Consumer Learning and Competitive Inefficiency

Woo-Jin Kim*

November 14, 2026

Abstract

We study price competition in a differentiated-product duopoly in which a consumer observes prices before deciding whether to acquire costly information about which product fits her better. Prices shape the value of learning and, in turn, the consumer's product choice. We characterize the unique mixed-strategy pricing equilibrium in which both firms are active. In equilibrium, sufficiently large price differences override product-fit information, generating ex post inefficiency even when information is free—a distortion absent under monopoly. How this inefficiency changes with information costs depends on market asymmetry: no learning maximizes total surplus when the allocation gain from favoring the advantaged firm outweighs the matching gain from learning. Otherwise, free information is optimal. Higher information costs initially reduce learning-induced product differentiation, intensify price competition, and raise consumer surplus. Once learning ceases, however, the advantaged firm captures the market, and its market power increases as information costs rise further, lowering consumer surplus. Finally, we endogenize the information friction and show that the firm never underprovides product-fit transparency relative to the social optimum.

1 Introduction

Consumers often face uncertainty about which competing product best fits their needs, even when they know the prices. A shopper comparing smartphones, for example, may see their prices at a glance, but evaluating their camera performance, battery life, or compatibility with her other devices takes time and attention. A natural intuition is that when information

*University of Southern California. Email: woojinki@usc.edu. I am deeply indebted to Guofu Tan for his invaluable guidance, continuous support, and encouragement throughout this project. I thank Odilon Camara, Alexandre de Corniere, Doh-Shin Jeon, Youngwoo Koh, Nicolas Lambert, Jonathan Libgober, Junha Park, Joao Ramos, Andrew Rhodes, Christopher Teh, and Mark Whitmeyer, as well as participants at the 37th Stony Brook International Conference on Game Theory, UCLA Student Theory Conference, and USC Theory Seminar, for their helpful comments. I am also grateful for the hospitality and valuable discussions during my visit to the Toulouse School of Economics.

is easier to get (e.g., through comparison tools or recommendation systems), consumers make better choices and are better off. Yet learning also changes how firms compete. After discovering that one product is a better match, the consumer may be willing to pay more for it rather than switch to a cheaper alternative. Learning can therefore make the products less substitutable and allow firms to sustain higher prices. This paper asks whether easier access to product information necessarily improves welfare in competitive markets.

To address this question, we study price competition in a differentiated-product duopoly with costly consumer learning. The two firms simultaneously set prices, which the consumer observes before deciding whether to incur an information cost to learn the relative match value between the two products. The relative match value is binary: Firm 1's product is the better match in one state and Firm 2's in the other. We allow for market asymmetry along two dimensions: how often each product is the better match and by how much. We refer to the firm with a weak ex ante advantage as Firm 1, meaning that its product is at least as good on average. After the learning decision, the consumer chooses between the two products under market coverage.

A key feature of this setting is that the benefit of learning depends on the price difference. When neither product has a decisive price advantage, the consumer may find it worthwhile to learn which product fits her better, despite the cost. When one product is sufficiently cheaper, however, learning becomes less valuable. The consumer's optimal learning and purchase decisions thus yield a demand function with two cutoffs in the price difference: between these cutoffs, she learns and purchases the better-matched product in each state; outside this range, she chooses one product without learning.[1]

We characterize the unique mixed-strategy pricing equilibrium when both firms are active. A key technical challenge is establishing the structure of price supports under market asymmetry; we show that both firms mix over connected intervals whose common length equals the distance between the two demand cutoffs. The intuition behind firms' mixing lies in two competing pricing incentives created by this demand structure. When the consumer learns, each firm retains positive demand in its favorable state, allowing it to charge a relatively high price. By setting a sufficiently low price relative to its rival, however, a firm can induce the consumer to buy from it without learning and capture the entire market. These incentives lead firms to randomize over two price branches: a lower-price branch that ensures positive demand and an upper-price branch that offers higher margins at the risk of losing the entire market.

The competitive equilibrium generates ex post inefficient product choices, even when information is free. Ex post efficiency requires the consumer to choose the better-matched

[1] We also find a related pattern in e-commerce search data, where consumers are more likely to inspect products when displayed prices are closer together. A one-standard-deviation decrease in within-search price differences is associated with a 1.66-percentage-point increase in the probability of product inspection. See Appendix D for data construction, empirical specifications, and regression results.

product in each state. In equilibrium, however, price dispersion generates some price differences large enough that the consumer chooses the same product regardless of the realized match value. Even a fully informed consumer can therefore choose the worse-matched product because of its lower price. By contrast, the monopoly benchmark yields a different result. As information costs vanish, the monopolist finds it profitable to charge the highest price at which the consumer is willing to learn, and the allocation approaches the first best. This highlights that the inefficiency arises from strategic price competition, not simply from the cost of information.

We next show that higher information costs increase consumer surplus while both firms remain active. As learning becomes more costly, the consumer becomes more willing to choose based on relative prices without learning which product is the better match. This makes the products closer substitutes and intensifies price competition. Both firms' equilibrium price distributions shift downward in the sense of first-order stochastic dominance, and their profits decline. More strongly, we show that both distributions continue to shift downward even after adding the cost of learning to prices. The consumer therefore benefits from stronger price competition whether she learns or not.

Lower prices benefit the consumer but do not directly increase total surplus. What matters is how information costs affect learning and product choice through equilibrium pricing. When neither firm has an ex ante advantage, a higher information cost leaves the learning probability and allocation unchanged, as the distribution of equilibrium price differences contracts at the same rate as the learning region. Hence, it only reduces product differentiation without favoring either firm. When Firm 1 has a strict ex ante advantage, however, it becomes easier for Firm 1 to induce purchases without learning. Firm 1 therefore puts more weight on its lower-price branch. In response, Firm 2 finds aggressive price cuts less attractive and instead shifts probability toward its upper-price branch. As a result, the consumer becomes more likely to purchase from Firm 1 without learning.

This shift in demand can improve total surplus because Firm 1's product is better on average. However, Firm 2 can still be the better match, and learning helps the consumer choose it when it is. Thus, total surplus depends on the gains from shifting demand toward Firm 1 (allocation effect) and from choosing the better-matched product (matching effect), net of learning costs. We find that total surplus is maximized either when information is free or at the threshold where the consumer stops learning and purchases from Firm 1. Intermediate information costs are never optimal because they incur learning costs without providing sufficient allocation or matching gains. Which outcome is better depends on market asymmetry. When the two products are ex ante similar, learning helps the consumer find the better match, and free information is optimal. In particular, without an ex ante advantage, higher information costs generate no allocation gain and only reduce total surplus. When Firm 1's advantage is sufficiently large, however, the allocation gain from no learning can exceed the matching gain from learning, making the no-learning outcome optimal.

When the information cost reaches the threshold at which the consumer stops learning, the equilibrium enters a quasi-monopoly regime: Firm 1 serves the entire market, while Firm 2 earns zero. Firm 1 charges a single price, the highest at which it can keep the consumer uninformed and buying from it. As long as learning remains possible, competitive pressure from Firm 2 and the consumer's ability to learn constrain Firm 1's price below its prior advantage. Nevertheless, higher information costs strengthen Firm 1's dominant position, allowing it to raise its price without inducing learning and thereby reducing consumer surplus. Total surplus, however, remains constant because the allocation does not change.

The main takeaway is that the welfare effects of consumer learning depend on firms' pricing responses. Empirical studies of online hotel-booking platforms find that personalized recommendations can steer consumers toward better-matched hotels and soften price competition, thereby reducing consumer surplus (Minaev (2021); Kaye (2026)). However, their results on total surplus differ, as Minaev (2021) finds an increase whereas Kaye (2026) finds a decrease. Our analysis provides a mechanism that helps explain these findings: lower information costs soften price competition, while their effect on total surplus depends on market asymmetry. These results suggest that policies facilitating access to product-fit information should account for firms' pricing incentives, market asymmetry, and the policy objective.[2]

Our findings also highlight the limitations of market shares as indicators of competitive pressure in antitrust analysis.[3] When both firms are active, higher information costs increase the advantaged firm's market share even as price competition intensifies and consumer surplus rises. The quasi-monopoly regime further illustrates this disconnect: although the rival makes no sales in equilibrium, it still constrains the advantaged firm's pricing. As information costs rise further, this competitive pressure weakens even though market shares remain unchanged.

Interestingly, we show that Firm 1 prefers the competitive market under free information to any quasi-monopoly outcome, even though it captures all industry profit in the latter. This raises the question of how much transparency Firm 1 would provide if it could influence the cost of consumer learning at its own expense. Gladia, a provider of speech-to-text APIs, sponsors a free tool that allows potential customers to upload their own audio and compare transcriptions from different providers, including Gladia and its competitors, before purchasing a service. Gladia covers the cost of these comparisons, including the fees charged by competing providers.[4] Another example is T-Mobile, which introduced Network Pass for network trials and Network Scorecard for location-specific comparisons of its network

[2]For example, Ofcom, the UK's communications regulator, launched *Map Your Mobile* in 2025 to help consumers compare mobile-network coverage and performance across providers.

[3]See U.S. Department of Justice and Federal Trade Commission (2023, *Merger Guidelines*, Sections 2.2 and 4.4.B) and Federal Trade Commission (2013, *Horizontal Merger Investigation Data: Fiscal Years 1996-2011*, p. 3, footnote 17) for related discussions of market shares and competitive significance.

[4]See `https://www.gladia.io/speech-to-text-api-comparison`.

performance with that of competing carriers.[5]

To examine this question and its welfare implications, we allow Firm 1 to choose the information cost before prices are set. Greater transparency means lower information costs but is more costly to provide. We show that the firm always chooses weakly more transparency than is socially optimal. Firm 1 does not internalize the social gains from product matching; instead, it values transparency because consumer learning supports higher prices. When Firm 1 is almost always the better match, learning adds little value, and the firm chooses full opacity. When its match advantage is instead sufficiently large, learning softens price competition, and the firm chooses full transparency even though full opacity maximizes net total surplus. From the consumer's perspective, Firm 1 provides too little transparency in the former case and too much in the latter. Thus, even when the consumer prefers greater transparency, mandating it when Firm 1 chooses full opacity would reduce social welfare.

Related Literature. A closely related literature studies costly consumer learning before prices are observed. Ravid, Roesler, and Szentes (2022) study a monopoly model with private learning and show that informational hold-up discourages learning: as information costs vanish, equilibria converge to the Pareto-worst equilibrium of the zero-cost game. By contrast, when the monopolist sets the price before the consumer learns, this hold-up is avoided, and the allocation can approach the first best in the limit.[6] In symmetric duopoly models, Albrecht and Whitmeyer (2026) show that competition can mitigate informational hold-up and restore ex post efficiency as information costs vanish. Biglaiser, Gu, and Li (2025) show that higher information costs reduce perceived product differentiation and equilibrium prices, while their effect on consumer surplus can be nonmonotone. In our model, observed prices directly affect both information acquisition and product choice. A sufficient price advantage can induce purchase without learning, creating an additional channel of price competition. We allow for firm asymmetry and examine how information costs shape market allocation through equilibrium pricing, as well as the resulting welfare implications.

While we study information endogenously acquired by the consumer, related studies examine how information about consumer preferences held by consumers or firms shapes pricing and welfare. Roesler and Szentes (2017) study a monopoly pricing model in which a buyer observes a signal about her valuation. They characterize buyer-optimal information structures and show how information can limit surplus extraction. Moscarini and Ottaviani (2001) study price competition for a buyer with private information about her relative valuation of two products. They characterize possible mixed-strategy pricing regimes and

[5] See T-Mobile, "T-Mobile Takes Hassle and Worry Out of Switching," August 31, 2022.

