T\, G^H(T)\ :\ \sum_{T\ni i}\gamma_T\ \geq\ r_i\ \ \forall i \in N\Big\}.$$

The dual minimum D^H satisfies the following properties which are immediate from the definition:

1. *Monotonicity:* if $r' \leq r$ coordinatewise, then $D^H(r') \leq D^H(r)$.
2. *Superadditivity:* for any $r, s \in \mathbb{R}^N_+$,

$$D^H(r+s)\ \geq\ D^H(r) + D^H(s). \tag{10}$$

Monotonicity holds, because relaxing the constraint can only reduce the dual minimum. Superadditivity holds because any collection of weights γ that feasible for $r+s$ is also feasible for r and for s.

Now, fix any two sets $A,B \subseteq N$. Using the identity of indicator vectors

$$\mathbf{1}^A + \mathbf{1}^B \;=\; \mathbf{1}^{A\cup B} + \mathbf{1}^{A\cap B},$$

we bound $D^H(\mathbf{1}^A + \mathbf{1}^B)$ from above and below.

- *Upper bound.* Take $\gamma_A = 1$, $\gamma_B = 1$, and $\gamma_T = 0$ otherwise. Then $\sum_{T\ni i} \gamma_T = (\mathbf{1}^A + \mathbf{1}^B)_i$ for every i, so γ is feasible for $\mathbf{1}^A + \mathbf{1}^B$, with cost

  $$\sum_T \gamma_T G^H(T) \;=\; G^H(A) + G^H(B).$$

  Hence

  $$D^H(\mathbf{1}^A + \mathbf{1}^B) \;\leq\; G^H(A) + G^H(B). \tag{11}$$

- *Lower bound.* By superadditivity (10) and the indicator identity,

  $$D^H\big(\mathbf{1}^A + \mathbf{1}^B\big) \;=\; D^H\big(\mathbf{1}^{A\cup B} + \mathbf{1}^{A\cap B}\big) \;\geq\; D^H(\mathbf{1}^{A\cup B}) + D^H(\mathbf{1}^{A\cap B}).$$

  Since V is strongly balanced, it is weakly balanced on both $A\cup B$ and $A\cap B$, and hence

  $$D^H(\mathbf{1}^A + \mathbf{1}^B) \;\geq\; G^H(A\cup B) + G^H(A\cap B). \tag{12}$$

Combining (11) and (12) gives

$$G^H(A) + G^H(B) \;\geq\; D(\mathbf{1}^A + \mathbf{1}^B) \;\geq\; G^H(A\cup B) + G^H(A\cap B),$$

which is exactly the definition of submodularity of G^H on H.

□

Proof of Proposition 6: Fix any subgame perfect Nash equilibrium with price vector $t = (t_1,\ldots,t_n)$. Define the complement-difference function on the grand market:

$$G_S := V_N - V_{N\setminus S} \quad \text{for all } S \subseteq N.$$

By Proposition 5, strong balancedness implies $G = G^N$ is submodular. From the equilibrium characterization:

- *Buyer-optimality:* For all $S \subseteq N$, $\sum_{i \in S} t_i \leq G_S$.
- *Exposure:* For each seller $i \in N$, there exists $S_i \subseteq N$ with $i \in S_i$ such that $\sum_{j \in S_i} t_j = G_{S_i}$.

For any set S with $\sum_{j \in S} t_i = G_S$, we say S is *tight*.

Step 1: Tight sets are closed under union. Let $A, B \subseteq N$ be tight sets: $\sum_{i \in A} t_i = G(A)$ and $\sum_{i \in B} t_i = G(B)$. By submodularity of G:

$$G(A) + G(B) \geq G(A \cup B) + G(A \cap B).$$

By inclusion-exclusion for the additive function $t(\cdot)$:

$$\sum_{i \in A} t_i + \sum_{i \in B} t_i = \sum_{i \in A \cup B} t_i + \sum_{i \in A \cap B} t_i.$$

Combining these:

$$\sum_{i \in A \cup B} t_i + \sum_{i \in A \cap B} t_i \geq G(A \cup B) + G(A \cap B).$$

By buyer-optimality (feasibility):

$$\sum_{i \in A \cup B} t_i \leq G(A \cup B) \quad \text{and} \quad \sum_{i \in A \cap B} t_i \leq G(A \cap B).$$

Therefore both inequalities are equalities, so $\sum_{i \in A \cup B} t_i = G(A \cup B)$.

Thus $A \cup B$ is tight. By induction, the union of finitely many tight sets is tight.

Step 2: Construct a chain of tight sets. By exposure, for each seller $i \in N$, there exists a tight set T_i with $i \in T_i$. Pick any ordering $i_1, \ldots, i_n$ of the sellers and define:

$$S_k := \bigcup_{r=1}^{k} T_{i_r} \quad \text{for } k = 1, \ldots, n, \quad S_0 := \emptyset.$$

By Step 1, each S_k is tight. Moreover:

- $S_0 \subseteq S_1 \subseteq \cdots \subseteq S_n = N$ (nested structure)
- Each S_k satisfies $\sum_{i \in S_k} t_i = G(S_k)$ (tight)
- $S_n = \bigcup_{i \in N} T_i = N$ (since each $T_i \ni i$)

Removing redundancies where $S_k = S_{k-1}$, we obtain a strictly increasing chain:

$$\emptyset = S_0 \subset S_1 \subset \cdots \subset S_m = N$$

where each S_j is tight and $m \leq n$.