[6] We compare the two monopoly timings in Appendix B using a simple model. Related monopoly models with learning after the seller's offer include Branco, Sun, and Villas-Boas (2012), who show that lower search costs can raise prices and reduce consumer surplus, and Mensch, Ravid, and Thereze (2026), who study monopolistic screening with flexible consumer learning and show that quality can be distorted downward even at the top.

their dependence on prior beliefs and signal precision. Armstrong and Zhou (2022) derive firm- and consumer-optimal signal structures, showing that the former increase product differentiation and soften competition, while the latter intensify competition but can distort product choice. Elliott et al. (2026) study how an intermediary provides information to firms. They find that consumer-optimal market segmentation achieves an efficient allocation, while producer-optimal segmentation can allow firms to extract all available surplus.

A large literature explains price dispersion through mixed-strategy pricing. Varian (1980) and Rosenthal (1980) show how firms randomize over prices when they serve captive consumers while competing for consumers who compare prices. Higher prices yield larger margins on captive demand, whereas lower prices attract price-comparing consumers. Subsequent work studies asymmetric captive customer bases and more general patterns of consumer consideration (Narasimhan (1988); Armstrong and Vickers (2019, 2022)), as well as asymmetries in production costs (Myatt and Ronayne (2026)).[7] Another strand of this literature studies mixed pricing when different consumer groups switch brands at different price gaps, which generates two discontinuities in demand (Shilony (1977); Raju, Srinivasan, and Lal (1990); Jing and Wen (2008)). Our model generates a more general two-cutoff demand structure, with the thresholds determined by the consumer's optimal decision. These papers also study equilibria in which the reservation price binds. With a nonbinding price cap, we characterize equilibrium pricing and outcomes across the full range of feasible parameters.

Our paper is also related to the literature on consumer search. In traditional sequential-search models, consumers incur costs to visit firms, and higher search costs typically soften price competition by giving firms more captive demand (Wolinsky (1986); Stahl (1989); Anderson and Renault (1999)). This effect can be reversed when prices are advertised before search. In price-directed search, lower prices can move a product earlier in the search process, and higher search costs strengthen firms' incentives to compete for being searched early and lead to lower prices (Armstrong and Zhou (2011); Choi, Dai, and Kim (2018); Haan, Moraga-González, and Petrikaitė (2018); Obradovits and Plaickner (2023)). In our model, higher information costs reduce effective product differentiation and intensify price competition. Related mechanisms arise from flexible investigation during sequential search in Jain and Whitmeyer (2021) and blind buying in Chen et al. (2021).

Our analysis of endogenous transparency is related to work on how firms strategically influence the product information available to consumers. Bar-Isaac, Caruana, and Cuñat (2010) show that a monopolist may use partial transparency as a means of non-price discrimination, while Gu and Xie (2013) find that competition strengthens high-quality firms' incentives to facilitate fit revelation but weakens those of low-quality firms. In search markets, Board and Lu (2018) show that full information provision can arise as search costs vanish when consumers' beliefs are private, while Au and Whitmeyer (2023) find that sufficiently

[7] Baye, Kovenock, and de Vries (1992) and Johnen and Ronayne (2021) study equilibrium structure and uniqueness in related models of price competition.

intense competition leads firms to provide the first-best level of information. Although the mechanisms differ from ours, related studies show that easier access to information need not improve market outcomes in competitive markets. Anderson and Renault (2009) show that comparative advertising can raise consumer surplus but reduce total surplus when quality asymmetry is sufficiently large. Hwang, Kim, and Boleslavsky (2026) find that requiring more information disclosure can reduce welfare, while Jain and Whitmeyer (2026) show that lower attention costs can worsen product choice by changing firms' disclosure incentives.

The remainder of the paper is organized as follows. Section 2 introduces the model. Section 3 characterizes the equilibrium, and Section 4 analyzes welfare. Section 5 and Section 6 study the quasi-monopoly regime and endogenous transparency, respectively. Section 7 extends the analysis to general distributions and flexible learning. Section 8 concludes. The appendices contain proofs and additional results.

2 Model

There are two firms, indexed by $i \in \{1,2\}$, producing differentiated goods at zero cost. Each firm chooses a single price $p_i \in [0,\bar{p}]$. A consumer is interested in purchasing one of the two products. Her valuation for Firm i's product is denoted by v_i. She obtains utility $v_i - p_i$ when purchasing from Firm i, while the outside option gives zero utility. Let $w := v_1 - v_2$ denote the consumer's relative match value. We write

$$(v_1, v_2) = \left(V + \frac{w}{2}, V - \frac{w}{2}\right),$$

where $V > 0$ is a common baseline value. Writing the valuations symmetrically around V is simply a normalization. We assume $V > \bar{p}$ so that the market is covered.[8]

We assume that $w = s > 0$ with probability $\mu \in (0,1)$ and $w = -1$ with probability $1-\mu$. Thus, Firm 1 is the better match in the first state, while Firm 2 is the better match in the second, with Firm 2's match advantage normalized to one. The parameter μ determines how often Firm 1 is the better match, while s determines the size of its advantage in that state. The prior mean of the relative match value is $\bar{w} := \mu s - (1-\mu)$. Throughout the paper, we assume that Firm 1 has a weak ex ante advantage, that is, $\bar{w} \geq 0$. The market is symmetric when $(\mu, s) = (1/2, 1)$.[9] A higher $\bar{w}$ therefore corresponds to greater market asymmetry as either μ or s increases, holding the other fixed.[10]

[8] The two valuations are perfectly negatively correlated, but this correlation is not essential under market coverage. Learning and product choice depend only on their difference w. For example, adding an independent common shock z to both valuations would change their correlation without changing w, and hence leave demand and equilibrium unchanged.

[9] When $\bar{w} = 0$ but $s \neq 1$, neither firm is favored by the prior, even though the frequencies and magnitudes of their respective match advantages differ.

[10] This specification can also be interpreted as a Hotelling model. Place Firm 1 at 0 and Firm 2 at 1 on the

The consumer does not initially observe w. After observing prices, she decides whether to acquire information about w. In the baseline model, information acquisition reveals w before the purchase decision. We also extend to a general distribution of match values with flexible information acquisition in Section 7 and show that the main qualitative results continue to hold.

The cost of learning is

$$C(k) := k \operatorname{Var}(w),$$

where $k \geq 0$ is the unit cost of information.[11] This specification separates the cost of acquiring information from the amount of uncertainty to be resolved, measured by $\operatorname{Var}(w)$. That is, for a given k, acquiring information is more costly when the prior over w is more dispersed.[12]

Timing.

1. The firms simultaneously choose prices.
2. After observing both prices, the consumer chooses whether to acquire information.
3. The match value w is realized. If the consumer acquires information, she observes w; otherwise, she remains uninformed.
4. The consumer makes her purchase decision.

We use subgame perfect equilibrium (SPE) as the solution concept. After every price pair, the consumer chooses an optimal information acquisition and purchase strategy, and the firms' price strategies are mutual best responses given this continuation behavior.

Modeling choice. Our formulation covers several settings. First, suppose Firm i has constant marginal cost c_i, and let $\xi := c_1 - c_2$. Writing prices in terms of markups, the consumer's payoff difference between buying from Firm 1 and Firm 2 is $(w - \xi) - [(p_1 - c_1) - (p_2 - c_2)]$. Thus, $w - \xi$ plays the role of the relative match value, taking values $s - \xi$ and $-1 - \xi$. Dividing the adjusted relative match value, markups, and learning cost by $1 + \xi$ does

unit interval, and let the consumer's location be $x \in \{0, 1\}$, with $x = 0$ with probability μ. If gross valuations are $B + \delta - tx$ and $B - t(1 - x)$, respectively, where B is a common baseline value, δ captures a valuation asymmetry between the firms, and $t > 0$ is the transportation cost, then $w = \delta + t(1 - 2x)$. Normalizing Firm 2's match advantage at $x = 1$, $t - \delta$, to one yields $w \in \{-1, s\}$, where $s = (t + \delta)/(t - \delta)$.

[11] The parameter k may reflect features of the information environment, such as product-information transparency, comparison tools, and recommender systems on online platforms.

[12] For example, when μ is close to one, Firm 1's product is already very likely to be the better match, and the consumer may need little effort to assess the products. When s is large, the consumer remains uncertain about which product is the better match, but one product may offer a substantial advantage if its specialized features suit her needs, while the alternative is only slightly better otherwise. Assessing these specialized features may require more extensive testing or comparisons, as when comparing specialized mattresses or running shoes with standard alternatives.

not change the consumer's decision and maps the (s, k, c_1, c_2) model into the baseline model $(s', k', 0, 0)$, where $s' = (s-\xi)/(1+\xi)$ and $k' = k(1+\xi)$. The equilibrium price distribution in the baseline model coincides with the distribution of the normalized markup $(p_i - c_i)/(1+\xi)$.

Also, we could introduce a differentiation parameter $\sigma > 0$ by letting $w \in \{-\sigma, \sigma s\}$, while keeping the learning cost independent of σ.[13] Dividing match values, prices, and learning costs by σ maps this setting into our model with information cost parameter k/σ. Equilibrium prices and profits are scaled by σ, as are consumer and total surplus relative to V. Thus, the comparative statics analysis with respect to k are preserved. While σ directly changes match differences, k affects product differentiation through consumer learning. We focus on the latter, since introducing σ does not change the equilibrium structure.

3 Equilibrium Characterization

In this section, we characterize the pricing equilibrium. We first analyze the consumer's optimal learning and purchase decisions given prices and derive the demand functions in Section 3.1. In Section 3.2, we show that no pure-strategy equilibrium exists when k is not too large and characterize the unique mixed-strategy equilibrium. Finally, in Section 3.3, we study how equilibrium and market allocation vary with k.

3.1 Consumer's Problem

We first analyze the consumer's information and purchase decisions. Let $\Delta := p_1 - p_2$. For a realized w and prices (p_1, p_2), her indirect utility is

$$u(w) := \max\left\{V + \frac{w}{2} - p_1,\, V - \frac{w}{2} - p_2\right\}.$$

Hence, with learning, she chooses Firm 1 when $w \geq \Delta$, while without learning, she chooses it when $\bar{w} \geq \Delta$.[14] Define the benefit of learning as

$$\begin{aligned} B(\Delta) &:= \mu u(s) + (1-\mu)u(-1) - u(\bar{w}) \\ &= \begin{cases} (1-\mu)(\Delta+1), & -1 < \Delta < \bar{w}, \\ \mu(s-\Delta), & \bar{w} \leq \Delta < s, \\ 0, & \Delta \leq -1 \text{ or } \Delta \geq s. \end{cases} \end{aligned}$$

[13] For example, two coffee products may become more differentiated in their flavors, with each becoming more appealing when it better matches the consumer's taste. Yet determining which coffee is the better match may still require the same effort, such as tasting both products.

[14] We assume that ties in the purchase decision are broken in favor of Firm 1. This does not affect the equilibrium analysis.

The benefit of learning is maximized at $\Delta = \bar{w}$, where the consumer is indifferent between the two firms based on the prior, and falls as the price difference moves away from this point. For $\Delta < \bar{w}$, the consumer would buy from Firm 1 without learning, so learning is useful only if it reveals $w = -1$ and switches her to Firm 2. Similarly, for $\Delta > \bar{w}$, learning is useful only if it reveals $w = s$.

Suppose that $0 < k < \bar{k} := 1/(1+s)$. Then $B(\bar{w}) > C(k)$, and $B(\Delta) = C(k)$ yields two cutoffs,

$$\alpha := -1 + k\mu(1+s)^2 \quad \text{and} \quad \beta := s - k(1-\mu)(1+s)^2.$$

Since $-1 < \alpha < \bar{w} < \beta < s$, learning is optimal if and only if $\Delta \in [\alpha, \beta]$. This is illustrated in Figure 1.

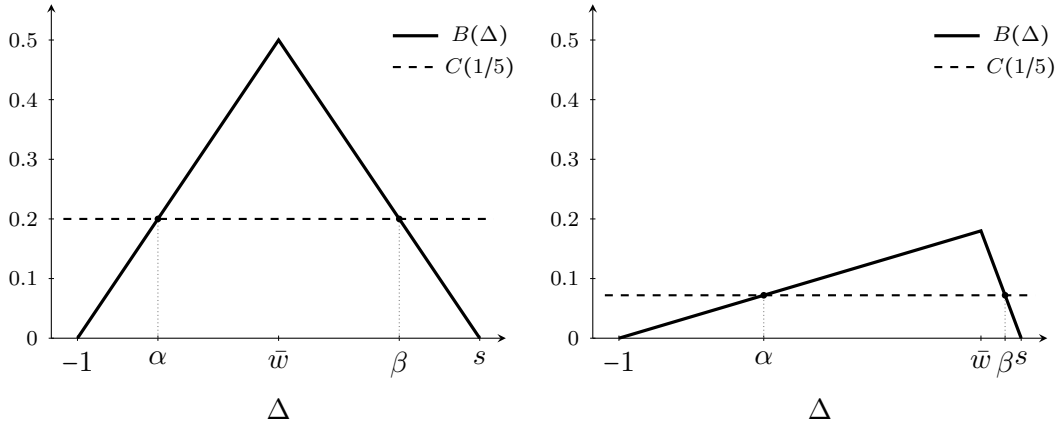


Figure 1: Benefit and cost of learning for $s = 1$ and $k = 0.2$, with $\mu = 0.5$ (left) and $\mu = 0.9$ (right). The intersections at α and β determine the learning region.

It is useful to distinguish two sources of no learning. For $-1 < \Delta < \alpha$ or $\beta < \Delta < s$, information could affect product choice, but the consumer does not learn because its benefit is smaller than its cost. By contrast, for $\Delta < -1$ or $\Delta > s$, the price difference alone determines the preferred product in both states, and information has no value for product choice. At $k = 0$, we have $(\alpha, \beta) = (-1, s)$, and the first source of no learning disappears while the second remains. In the latter region, the consumer is indifferent between learning and not learning because information is free and does not affect her purchase decision. Thus, her learning choice is not uniquely determined. Henceforth, when $k = 0$, we can let the consumer always learn.[15]

We next derive demand. Let $D_i(\Delta)$ denote Firm i's demand, defined as the probability that the consumer purchases from Firm i under her optimal information and purchase decisions, given the price difference Δ. If $\Delta < \alpha$, the consumer does not learn and, since $\Delta < \bar{w}$, purchases from Firm 1. If $\alpha < \Delta < \beta$, she learns and purchases from Firm 1 when $w = s$ and from Firm 2 when $w = -1$. If $\Delta > \beta$, she does not learn and, since $\Delta > \bar{w}$, purchases from

[15] This is without loss for the equilibrium analysis.

Firm 2. Thus,

$$D_1(\Delta) = \begin{cases} 1, & \Delta < \alpha, \\ d_\alpha & \Delta = \alpha, \\ \mu, & \alpha < \Delta < \beta, \\ d_\beta & \Delta = \beta, \\ 0, & \Delta > \beta, \end{cases} \qquad D_2(\Delta) = 1 - D_1(\Delta).$$

At $\Delta = \alpha$, the consumer is indifferent between learning and being uninformed. No learning yields Firm 1 demand one, whereas learning yields demand μ, so any admissible tie-breaking rule satisfies $d_\alpha \in [\mu, 1]$. Similarly, $d_\beta \in [0, \mu]$.[16] As we will see below, the equilibrium price distributions and outcomes are independent of the particular admissible tie-breaking rule.

Define the length of the learning region as $L := \beta - \alpha$. As the information cost k increases, this region shrinks because α increases and β decreases, as illustrated in Figure 2. The two products therefore become closer substitutes, as the consumer relies more on relative prices and less on product-fit information when choosing between them. This affects firms' pricing incentives and hence price competition, as we show below.

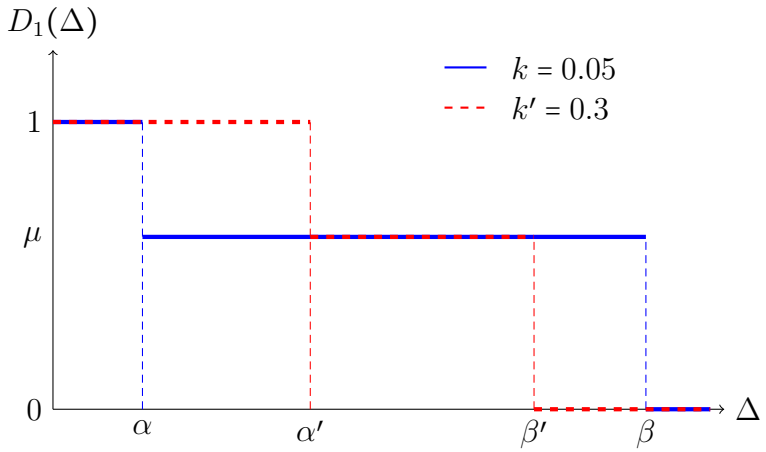


Figure 2: Firm 1's demand as a function of Δ with $(\mu, s) = (0.6, 1)$.

When $k \geq \bar{k}$, no learning is optimal for every price pair and equilibrium outcomes no longer depend on k. We leave the equilibrium analysis of this region to Appendix C and focus on $k < \bar{k}$ in the main analysis.

[16] When $k \in (0, \bar{k})$, purchase tie-breaking does not arise under the consumer's optimal behavior, since the preceding demand characterization implies that she strictly prefers one firm within each optimal information regime. When $k = 0$, the consumer always learns. At $\Delta = -1$, she buys from Firm 1 in both states, since the purchase tie is broken in favor of Firm 1, giving $d_\alpha = 1$. At $\Delta = s$, she buys from Firm 1 when $w = s$, giving $d_\beta = \mu$.

3.2 Mixed-Strategy Pricing Equilibrium

We first consider the existence of pure-strategy pricing equilibria. Define

$$\hat{k} := \frac{1+\mu s}{\mu(2-\mu)(1+s)^2},$$

which satisfies $\hat{k} \le \bar{k}$, with equality if and only if $\bar{w} = 0$. Throughout the baseline analysis, we assume that $\bar{p} > 3 + 4s$.

Lemma 1. *Suppose that $0 \le k < \hat{k}$. Under any admissible tie-breaking rule, no pure-strategy pricing equilibrium exists.*

Proof. Away from the two cutoffs α and β, demand is locally flat, so a firm with positive demand can profitably raise its price slightly unless it is at the price cap. If $p_2 = \bar{p} > p_1$, Firm 1 can raise its price since $\Delta < 0 < \beta$. If $p_1 = \bar{p} > p_2$, Firm 2 can raise its price if it has positive demand, or lower it to obtain positive demand otherwise. If both firms charge $\bar{p}$, at least one has demand at most $1/2$. Since $\alpha \ge -1$ and $\beta \le s$, that firm can capture the entire market by lowering its price by less than $1 + s$. Its resulting profit exceeds $\bar{p}/2$, since $\bar{p} > 2(1+s)$. Hence any pure-strategy equilibrium must satisfy $\Delta \in \{\alpha, \beta\}$.

Suppose first that $\Delta = \beta$. If $d_\beta < \mu$, Firm 1 can slightly lower its price and increase its demand from d_β to μ. If $d_\beta = \mu$, Firm 2 can instead slightly raise its price while keeping demand $1 - \mu$. Thus $\Delta = \beta$ cannot be sustained. Now suppose that $\Delta = \alpha$. If $d_\alpha < 1$, Firm 1 can profitably lower its price when $p_1 > 0$, or raise it when $p_1 = 0$. If $d_\alpha = 1$ and $p_2 > 0$, Firm 2 can slightly lower its price and obtain positive demand. Thus any remaining candidate must have $p_2 = 0$, and hence $p_1 = \alpha$. If $\alpha < 0$, such a price is infeasible. If $\alpha \ge 0$, Firm 1's profit is α. Since

$$\alpha - \mu\beta = \mu(2-\mu)(1+s)^2\left(k - \hat{k}\right) < 0,$$

Firm 1 can instead charge $p_1 = \beta - \varepsilon$ and earn $\mu(\beta - \varepsilon) > \alpha$ for sufficiently small $\varepsilon > 0$. □

In other words, when k is not too large, consumer learning provides enough product differentiation that Firm 1 prefers sharing the market with Firm 2 at a higher price to capturing the entire market at a lower price. This incentive allows both firms to remain active in equilibrium, as we will see.

Since no pure-strategy equilibrium exists for $k < \hat{k}$, we now characterize the mixed-strategy equilibrium. Let F_i denote Firm i's price distribution. To see how the equilibrium price distributions are determined, suppose for the moment that they are atomless and have connected supports $\mathrm{supp}(F_1) = [a, a+L]$ and $\mathrm{supp}(F_2) = [b, b+L]$, where $a, b > 0$, $a < b + \beta < a + L$, and $b < a - \alpha < b + L$. The proposition below shows that every equilibrium must have this form.

In a mixed-strategy equilibrium, every price p in Firm i's support must yield the same expected profit Π_i, and no price outside its support can yield a higher profit. Given Firm 2's

price distribution F_2, Firm 1's expected profit from charging p is

$$\begin{aligned}\pi_1(p) &= p\,\mathbb{E}_{p_2\sim F_2}[D_1(p-p_2)]\\ &= p\Big[1-(1-\mu)F_2(p-\alpha)-\mu F_2(p-\beta)\Big].\end{aligned}$$

Similarly, given F_1, Firm 2's expected profit is

$$\pi_2(p) = p\Big[1-\mu F_1(p+\beta)-(1-\mu)F_1(p+\alpha)\Big].$$

The prices $b+\beta$ for Firm 1 and $a-\alpha$ for Firm 2 divide their respective supports into two branches. We refer to these prices as the switching prices. Define $\mathcal{L}_1 := [a, b+\beta]$ and $\mathcal{U}_1 := [b+\beta, a+L]$, and similarly $\mathcal{L}_2 := [b, a-\alpha]$ and $\mathcal{U}_2 := [a-\alpha, b+L]$.

Accordingly, Firm 1's indifference condition is $\Pi_1 = p[1-(1-\mu)F_2(p-\alpha)]$ for $p \in \mathcal{L}_1$, while $\Pi_1 = \mu p[1-F_2(p-\beta)]$ for $p \in \mathcal{U}_1$. Since $F_2(b) = 0$, the latter condition at $p = b+\beta$ gives $\Pi_1 = \mu(b+\beta)$. Solving these two conditions for F_2 yields

$$F_2(p) = \begin{cases} 0, & p < b,\\ 1-\dfrac{b+\beta}{p+\beta}, & b \le p < a-\alpha,\\ \dfrac{1}{1-\mu}\left(1-\dfrac{\mu(b+\beta)}{p+\alpha}\right), & a-\alpha \le p < b+L,\\ 1, & p \ge b+L. \end{cases} \tag{1}$$

The same argument for Firm 2 gives $\Pi_2 = (1-\mu)(a-\alpha)$ and

$$F_1(p) = \begin{cases} 0, & p < a,\\ 1-\dfrac{a-\alpha}{p-\alpha}, & a \le p < b+\beta,\\ \dfrac{1}{\mu}\left(1-\dfrac{(1-\mu)(a-\alpha)}{p-\beta}\right), & b+\beta \le p < a+L,\\ 1, & p \ge a+L. \end{cases} \tag{2}$$

Continuity of F_1 at $b+\beta$ and of F_2 at $a-\alpha$ gives

$$1-\frac{b+\beta}{a+L} = \frac{1-\mu(b+\beta)/a}{1-\mu}, \tag{3}$$

$$1-\frac{a-\alpha}{b+L} = \frac{1-(1-\mu)(a-\alpha)/b}{\mu}. \tag{4}$$

These two equations determine the lower support endpoints (a, b).

The proposition below shows that this candidate is the unique equilibrium in price dis-

tributions for $0 \le k < \hat{k}$, with both firms making positive profits.