Step 3: Telescope along the chain. For each $j \in \{1,\ldots,m\}$, define the block $B_j := S_j \setminus S_{j-1}$. Since S_{j-1} and S_j are both tight:

$$\sum_{i \in B_j} t_i = \sum_{i \in S_j} t_i - \sum_{i \in S_{j-1}} t_i = G(S_j) - G(S_{j-1}).$$

Summing over all blocks:

$$\sum_{i \in N} t_i = \sum_{j=1}^{m} \sum_{i \in B_j} t_i = \sum_{j=1}^{m} [G(S_j) - G(S_{j-1})].$$

This is a telescoping sum:

$$= G(S_m) - G(S_0) = G(N) - G(\emptyset) = (V_N - V_\emptyset) - (V_N - V_N) = V_N.$$

Therefore, the total revenue equals V_N, so the buyer's surplus is zero. □

Proof of Proposition 7:

(i) Suppose to the contrary that there is a symmetric equilibrium at fee $t \neq \bar{t}$.

- *If $t < \bar{t}$, sellers want to raise their fee.* Let $k^* \in \arg\min_{k<n} (\hat{V}_n - \hat{V}_k)/(n-k)$. Since

  $$t < \bar{t} \; = \; \frac{\hat{V}_n - \hat{V}_{k^*}}{n - k^*},$$

  we have
  $$(n-k^*)t \; < \; \hat{V}_n - \hat{V}_{k^*} \quad \Longrightarrow \quad \hat{V}_n - nt \; > \; \hat{V}_{k^*} - k^* t.$$

  Thus the buyer strictly prefers the full n-seller bundle at price t over the k^*-seller bundle at the same price t. In particular, there exists $t' \in (t, \bar{t})$ such that if one seller unilaterally increases her fee to some t', the buyer still strictly prefers to keep all n. That seller's profit rises from t to t', a profitable deviation. This deviation is profitable because the buyer's portfolio choice remains unchanged (she still purchases from all n sellers), so the deviating seller's revenue increases while all other sellers' revenues remain constant. Hence $t < \bar{t}$ cannot be an equilibrium.

- *If $t > \bar{t}$, the buyer wants to drop sellers.* Again let $k^* \geq 0$ be the smallest k that minimizes $(\hat{V}_n - \hat{V}_k)/(n-k)$. Now

$$t > \bar{t} = \frac{\hat{V}_n - \hat{V}_{k^*}}{n - k^*} \implies \hat{V}_n - nt < \hat{V}_{k^*} - k^* t.$$

Thus the buyer strictly prefers contracting only with k^* sellers at fee t rather than all n. She would refuse the n-bundle, breaking the putative equilibrium.

Combining (i) and (ii), the only possible symmetric equilibrium fee is $t = \bar{t}$.

(ii) Omitted.

(iii) Assume $(\hat{V}_k)_{k=0}^n$ is convex in the sense that

$$\hat{V}_{k+2} - \hat{V}_{k+1} \geq \hat{V}_{k+1} - \hat{V}_k \quad \text{for all } k = 0, \ldots, n-2.$$

Fix integers $k \geq \ell$ and $j \geq 0$. Since the increments are nondecreasing in the index, we have

$$\hat{V}_{k+j} - \hat{V}_{\ell+j} = \sum_{r=0}^{k-\ell-1} \left(\hat{V}_{\ell+j+r+1} - \hat{V}_{\ell+j+r}\right) \geq \sum_{r=0}^{k-\ell-1} \left(\hat{V}_{\ell+r+1} - \hat{V}_{\ell+r}\right) = \hat{V}_k - \hat{V}_\ell.$$

Thus, convexity of $\hat{V}_k$ implies that

$$\hat{V}_{k+j} + \hat{V}_\ell \geq \hat{V}_{\ell+j} + \hat{V}_k. \qquad (*)$$

To show V is supermodular, we must prove that for all $S, T \subseteq N$,

$$V_{S \cup T} + V_{S \cap T} \geq V_S + V_T.$$

By symmetry of V, this is equivalent to

$$\hat{V}_{|S \cup T|} + \hat{V}_{|S \cap T|} \geq \hat{V}_{|S|} + \hat{V}_{|T|}. \qquad (\dagger)$$

Now, in inequality $(*)$, set

$$k = |T|, \qquad \ell = |S \cap T|, \qquad j = |S| - |S \cap T| \geq 0.$$

Substituting these into $(*)$ and, using that $|S \cup T| = |S| + |T| - |S \cap T|$, yields exactly the definition of supermodularity $(\dagger)$ when V is symmetric. The claim now follows directly from Proposition (6).