Proposition 1. *Suppose that $0 \le k < \hat{k}$. Under any admissible tie-breaking rule, there is a unique equilibrium in price distributions satisfying*

(i) The equilibrium distributions are atomless with supports $\text{supp}(F_1) = [a, a + L]$ *and* $\text{supp}(F_2) = [b, b + L]$, *where* $a < b + \beta < a + L$ *and* $b < a - \alpha < b + L$.

(ii) The equilibrium CDFs are given by (1)-(2), *where* (a, b) *is the unique positive solution to* (3)-(4).

(iii) Equilibrium profits are $\Pi_1 = \mu(b + \beta)$ *and* $\Pi_2 = (1 - \mu)(a - \alpha)$.

Figure 3 illustrates the equilibrium price distributions in an asymmetric and a symmetric case.

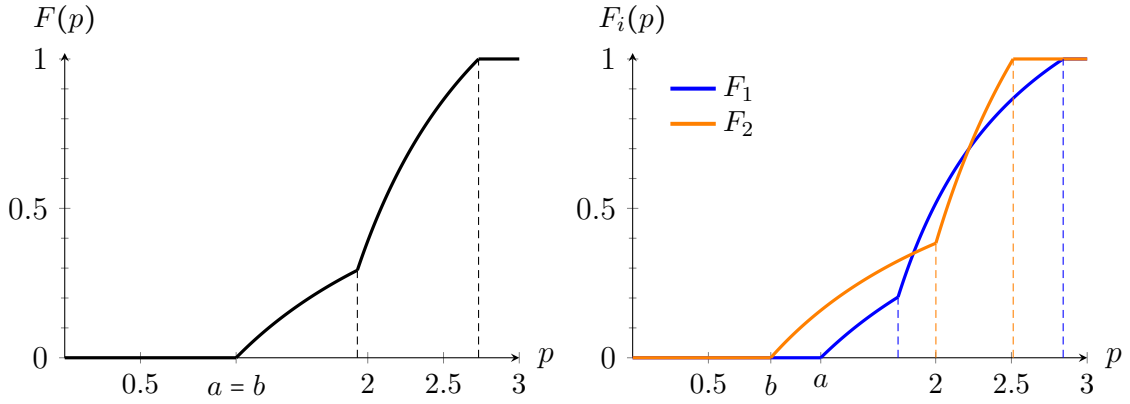


Figure 3: Equilibrium price distributions. The left panel shows the symmetric case with $(\mu, s, k) = (0.5, 1, 0.1)$, and the right panel shows the asymmetric case with $(\mu, s, k) = (0.6, 1, 0.1)$.

The proof proceeds in several steps. The main challenge is to establish the structure of the price supports under asymmetry. We first rule out atoms and show that both supports have length $L = \beta - \alpha$. We then exclude gaps, which would allow a firm to raise its price without losing demand. The indifference conditions give the CDFs above, and a separate argument shows that the continuity conditions uniquely determine the support endpoints. Finally, we verify that neither firm can profitably deviate outside its support.

The switching prices are the points at which the trade-off between margin and demand changes. At $p_1 = b + \beta$, the price difference Δ lies in $[\alpha, \beta]$ for every p_2 in Firm 2's support $[b, b + L]$. Thus, for almost every p_2, the consumer learns and Firm 1 has demand μ, yielding profit $\mu(b + \beta)$. When Firm 1 charges a price in its lower branch $\mathcal{L}_1$, we have $\Delta \le \beta$, and Firm 1 never has zero demand. It gives up some margin but can increase its demand from μ to one when Firm 2 charges a high price. For prices in the upper branch $\mathcal{U}_1$, by contrast, we have $\Delta \ge \alpha$, and Firm 1 can no longer serve the whole market. It earns a higher margin

but faces the risk that its demand falls from μ to zero when Firm 2 charges a low price. The same logic applies to Firm 2. This change in the trade-off generates the two pieces of each equilibrium CDF.[17]

Price branches	Possible outcome
$\mathcal{L}_1 \times \mathcal{L}_2 \implies \alpha \le \Delta \le \beta$	learning
$\mathcal{U}_1 \times \mathcal{U}_2 \implies \alpha \le \Delta \le \beta$	learning
$\mathcal{L}_1 \times \mathcal{U}_2 \implies \Delta \le \beta$	learning, or no learning and Firm 1 is chosen
$\mathcal{U}_1 \times \mathcal{L}_2 \implies \Delta \ge \alpha$	learning, or no learning and Firm 2 is chosen

Table 1: Possible outcomes by price-branch pair.

The threshold $\hat{k}$ is the information cost below which the equilibrium features learning with positive probability and both firms earn positive profits.[18] Once k reaches $\hat{k}$, the consumer no longer learns in equilibrium and the market enters the quasi-monopoly regime, which we will discuss in Section 5. We maintain $k < \hat{k}$ and focus on the mixed-strategy equilibrium characterized above until then.

Lemma 2. *Let $F_i(\cdot; k)$ denote Firm i's equilibrium price distribution. If $0 \le k_1 < k_2$, then*

$$F_i(p; k_2) \ge F_i(p; k_1) \qquad \text{for every } p \in \mathbb{R} \text{ and } i \in \{1, 2\}.$$

Both firms' equilibrium price distributions shift downward in the sense of first-order stochastic dominance as k increases. This implies that their equilibrium profits decrease, since the switching prices $b + \beta$ and $a - \alpha$ both fall. Moreover, their price supports become narrower, as each has length $L = \beta - \alpha$, which decreases with k.

The intuition is as follows. As k increases, the learning region shrinks and the consumer relies more on relative prices than on personal fit. Thus, the products become less differentiated, intensifying price competition and shifting both price distributions downward. Furthermore, a small change in price is more likely to make a firm lose or capture the entire market. Further price cuts then bring no additional benefit, while high prices become riskier. This makes extreme prices less attractive and narrows the range of prices used in equilibrium.

[17] The two equilibrium price distributions need not be ordered by first- or second-order stochastic dominance. For example, when $(\mu, s, k) = (0.4, 2, 0.1)$, we have $a \simeq 1.253 < b \simeq 1.390$, but $\mathbb{E}[p_1] \simeq 2.533 > \mathbb{E}[p_2] \simeq 2.350$. Thus neither distribution second-order stochastically dominates the other.

[18] $\hat{k}$ is U-shaped in μ. As μ rises, Firm 1 becomes more likely to be the better match, so the consumer can rely more on the prior and $\hat{k}$ initially falls. When μ is close to one, however, little uncertainty remains, making learning cheap, and $\hat{k}$ rises. By contrast, $\hat{k}$ decreases with s, as the increase in the cost of learning dominates the increase in its benefit.

3.3 Information Costs and Market Allocation

We now examine how information costs affect equilibrium pricing and the consumer's learning and purchase behavior.

When neither firm has an ex ante advantage, $\bar{w} = 0$, we show that changing k simply rescales the equilibrium. Let $\Delta(k)$ denote the equilibrium price gap at information cost k.

Lemma 3. *Suppose that $\bar{w} = 0$ and define $\lambda(k) := 1 - k(1+s)$. Then,*

$$F_i(p;k) = F_i\left(\frac{p}{\lambda(k)};0\right), \qquad \Delta(k) \overset{d}{=} \lambda(k)\Delta(0), \qquad \Pi_i(k) = \lambda(k)\Pi_i(0).$$

The lemma implies that the probability that the price gap lies in the learning region in equilibrium does not depend on k. Define $P_0 := \Pr(-1 \le \Delta(0) \le s)$. When $\bar{w} = 0$, we have $\alpha = -\lambda$, $\beta = s\lambda$, and $\Delta(k) \overset{d}{=} \lambda\Delta(0)$. Hence,

$$\Pr(\alpha \le \Delta(k) \le \beta) = P_0$$

for every $k > 0$. Similarly, each of the two tail probabilities is also independent of k: $\Pr(\Delta(k) < \alpha) = \Pr(\Delta(0) < -1)$ and $\Pr(\Delta(k) > \beta) = \Pr(\Delta(0) > s)$. That is, as k increases, the learning region becomes narrower, but the distribution of price gap Δ contracts by exactly the same factor $\lambda(k)$. Intuitively, when $\bar{w} = 0$, a higher information cost does not favor either firm; it only reduces the differentiation created by learning, leaving the probability of learning unchanged. The consumer makes the same learning and purchase decisions in equilibrium, and therefore the equilibrium allocation remains unchanged as the information cost increases.

Symmetric equilibrium. The market is symmetric when $(\mu, s) = (1/2, 1)$. The corresponding equilibrium follows directly from Proposition 1. This pricing equilibrium also arises in Biglaiser, Gu, and Li (2025) and Albrecht and Whitmeyer (2026).

Corollary 1. *Suppose that $(\mu, s) = (1/2, 1)$. When $0 \le k < 1/2$, the unique equilibrium in price distributions is symmetric. Both firms use the common support $[\sqrt{2}\beta, (\sqrt{2}+2)\beta]$ and the common CDF*

$$F(p) = \begin{cases} 0, & p < \sqrt{2}\beta, \\ 1 - \dfrac{(\sqrt{2}+1)\beta}{p+\beta}, & \sqrt{2}\beta \le p < (\sqrt{2}+1)\beta, \\ 2 - \dfrac{(\sqrt{2}+1)\beta}{p-\beta}, & (\sqrt{2}+1)\beta \le p \le (\sqrt{2}+2)\beta, \\ 1, & p > (\sqrt{2}+2)\beta. \end{cases}$$

Each firm earns $\Pi_1 = \Pi_2 = (\sqrt{2}+1)\beta/2$.

Proof. Under full symmetry, $\alpha = -\beta$ and $L = 2\beta$. If (a, b) solves the endpoint system, then (b, a) does as well. Since the solution is unique, it follows that $a = b$. The equations (3)-(4) reduce to the same equation, $a^2 = 2\beta^2 \Leftrightarrow a = \sqrt{2}\beta$. Substituting into the equilibrium CDFs gives the stated expressions. □

In this case, we can compute the learning probability explicitly. By symmetry, the two no-learning tails have the same probability, and the equilibrium CDF gives

$$\frac{1-P_0}{2} = \Pr(\Delta(0) > 1) = \int_{\sqrt{2}+1}^{\sqrt{2}+2} F(p-\beta)\, dF(p) = (\sqrt{2}+1)^2 \log\left(\frac{\sqrt{2}+1}{2}\right) - 1.$$

Thus $P_0 \simeq 0.8059$. For every $k \in (0, 1/2)$, the consumer therefore learns with probability approximately 80.6 percent in equilibrium, while each no-learning tail occurs with probability approximately 9.7 percent, as shown in Figure 4. The figure also illustrates how the distribution of Δ becomes less dispersed in the sense of a mean-preserving contraction as k increases.[19]

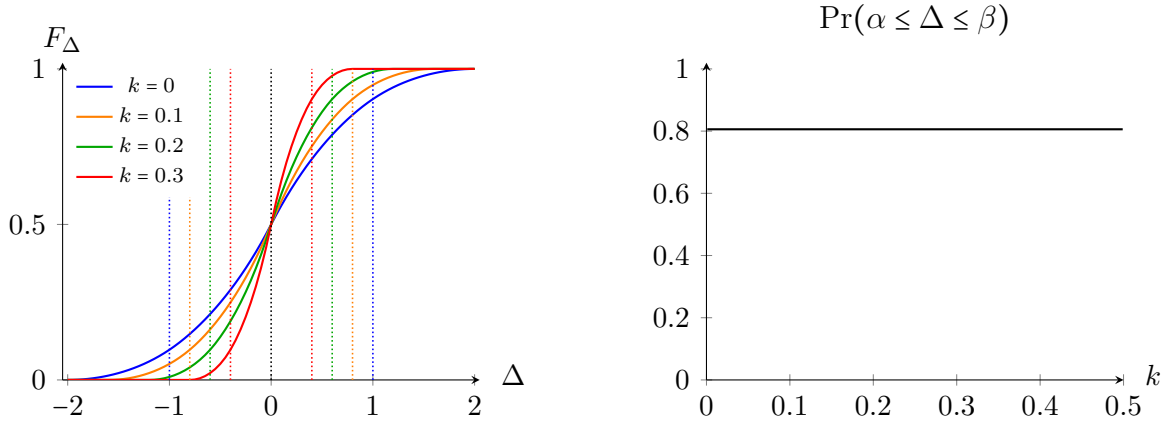


Figure 4: Distribution of Δ and the learning probability in the symmetric case $(\mu, s) = (0.5, 1)$. The dotted lines indicate the learning cutoffs α and β.

We now turn to the case where Firm 1 has an ex ante advantage ($\bar{w} > 0$). Here, a higher information cost not only reduces product differentiation but also changes the equilibrium allocation through its asymmetric effects on consumer learning and purchase decisions.

Unlike when $\bar{w} = 0$, the distribution of Δ no longer rescales with the learning region as k increases, even though its support still shrinks under stronger competition.[20] Consequently,

[19] More generally, when $\bar{w} = 0$, the same ordering applies to the centered price gap $\Delta(k) - \mathbb{E}[\Delta(k)]$. For $k_1 < k_2$, let $X := \Delta(k_1) - \mathbb{E}[\Delta(k_1)]$ and $c := \lambda(k_2)/\lambda(k_1) \in (0, 1)$. Then $\Delta(k_2) - \mathbb{E}[\Delta(k_2)] \overset{d}{=} cX$. For every convex φ, $\varphi(cX) \le c\varphi(X) + (1-c)\varphi(0)$ holds. Since $\mathbb{E}[X] = 0$, Jensen's inequality gives $\varphi(0) \le \mathbb{E}[\varphi(X)]$. Hence $\mathbb{E}[\varphi(cX)] \le \mathbb{E}[\varphi(X)]$. Since X and cX have the same mean, cX is a mean-preserving contraction of X.

[20] In Lemma A.2, we show that $\mathrm{supp}(\Delta(k_2)) \subset \mathrm{supp}(\Delta(k_1))$ for $k_2 > k_1$.

the equilibrium learning probability need not be monotone in k. To understand how these forces affect learning and, in turn, the allocation, we next consider how the firms strategically randomize over prices.

Lemma 4. *Suppose that $\bar{w} > 0$. Define $m_1(k) := \Pr(p_1 \in \mathcal{L}_1; k)$ and $m_2(k) := \Pr(p_2 \in \mathcal{L}_2; k)$. Then $m_1(k)$ increases strictly to 1 and $m_2(k)$ decreases strictly to 0 as $k \uparrow \hat{k}$.*

Because $\bar{w} > 0$, the two firms respond differently to a higher information cost. For Firm 1, the prior is already in its favor. As k increases, the cutoff α rises, so Firm 1 does not need to lower its price as much relative to Firm 2 to induce no learning and capture the entire market. Lower prices therefore become more attractive, and Firm 1 puts more probability on its lower branch. Firm 2 faces the opposite pressure. A lower price can also capture the entire market when the price gap exceeds β, and the fall in β makes this easier. But Firm 1 is also pricing aggressively, and Firm 2 must lower its price further to make such events occur. Lowering its price simply to avoid zero demand therefore becomes less attractive for Firm 2. It shifts probability away from its lower branch and toward its upper branch, even at the risk of losing demand altogether.

In particular, Lemma 4 suggests how the two no-learning outcomes change with k. From Table 1, $\Delta < \alpha$ can occur only when Firm 1 prices on its lower branch and Firm 2 on its upper branch, whereas $\Delta > \beta$ requires the opposite branch pair. Since the firms randomize independently,

$$\Pr(p_1 \in \mathcal{L}_1 \,\&\, p_2 \in \mathcal{U}_2) = m_1(1 - m_2), \qquad \Pr(p_1 \in \mathcal{U}_1 \,\&\, p_2 \in \mathcal{L}_2) = (1 - m_1)m_2.$$

As k rises, the first probability increases and the second decreases. These branch shifts therefore push in the direction of increasing $\Pr(\Delta < \alpha)$ and decreasing $\Pr(\Delta > \beta)$. The next lemma shows that the equilibrium tail probabilities indeed move in these directions.

Lemma 5. *Suppose that $\bar{w} > 0$. In equilibrium,*

(i) $\Pr(\Delta < \alpha)$ increases strictly to 1, whereas $\Pr(\Delta > \beta)$ decreases strictly to 0, as $k \uparrow \hat{k}$.

(ii) Firm 1's expected demand $\bar{D}_1(k)$ increases strictly to one, whereas Firm 2's expected demand $\bar{D}_2(k)$ decreases strictly to zero, as $k \uparrow \hat{k}$.

As k increases, Firm 1 puts more probability on its lower price branch, while Firm 2 puts more probability on its upper branch. Together, these changes make Firm 1's price lower relative to Firm 2's more frequently, inducing the consumer to stop learning and buy from Firm 1 more often. This also shifts equilibrium demand toward Firm 1. To see this, Firm 1's equilibrium demand can be written as

$$\begin{aligned}\bar{D}_1(k) &= \Pr(\Delta < \alpha) + \mu \Pr(\alpha \le \Delta \le \beta) \\ &= \mu + (1 - \mu) \Pr(\Delta < \alpha) - \mu \Pr(\Delta > \beta).\end{aligned}$$

Figure 5 shows how the distribution of price differences and the consumer's learning and purchase behavior change with k when $\bar{w} > 0$.

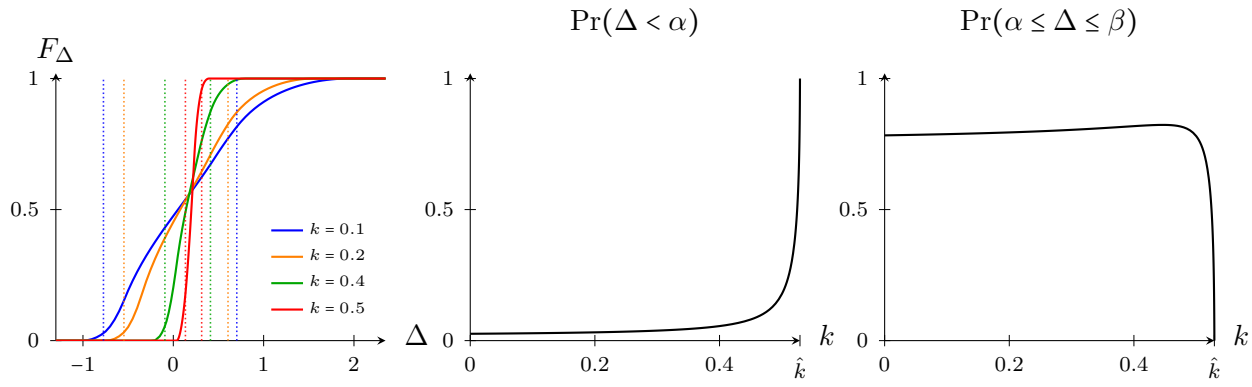


Figure 5: Distribution of Δ and outcome probabilities with $(\mu, s) = (0.7, 0.8)$. The dotted lines indicate the learning cutoffs α and β.

4 Welfare Analysis

This section addresses two distinct welfare questions. First, does the competitive equilibrium ever lead the consumer to choose the better-matched product in each state? Second, how do consumer surplus and total surplus vary with the information cost k? For example, a free information equilibrium can fail to attain the first best and yet maximize total surplus across k.

4.1 Failure of Ex Post Efficiency

We use the ex post efficient allocation as our first-best benchmark. It assigns the consumer to the better-matched product in each realized state: Firm 1 when $w = s$ and Firm 2 when $w = -1$, which yields

$$TS^{FB} := V + \frac{\mu s + (1 - \mu)}{2}.$$

In equilibrium, misallocation occurs when the price difference lies outside the learning region. If $\Delta < \alpha$, the consumer chooses Firm 1 even when $w = -1$, resulting in a welfare loss of one in that state. If $\Delta > \beta$, she chooses Firm 2 even when $w = s$, resulting in a loss of s. The corresponding expected losses are $(1 - \mu) \Pr(\Delta < \alpha)$ and $\mu s \Pr(\Delta > \beta)$. Subtracting these losses and the expected information cost from the first best gives

$$TS^{FB} - TS(k) = (1 - \mu) \Pr(\Delta < \alpha) + \mu s \Pr(\Delta > \beta) + C(k) \Pr(\alpha \le \Delta \le \beta).$$

The following proposition shows that both types of misallocation occur with positive probability in equilibrium.

Proposition 2. *In equilibrium,*

$$[\alpha, \beta] \subset \text{int}(\text{supp}(\Delta)).$$

Thus, the equilibrium fails to achieve ex post efficiency even when information is free and the consumer learns with probability one.

The strategic interaction between the firms generates price dispersion, with some price gaps falling outside the learning region. When the consumer learns, each firm retains positive demand when its product is the better match, allowing it to charge a relatively high price without losing that demand. At the same time, a sufficiently low price relative to the other firm can induce the consumer to buy from that firm without learning. This gives another incentive to lower prices.

The case of free information is of particular interest. At $k = 0$, the learning cutoffs become $(\alpha, \beta) = (-1, s)$. When $\Delta < \alpha$ or $\Delta > \beta$, the price difference is large enough that learning does not change the consumer's choice and the same firm is chosen in both states. Thus, even a fully informed consumer chooses the worse-matched product with positive probability in each state.[21]

Remark 1. By contrast, we show in Appendix B that this misallocation does not arise under monopoly when the consumer observes the price before deciding whether to learn her valuation for the product. For sufficiently low information costs, the monopolist finds it profitable to charge the highest price that induces learning. As a result, total surplus approaches the first-best level as information costs vanish. This conclusion need not hold when the consumer makes her learning decision without observing the price, as in Ravid, Roesler, and Szentes (2022). We also compare these two benchmarks.

4.2 Welfare and the Cost of Information

Having shown that competition generates misallocation, we now examine how welfare changes with the information cost. Let $X \in \{0, 1\}$ indicate the consumer's purchase, with $X = 1$ if she buys from Firm 1. Since the realized product value is $V + (2X - 1)w/2$, total surplus can

[21] Although the equilibrium fails to attain the first best for every $\mu < 1$, the welfare loss vanishes as $\mu \to 1$. Indeed, as $\mu \to 1$, we have $a \to s$, $b \to 0$, and $\beta \to s$. Firm 1's equilibrium price converges in probability to s, and $\Pr(\Delta > \beta) \to 0$. Hence, both sources of misallocation vanish: Firm 2 is the better match with probability $1 - \mu \to 0$, while the probability that the consumer purchases from Firm 2 when $w = s$ also converges to zero. The expected information cost vanishes as well. Thus, $\lim_{\mu \to 1} TS(k; \mu, s) = V + s/2 = TS^{FB}|_{\mu=1}$.

be written as

$$\begin{aligned} TS(k) &= \mathbb{E}\left[V + \frac{(2X-1)w}{2}\right] - C(k)\Pr(\alpha \le \Delta \le \beta) \\ &= V + \underbrace{\bar{w}\left(\bar{D}_1(k) - \frac{1}{2}\right)}_{\text{allocation effect}} + \underbrace{\mathrm{Cov}(X,w)}_{\text{matching effect}} - C(k)\Pr(\alpha \le \Delta \le \beta). \end{aligned} \tag{5}$$

The allocation effect captures the value of shifting demand toward Firm 1, which is the better product on average when $\bar{w} > 0$. The matching effect captures the additional value from choosing the better-matched product in each state. When the consumer chooses the better-matched product, her expected product value is $\mathbb{E}[\max\{v_1, v_2\}]$. If instead she randomly chose Firm 1 with probability μ and Firm 2 with probability $1-\mu$, independently of w, the expected value would be $\mu\mathbb{E}[v_1] + (1-\mu)\mathbb{E}[v_2]$. The difference is $\mu(1-\mu)(1+s) > 0$, which is the gain from better matching. In equilibrium,

$$\mathrm{Cov}(X,w) = \mu(1-\mu)(1+s)\Pr(\alpha \le \Delta \le \beta),$$

because the consumer's product choice can improve only when she learns. The last term in (5) is the expected learning cost.