□

Proof of Proposition 9: Under strict concavity, Proposition 7 implies $k^* = s-1$ for all s, yielding:

$$\bar{t}_s = \frac{\hat{V}_s - \hat{V}_{s-1}}{1} = \Delta\hat{V}_s$$

Therefore, sellers enter if $\Delta\hat{V}_s > F$ and stop entering if $\Delta\hat{V}_{s+1} < F$, which coincides with the social optimality condition (7), which for a concave $\hat{V}_s$ is also sufficient.

Under strict convexity, $k^* = 0$ for all s, yielding $\bar{t}_s = \hat{V}_s/s$. Entry occurs when $\hat{V}_s/s \geq F$. But social efficiency requires $\Delta\hat{V}_s \geq F$. Since convexity implies $\Delta\hat{V}_s > \hat{V}_s/s$ for all s, there exist parameter regions where entry is socially valuable ($\Delta\hat{V}_s > F$) but unprofitable ($\hat{V}_s/s < F$).

With mixed curvature, $\bar{t}_s$ may exceed or fall short of $\Delta\hat{V}_s$ depending on the location of curvature changes, creating divergence between private and social entry incentives. □

References

Acemoglu, D., A. Makhdoumi, A. Malekian, and A. Ozdaglar (2022, November). Too much data: Prices and inefficiencies in data markets. *American Economic Journal: Microeconomics 14*(4), 218–56.

Admati, A. R. and P. Pfleiderer (1986). A monopolistic market for information. *Journal of Economic Theory 39*(2), 400–438.

Ali, S. N., N. Haghpanah, X. Lin, and R. Siegel (2022). How to sell hard information. *The Quarterly Journal of Economics 137*(1), 619–678.

Arribas, I. and A. Urbano (2005). Nash equilibria in a model of multiproduct price competition: an assignment problem. *Journal of Mathematical Economics 41*(3), 351–385.

Bergemann, D., A. Bonatti, and T. Gan (2022). The economics of social data. *The RAND Journal of Economics 53*(2), 263–296.

Bergemann, D., A. Bonatti, and A. Smolin (2018). The design and price of information. *The American Economic Review 108*(1), 1–48.

Bondareva, O. N. (1963). Some applications of linear programming methods to the theory of cooperative games. *Problemy Kibernetiki 10*, 119–139.

Börgers, T., A. Hernando-Veciana, and D. Krähmer (2013). When are signals complements or substitutes? *Journal of Economic Theory 148*(1), 165–195.

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

DeMarzo, P. M., I. Kremer, and A. Skrzypacz (2019, June). Test design and minimum standards. *American Economic Review 109*(6), 2173–2207.

Dixit, A. K. and J. E. Stiglitz (1977). Monopolistic competition and optimum product diversity. *American Economic Review 67*(3), 297–308.

Gentzkow, M. and E. Kamenica (2017a). Bayesian persuasion with multiple senders and rich signal spaces. *Games and Economic Behavior 104*, 411 – 429.

Gentzkow, M. and E. Kamenica (2017b). Competition in persuasion. *Review of Economic Studies 84*(1), 300 – 322.

Li, F. and P. Norman (2021). Sequential persuasion. *Theoretical Economics 16*(2), 639–675.

Lizzeri, A. (1999). Information revelation and certification intermediaries. *The RAND Journal of Economics 30*(2), 214–231.

Mankiw, N. G. and M. D. Whinston (1986). Free entry and social inefficiency. *The RAND Journal of Economics 17*(1), 48–58.

McAfee, R. P. (1993). Mechanism design by competing sellers. *Econometrica 61*(6), 1281–1312.

Raiffa, H. and R. Schlaifer (1961). *Applied statistical decision theory*. Division of Research, Graduate School of Business Administration, Harvard University.

Raith, M. (1996). A general model of information sharing in oligopoly. *Journal of Economic Theory 71*(1), 260–288.

Shapley, L. S. (1967). On balanced sets and cores. *Naval Research Logistics Quarterly 14*(4), 453–460.

Singh, N. and X. Vives (1984). Price and quantity competition in a differentiated duopoly. *RAND Journal of Economics 15*(4), 546–554.

Spence, A. M. (1976). Product selection, fixed costs, and monopolistic competition. *Review of Economic Studies 43*(2), 217–235.

Stahl, K. and R. Strausz (2017). Certification and market transparency. *Review of Economic Studies 84*(4), 1842–1866.

Tauman, Y., A. Urbano, and J. Watanabe (1997). A model of multiproduct price competition. *Journal of Economic Theory 77*(2), 377–401.

Vives, X. (1984). Duopoly information equilibrium: Cournot and bertrand. *Journal of Economic Theory 34*(1), 71–94.

Vives, X. (1988). Aggregation of information in large cournot markets. *Econometrica 56*(4), 851–876.

Wu, W. (2023). Sequential bayesian persuasion. *Journal of Economic Theory 214*, 105763.

Yang, K. H. (2022). Selling consumer data for profit: Optimal market-segmentation design and its consequences. *The American Economic Review 112*(4), 1364–1393.