We first consider the case without an ex ante advantage. In this case, the results from the previous section give a simple comparative-static result for total surplus.

Corollary 2. *Suppose that $\bar{w} = 0$. Then, total surplus is strictly decreasing in k.*

When neither firm has an ex ante advantage, the equilibrium allocation does not change with the information cost. Specifically, when $\bar{w} = 0$, the allocation effect in (5) vanishes. Since the probability of learning remains P_0 and $C(k) = ks$, we obtain

$$TS(k) = TS(0) - ksP_0,$$

which is strictly decreasing in k. Thus, a higher k only raises expected information costs and reduces total surplus.

The case with an ex ante advantage is different because the allocation itself changes with k: a higher k shifts no-learning purchases toward Firm 1. This shift can raise expected product value, but learning also improves matching when Firm 2 is the better match. Accounting for the cost of learning, the overall effect of k on total surplus is therefore ambiguous.

The welfare-maximizing information cost nevertheless has a simple characterization in terms of a unique cutoff in market asymmetry.

Proposition 3. *Fix $s > 0$. There exists a unique $\mu^*(s) \in (1/(1+s), 1)$ such that $TS(k)$ is maximized at $k = 0$ if $\mu < \mu^*(s)$ and at $k = \hat{k}$ if $\mu \ge \mu^*(s)$. Moreover, $\mu^*(s)$ is continuously*

differentiable, $\mu^{*\prime}(s) < 0$*, and* $\lim_{s\to\infty} \mu^*(s) = 0$.[22]

The proposition shows that total surplus is maximized either at free information or at the no-learning outcome. The decomposition in (5) helps explain why free information need not maximize total surplus. We first compare these two outcomes. At $k = \hat{k}$, the consumer always purchases from Firm 1 without learning, so $\bar{D}_1(\hat{k}) = 1$ and $\mathrm{Cov}(X, w) = 0$. Hence $TS(\hat{k}) = V + \bar{w}/2$. At $k = 0$, define $R_0 := \Pr(\Delta(0) > s)$. Since $\bar{D}_1(0) = 1 - (1-\mu)P_0 - R_0$, we obtain

$$\begin{aligned} TS(0) - TS(\hat{k}) &= \underbrace{\mu(1-\mu)(1+s)P_0}_{\text{matching gain}} - \underbrace{\bar{w}[1-\bar{D}_1(0)]}_{\text{allocation loss}} \\ &= (1-\mu)P_0 - \bar{w}R_0. \end{aligned} \tag{6}$$

In both cases, the consumer incurs no information cost: information is free at $k = 0$, while she does not learn at $k = \hat{k}$. Relative to the no-learning outcome, the first term $(1-\mu)P_0$ is the gain from learning: the consumer can choose Firm 2 when it is the better match, rather than always purchasing from Firm 1. The second term $\bar{w}R_0$ is the loss when the price difference exceeds s, so the consumer purchases from Firm 2 regardless of which product is better, even though Firm 1's product has a higher expected value.

The comparison in (6) shows why market asymmetry matters. As μ increases, Firm 2 is less likely to be the better match, reducing the potential gain from learning. Meanwhile, Firm 1's ex ante advantage increases, making purchases from Firm 2 regardless of the match more costly. This tends to make the no-learning outcome more socially desirable than free information. Although the equilibrium probabilities P_0 and R_0 also change with μ and s, Proposition 3 shows that no learning is optimal for sufficiently large μ. A larger s also increases Firm 1's ex ante advantage and can compensate for a smaller μ. Indeed, the cutoff $\mu^*(s)$ decreases with s. Moreover, $\lim_{s\to\infty} \mu^*(s) = 0$: for any fixed $\mu > 0$, no learning maximizes total surplus when s is sufficiently large.

An intermediate information cost never maximizes total surplus. Compared with $k = \hat{k}$, learning at sufficiently large $k < \hat{k}$ helps the consumer avoid mistakes when Firm 2 is the better match. However, the cost of learning exceeds the gain from learning, and total surplus is lower than at $\hat{k}$.[23] When k is relatively low, compared with the two extremes, it does not provide enough gains from matching or allocation, given the positive cost of learning.

An argument similar to the one used above shows that even when $\bar{w} > 0$ is fixed, no learning becomes socially optimal when μ is sufficiently large.

Corollary 3. *Fix* $\bar{w} > 0$ *and vary* (μ, s) *subject to* $\mu s - (1-\mu) = \bar{w}$*. There exists a unique* $\mu^\dagger(\bar{w}) \in (0,1)$ *such that total surplus is maximized at* $k = 0$ *if* $\mu < \mu^\dagger(\bar{w})$ *and at* $k = \hat{k}$ *if*

[22] Although we characterize the mixed equilibrium for $k < \hat{k}$, Section 5 shows that, when $\bar{w} > 0$, an equilibrium exists at $k = \hat{k}$ and the equilibrium outcome is unique. The consumer does not learn and purchases from Firm 1, yielding $TS(\hat{k}) = \lim_{k\uparrow\hat{k}} TS(k)$.

[23] We have $TS(k) - TS(\hat{k}) = [1 - \mu - C(k)]\Pr(\alpha \le \Delta \le \beta) - \bar{w}\Pr(\Delta > \beta) < 0$ whenever $C(k) \ge 1 - \mu$.

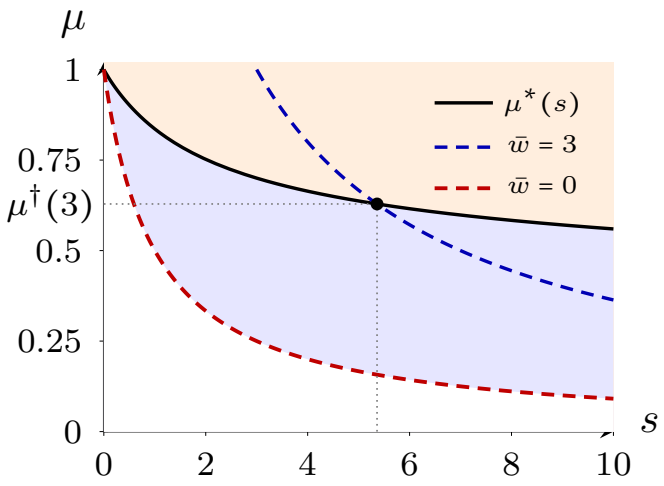


Figure 6: Welfare cutoff $\mu^*(s)$ and curves with fixed prior mean $\bar{w} = 0$ and $\bar{w} = 3$. The blue and orange regions indicate where total surplus is maximized at $k = 0$ and $k = \hat{k}$, respectively.

$\mu \geq \mu^\dagger(\bar{w})$.

Thus, for each fixed $\bar{w} > 0$, the curve $\mu = (1+\bar{w})/(1+s)$ intersects $\mu = \mu^*(s)$ exactly once, at $\mu = \mu^\dagger(\bar{w})$. Figure 6 illustrates this for $\bar{w} = 3$. In the region between the curves to the left of the intersection, no learning is optimal even when $\bar{w} < 3$. To the right, free information is optimal even when $\bar{w} > 3$. When μ is lower, Firm 2 is the better match more often and learning remains valuable despite a larger s.

Consumer surplus is defined as

$$CS(k) = TS(k) - \Pi_1(k) - \Pi_2(k),$$

where $\Pi_1(k) = \mu(b+\beta)$ and $\Pi_2(k) = (1-\mu)(a-\alpha)$.

Proposition 4. *In equilibrium, consumer surplus is strictly increasing in k.*

Notably, the consumer is better off even though learning becomes more expensive. A higher information cost reduces product differentiation, inducing firms to compete more aggressively on price. The proof establishes a stronger result than the first-order stochastic dominance in Lemma 2: for each firm i and every quantile $u \in (0,1)$, $Q_i(u,k) + C(k)$ is strictly decreasing in k, where $Q_i(u,k)$ denotes Firm i's uth price quantile. Thus, at every quantile, the decline in price exceeds the increase in the cost of learning.

To see why this implies higher consumer surplus, the consumer's equilibrium payoff at given prices can be written as

$$\max\Big\{\mathbb{E}[v_1] - p_1,\ \mathbb{E}[v_2] - p_2,\ \mathbb{E}[\max\{v_1, v_2\}] - \mu p_1 - (1-\mu)p_2 - C(k)\Big\}.$$

The first two terms are the payoffs from purchasing from Firm 1 and Firm 2 without learning, respectively, while the third represents the payoff from learning. For any fixed pair of quantiles (u_1, u_2), both no-learning payoffs increase with k because $Q_i(u_i, k)$ is strictly decreasing in k for $i = 1, 2$. The learning payoff also increases because $\mu[Q_1(u_1, k) + C(k)] + (1-\mu)[Q_2(u_2, k) + C(k)]$ is strictly decreasing in k. Thus, all three payoffs increase, and so does the consumer's expected surplus.

5 Quasi-Monopoly

As the information cost approaches the cutoff $\hat{k}$, the interior mixed equilibrium reaches its boundary. The consumer becomes increasingly likely to purchase from Firm 1 without learning, and Firm 1's equilibrium demand converges to one, while Firm 2's converges to zero. We refer to this no-learning outcome as the *quasi-monopoly* regime.[24]

Proposition 5. *Suppose that $\bar{w} > 0$ and $k \in [\hat{k}, \bar{k})$. Under any admissible tie-breaking rule, there is an equilibrium in which Firm 1 sets $p_1 = \alpha$ and Firm 2 mixes on $[0, L]$ according to*

$$F_2(p) = \frac{p}{(1-\mu)(\alpha + p)}, \qquad p \in [0, L),$$

with $F_2(L) = 1$. The consumer does not acquire information and purchases from Firm 1 almost surely. Equilibrium profits and welfare satisfy

$$(\Pi_1, \Pi_2) = (\alpha, 0), \qquad TS = \mathbb{E}[v_1], \qquad CS = \mathbb{E}[v_1] - \alpha.$$

The equilibrium in Proposition 5 is obtained as the limit of the interior mixed-strategy equilibrium at $\hat{k}$. In this equilibrium, Firm 1 serves the entire market and Firm 2 earns zero profit. Without learning, the consumer would choose based on her prior, and Firm 1 could charge $p_1 = \bar{w}$, its prior advantage. When learning is possible, however, Firm 1 must charge a lower price to keep the consumer uninformed. Although Firm 2 receives zero demand in equilibrium, if Firm 1 charges more than α, Firm 2 can profitably deviate to a sufficiently low price, inducing the consumer to learn or purchase from Firm 2 without learning. Thus, Firm 1 charges $p_1 = \alpha < \bar{w}$, the highest price at which it can keep the consumer uninformed and serve the entire market against competitive pressure from Firm 2.[25]

Equilibrium price strategies in this region are generally not unique.[26] Nevertheless, the

[24] When $\bar{w} = 0$, this regime does not arise since $\hat{k} = \bar{k}$.

[25] A similar structure arises in the asymmetric-cost Bertrand equilibrium of Blume (2003). There, the low-cost firm serves the market, while the high-cost firm earns zero profit but randomizes in a way that disciplines the low-cost firm's price.

[26] Let $\tilde{F}_2$ be any CDF supported on $[0, L]$ such that $\tilde{F}_2(0) = 0$ and $\tilde{F}_2(p) \geq F_2(p)$ for every $p \in [0, L)$. Then $p_1 = \alpha$ together with $\tilde{F}_2$ is also an equilibrium under any admissible tie-breaking rule. Relative to F_2, $\tilde{F}_2$ puts weakly more probability on lower prices, which can only reduce Firm 1's expected demand from any

equilibrium outcome is unique: the consumer remains uninformed and purchases from Firm 1 with probability one, Firm 1 charges $p_1 = \alpha$, and the firms earn $(\Pi_1, \Pi_2) = (\alpha, 0)$. Thus the allocation, transaction price, payoffs, and welfare are the same across all equilibria.[27]

A higher information cost strengthens Firm 1's dominant position by deterring consumer learning. As learning becomes more costly, the consumer is willing to rely on the prior, allowing Firm 1 to strategically exploit this by raising the price α while continuing to serve the entire market. This higher price reduces consumer surplus. It does not, however, change the allocation: the consumer remains uninformed and purchases from Firm 1 almost surely throughout the quasi-monopoly regime. Since no information is acquired, the higher information cost is never incurred. Total surplus therefore remains $\mathbb{E}[v_1]$ and the increase in Firm 1's price exactly redistributes surplus from the consumer to Firm 1.

Figure 7 summarizes consumer surplus and total surplus over the full range of information costs. Consumer surplus is uniquely maximized at the threshold $\hat{k}$: it is strictly increasing along the interior mixed-equilibrium branch and strictly decreasing throughout the quasi-monopoly regime. Thus, the consumer benefits most when price competition is strongest, even though only one firm serves the market. Total surplus, by contrast, remains constant throughout the quasi-monopoly regime. As market asymmetry increases, the information cost that maximizes total surplus shifts from $k = 0$ to the no-learning region $[\hat{k}, \bar{k}]$. Thus, the no-learning outcome at $k = \hat{k}$ can maximize both consumer and total surplus.

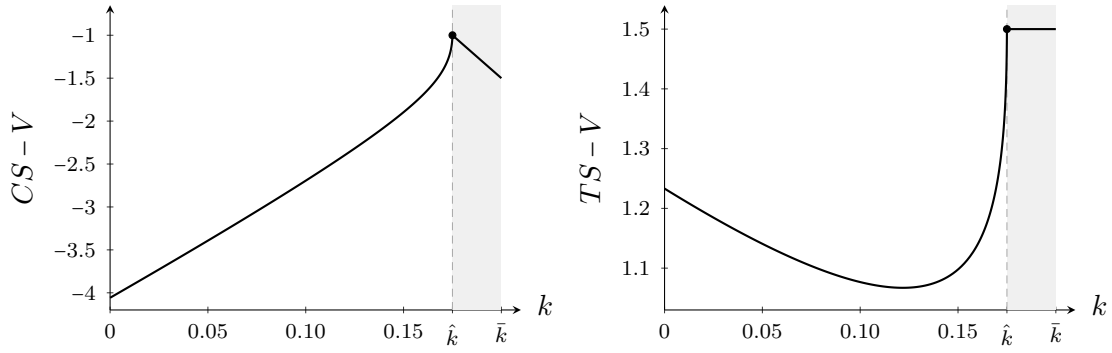


Figure 7: Consumer and total surplus relative to V for $(\mu, s) = (0.8, 4)$ over $k \in [0, \bar{k}]$.

Lemma 6. *For $k \in [0, \bar{k})$, industry profit is uniquely maximized at $k = 0$. Moreover, even Firm 1's profit is uniquely maximized at $k = 0$.*

As k rises along the interior mixed-equilibrium branch, stronger price competition reduces both firms' profits. In the quasi-monopoly regime, Firm 2 earns zero profit, while Firm 1's

upward deviation, while $\tilde{F}_2(0) = 0$ ensures that Firm 1 sells almost surely at $p_1 = \alpha$ and earns α. Firm 2 remains a best response because it earns zero at every price against $p_1 = \alpha$. If $d_\alpha = 1$, the pure price profile $(p_1, p_2) = (\alpha, 0)$ is also an equilibrium with the same outcome.

[27] We establish this result in Proposition A.1.

profit increases with k. Interestingly, Firm 1's profit is nevertheless uniquely maximized at $k = 0$, as illustrated in Figure 8. Under free information, consumer learning supports higher prices, making the competitive market more profitable for Firm 1 than any quasi-monopoly outcome, even though it may earn less than Firm 2.[28]

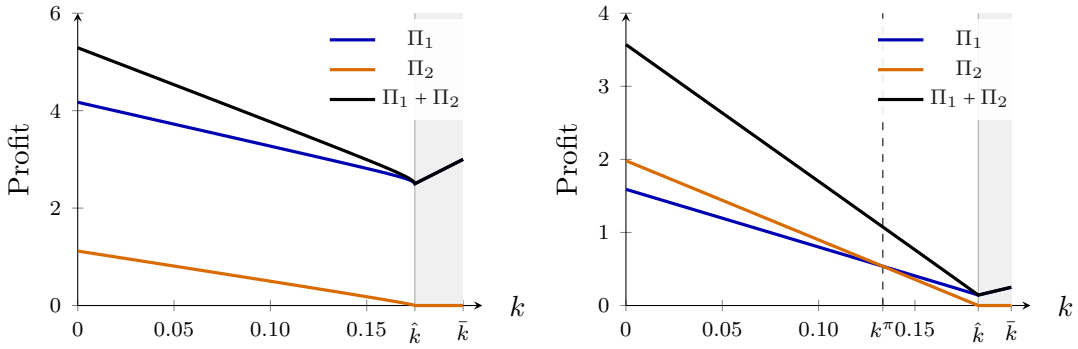


Figure 8: Equilibrium profits for $(\mu, s) = (0.8, 4)$ (left) and $(\mu, s) = (0.25, 4)$ (right). In the right panel, the two firms' profits are equal at $k = k^\pi$.

6 Endogenous Transparency

So far, we have treated the information friction as exogenous. In many markets, however, a firm can influence how easily consumers compare its product with competing products, as illustrated by the examples of Gladia and T-Mobile discussed in the introduction. This can occur through access to performance information, product trials, comparison tools, or other features of the information environment. Greater transparency can help the consumer choose the better-matched product, but consumer learning can also support higher equilibrium prices. We examine private and social incentives for transparency by allowing Firm 1 to choose transparency at a cost.

Environment and timing. Firm 1 chooses the information friction $k \in [0, K]$, where a lower k represents greater transparency. We assume $K \geq 1 (> \bar{k})$, so that the feasible set includes values of k for which no learning is optimal regardless of prices. We refer to $k = 0$ as full transparency and $k = K$ as full opacity. Firm 1 incurs a cost $c(k)$ of providing transparency, which may capture the resources needed to collect and provide fit-relevant information. We assume that $c(\cdot)$ is twice continuously differentiable on $[0, K]$ and $c'(k) < 0 < c''(k)$ for $k \in (0, K)$. Thus, providing more transparency is more costly.

The game proceeds as follows:

[28] At $k = 0$, $\Pi_2(0) > \Pi_1(0)$ if and only if $\mu < 1/(1+\sqrt{s})$. In this region, Firm 1 has a larger but less frequent match advantage, so Firm 2 benefits from being the better match more often when information is cheap. If $\bar{w} > 0$, the profits cross exactly once at some $k^\pi \in (0, \hat{k})$, as higher information costs favor Firm 1's prior advantage.

1. Firm 1 publicly chooses k and incurs the cost $c(k)$.

2. The two firms simultaneously choose prices.

3. The consumer observes both prices, decides whether to learn, and chooses a product.

Since the provision cost is sunk before prices are chosen, the pricing subgame coincides with the baseline game.[29] Let k^f maximize Firm 1's net profit $\tilde{\Pi}_1(k) = \Pi_1(k) - c(k)$, and let k^c and k^w maximize consumer surplus and net total surplus $TS(k) - c(k)$, respectively.

When transparency is costless, Firm 1 uniquely chooses full transparency. This is a direct consequence of Lemma 6.

Corollary 4. *If $c(k) \equiv 0$, then $k^f = 0$.*

Consumer learning differentiates the products and weakens price competition, and hence transparency is privately valuable to Firm 1. With costly provision, Firm 1 trades off this benefit against the cost of providing transparency. We show that this tradeoff generically leads to a unique choice.

Proposition 6. *Firm 1's optimal transparency choice k^f is generically unique. Moreover, $k^f \leq k^w$.*

Below $\hat{k}$, greater transparency raises Firm 1's profit but is more costly to provide. Its net profit is strictly concave on the interior mixed branch. For $k \geq \hat{k}$, a higher k weakly raises Firm 1's profit and lowers the provision cost, so only full opacity can be optimal in this region. Hence, its choice is unique unless K and the unique maximizer on $[0, \hat{k}]$ yield the same net profit.

The second part of the proposition highlights a difference between private and social incentives for transparency. A lower k can improve product matching by allowing the consumer to switch to Firm 2 when it is the better match. Total surplus can thus increase. By contrast, Firm 1 does not value this reallocation itself. It benefits because consumer learning supports higher prices. We show that $TS(k) - \Pi_1(k)$ is strictly increasing when both firms are active: whenever lowering k increases total surplus, Firm 1's profit increases by even more. We also show that if Firm 1 chooses full opacity, full opacity maximizes net total surplus. Since both objectives include the same provision cost, Firm 1 chooses weakly more transparency than is socially optimal.

This argument becomes clearer when $\bar{w} = 0$. In this case, we can write

$$\Pi_1(k) = \lambda(k)\Pi_1(0), \qquad TS(k) = TS(0) - ksP_0,$$

[29] For $k \geq \bar{k}$, Proposition C.1 shows that Firm 1 earns its prior advantage $\bar{w}$ without learning, and equilibrium outcomes are uniquely determined. Together with the outcome uniqueness in the quasi-monopoly regime established in Proposition A.1, the continuation values are well defined for every $k \in [0, K]$.

for $k \leq \bar{k}$. Notice that the two marginal gains from transparency are

$$|\Pi_1'(k)| = \underbrace{b_0 + s}_{\text{marginal private gain}} > \underbrace{sP_0}_{\text{marginal social gain}} = |TS'(k)|,$$

where $b_0 + s$ is Firm 1's switching price and b_0 is Firm 2's lowest equilibrium price at $k = 0$.

When $\bar{w} = 0$, changing $k < \bar{k}$ simply rescales the equilibrium and the allocation remains unchanged. Total surplus rises only because information is cheaper when the consumer learns. A lower k raises Firm 1's profit through weaker competition: consumer learning allows both its match advantage and its rival's equilibrium price to be reflected in its own price. The inequality above shows that this pricing effect makes Firm 1's marginal gain from transparency larger than the corresponding gain in total surplus. Consumer surplus is maximized at every $k \in [\bar{k}, K]$: it rises with k and reaches its maximum at $\bar{k}$. For all $k \geq \bar{k}$, both equilibrium prices are zero and consumer surplus remains unchanged.

When $\bar{w} > 0$, consumer surplus is uniquely maximized at $k^c = \hat{k}$. The following lemma shows how different sources of market asymmetry affect the firm's transparency choice relative to the consumer and social optima.

Lemma 7. *The followings hold.*

(i) Fix $s > 0$. For all μ sufficiently close to one, $k^c = \hat{k} < k^f = k^w = K$.

(ii) Fix $\mu \in (0, 1)$. For all sufficiently large s, $k^f = 0 < k^c = \hat{k} < k^w = K$.

In both parts, k^f, k^c, and k^w are unique.

When μ is close to one, the prior is already a good guide to product choice, and learning adds little matching value. Given the cost of providing transparency, full opacity is socially optimal. The value of learning to Firm 1 is also small: as $\mu \to 1$, Firm 1's profit converges to s, its match advantage, for $k \in [0, K]$. Thus, whether or not the consumer learns, Firm 1 can essentially extract the same match advantage, and the additional profit created by learning vanishes. Firm 1 therefore saves on provision costs by choosing full opacity. From the consumer's perspective, however, Firm 1 provides too little transparency: raising k beyond k^c does not change the allocation but allows Firm 1 to charge a higher no-learning price.

A larger s has a different effect. Firm 2 remains the better match with probability $1-\mu > 0$, and learning continues to differentiate the products, while Firm 1 has a larger match advantage when its product is preferred. Consumer learning allows Firm 1 to reflect this larger match advantage in its price, making transparency increasingly valuable. The resulting profit gain eventually makes full transparency optimal for Firm 1. When s is sufficiently large, total surplus is maximized in the no-learning region, as shown in Proposition 3. Hence, full opacity maximizes net total surplus. As a result, Firm 1 provides more transparency than at either k^c or k^w. Figure 9 illustrates the two cases.

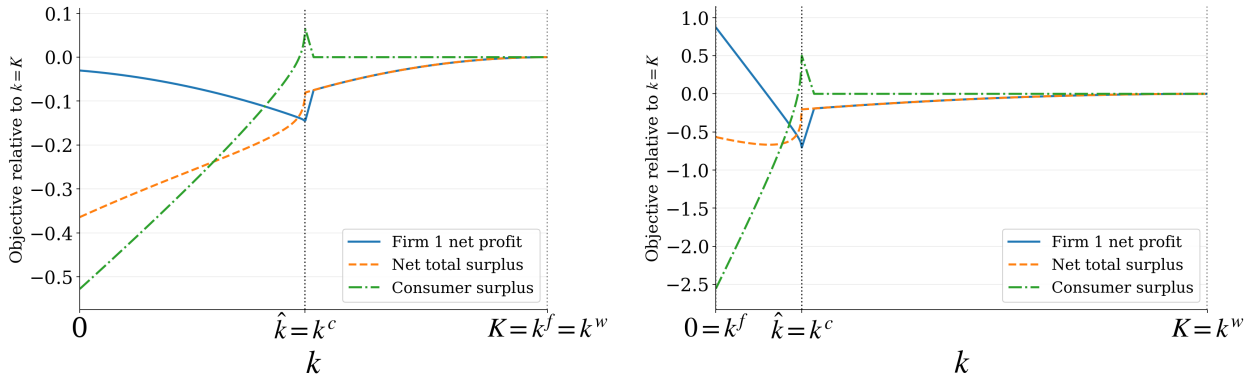


Figure 9: Transparency choices when $\bar{w} > 0$. $K = 1$ and $c(k) = 0.3(1-k)^2$. The left panel illustrates a large probability advantage: with $(\mu, s) = (0.91, 1)$, it has $k^c = \hat{k} \simeq 0.481 < k^f = k^w = K$. The right panel illustrates a large match advantage: with $(\mu, s) = (0.8, 4)$, it has $k^f = 0 < k^c = \hat{k} = 0.175 < k^w = K$. Each curve is measured relative to its value at $k = K$.

7 General Match Values and Flexible Learning

In the baseline model, the consumer either acquires no information or learns her match value perfectly. We now allow w to have a general distribution and let the consumer choose any Blackwell experiment after observing prices. She can therefore choose how much to learn, rather than only whether to learn. Under a quadratic information cost, we characterize a unique pure-strategy pricing equilibrium for sufficiently high information costs and, under a uniform prior, for all information costs. We show that the main qualitative results of the baseline model continue to hold.

Let w have a nondegenerate distribution F with $\mathbb{E}_F|w| < \infty$ and mean $\bar{w} \geq 0$. The firms simultaneously choose prices $p_i \in [0, \bar{p}]$. We assume $V > \bar{p}$, so the market is covered. After observing prices, the consumer chooses a Blackwell experiment about w, which specifies a distribution over signals conditional on each realization of w. The state is then realized, she observes the signal, and updates her belief about w by Bayes' rule. Since product values are linear in w, the consumer's expected utility from either purchase depends on her posterior belief only through its mean. Let m denote the posterior mean and G the induced distribution of m.

We assume that the information cost is proportional to the variance of posterior means:

$$C(G; k) = \frac{k}{2}\mathbb{E}_G[(m - \bar{w})^2]$$

for $k > 0$.[30] No learning gives $m = \bar{w}$ with probability one and costs zero. Under full revelation, $m = w$, and the cost reduces to $\frac{k}{2}\mathrm{Var}_F(w)$, the natural analogue of baseline

[30]For general characterizations of posterior-separable and posterior-mean separable information costs, see Caplin, Dean, and Leahy (2022) and Mensch and Malik (2024), respectively.

information cost. At $k = 0$, information is free, and we define $C(G;0) = 0$. We continue to use subgame perfect equilibrium and focus on equilibria in pure prices.

Consumer information choice. Let $\Delta = p_1 - p_2$. Given prices, the consumer purchases from Firm 1 when $m \geq \Delta$ and from Firm 2 when $m < \Delta$. After dropping the term $V - \bar{w}/2 - p_2$, her information problem is

$$\sup_{G \in \mathrm{MPC}(F)} \mathbb{E}_G[(m-\Delta)_+] - C(G;k), \tag{7}$$

where $x_+ = \max\{x, 0\}$ and $\mathrm{MPC}(F)$ denotes the set of mean-preserving contractions of F, or equivalently the distributions smaller than F in convex order.[31] Note that when $k = 0$, $G = F$ is always optimal since $(m-\Delta)_+$ is convex.

Consider an experiment with several signals that lead to purchase from the same firm. Pooling these signals into one replaces their posterior means by a weighted average. Because the payoff is linear in the posterior mean, the expected purchase payoff is unchanged, and the optimal purchase remains unchanged. Pooling also weakly reduces the variance of posterior means and hence the information cost. Thus, it suffices to consider experiments with at most two signals, each inducing a purchase from a different firm.

Section A.5 solves the consumer's problem using these two-signal experiments. For every prior F satisfying the assumptions above, there exists a finite $k_F > 0$ such that demand is uniquely determined at every feasible price pair whenever $k \geq k_F$. For such k, let $L = L(k) = 1/k$.

$$D_1(\Delta) = \begin{cases} 1, & \Delta \leq \bar{w} - L/2, \\ \frac{1}{2} + \frac{\bar{w}-\Delta}{L}, & |\Delta - \bar{w}| < L/2, \\ 0, & \Delta \geq \bar{w} + L/2, \end{cases} \qquad D_2(\Delta) = 1 - D_1(\Delta). \tag{8}$$

The consumer learns only when the price difference Δ is sufficiently close to $\bar{w}$. Here, L is the length of this learning interval. As information becomes more costly, the learning interval narrows, and demand becomes more sensitive to relative prices: for example, within the learning interval, the slope of Firm 1's demand with respect to Δ is $-1/L = -k$.[32] Thus consumer learning provides less product differentiation and weakens price competition.

Define $k^\dagger = 3/(2\bar{w})$ when $\bar{w} > 0$, and $k^\dagger = \infty$ otherwise. As in the baseline model, the equilibrium features two regimes: an interior equilibrium in which both firms earn positive profits for $k < k^\dagger$, and a quasi-monopoly regime in which only Firm 1 serves the market without learning for $k \geq k^\dagger$.

[31] An experiment under prior F can generate a distribution of posterior means G if and only if $G \in \mathrm{MPC}(F)$; see Strassen (1965) and Gentzkow and Kamenica (2016).

[32] Unlike in the baseline, L remains positive for every finite k. With quadratic information costs, the benefit of acquiring a small amount of information can exceed its cost, even when k is large. Thus, some learning remains optimal when the price difference is sufficiently close to $\bar{w}$.

Proposition 7. *Suppose that $k \geq k_F$ and $\bar{p} > \bar{w} + 1/k_F$. There is a unique pure-strategy pricing equilibrium.*[33]

(i) For $k < k^\dagger$, equilibrium prices are $(p_1^*, p_2^*) = (L/2 + \bar{w}/3, L/2 - \bar{w}/3)$ with equilibrium demands $\bar{D}_i = p_i^*/L$ and profits $\Pi_i = (p_i^*)^2/L$ for $i = 1, 2$. Total surplus and consumer surplus are $TS = V + L/8 + 5\bar{w}^2/(18L)$ and $CS = V - 3L/8 + \bar{w}^2/(18L)$.

(ii) For $k \geq k^\dagger$, equilibrium prices are $(p_1^*, p_2^*) = (\bar{w} - L/2, 0)$ with equilibrium demands $(\bar{D}_1, \bar{D}_2) = (1, 0)$ and profits $(\Pi_1, \Pi_2) = (\bar{w} - L/2, 0)$. The consumer buys from Firm 1 without learning. Total surplus and consumer surplus are $TS = V + \bar{w}/2$ and $CS = V - \bar{w}/2 + L/2$.

Under the uniform prior $w \sim U[\bar{w} - 1, \bar{w} + 1]$, where $0 \leq \bar{w} < 1$, we can characterize the equilibrium for every $k \geq 0$. In this case, the learning interval has length $L(k) = 2 - k$ for $0 \leq k \leq 1$ and $L(k) = 1/k$ for $k > 1$.

Lemma 8. *Under the uniform prior, suppose that $\bar{p} > 2$. For every $k \geq 0$, there is a unique pure-strategy pricing equilibrium. The demand function in (8) and the equilibrium outcomes in Proposition 7 apply with $L = L(k)$.*[34]

For the remainder of this section, we focus on the uniform prior for comparative statics with respect to k. In the interior equilibrium, the price difference is $\Delta^* = 2\bar{w}/3$ for $k < k^\dagger$. This difference generates misallocation even when information is free. To see this, focus on the full-revelation experiment at $k = 0$.[35] In equilibrium, the consumer purchases from Firm 1 if $w \geq \Delta^*$ and from Firm 2 otherwise. Under the first-best allocation, however, the consumer purchases from Firm 1 when $w \geq 0$ and from Firm 2 otherwise. Misallocation arises when $\bar{w} > 0$: with positive probability, the consumer purchases from Firm 2 even though Firm 1 is the better match.[36] Thus, as in the baseline model, equilibrium price differences prevent the equilibrium from attaining the first best even under costless learning. Here, however, there is no price dispersion, and market asymmetry generates misallocation in only one direction, as shown in Figure 10.

When $k < k^\dagger$, Firm 1's prior advantage exceeds the equilibrium price difference: $\bar{w} - \Delta^* = \bar{w}/3 > 0$ when $\bar{w} > 0$. Since L decreases with k, its equilibrium demand $\bar{D}_1 = 1/2 + (\bar{w} - \Delta^*)/L$

[33] Note that, once $k \geq k_F$, demand and equilibrium outcomes depend on F only through $\bar{w}$. For lower information costs, the feasibility constraint on posterior means may bind, so the consumer's optimal information choice and demand can depend on the shape of F. As a result, both the existence of a pure-strategy equilibrium and its outcomes may also depend on the shape of F, for example, on its curvature.

[34] The equilibrium prices are unique, but the consumer's optimal experiment need not be unique, since several experiments may induce the same optimal distribution of posterior means.

[35] In the proof, we show that every optimal experiment induces the same allocation almost surely at $k = 0$.

[36] When $\bar{w} = 0$, the consumer always purchases her preferred product, and the equilibrium achieves the first best.

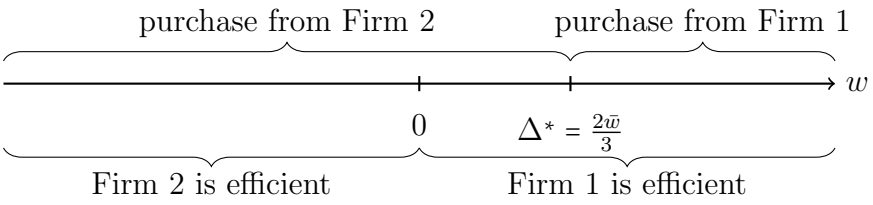


Figure 10: Product choice when information is free.

increases, raising the allocation effect in (5). The matching effect, however, decreases as the consumer relies more on her prior rather than the realized match value.[37]

Indeed, total surplus is maximized either at $k = 0$ or in the no-learning regime. An intermediate information cost provides neither the full matching gain from free information nor the allocation gain from no learning, while the consumer still incurs information costs.[38] We have

$$TS(0) - TS(k^\dagger) = \frac{(3-\bar{w})(3-5\bar{w})}{36}.$$

Thus, when market asymmetry is sufficiently large, the no-learning outcome is optimal because the allocation gain is large enough. Otherwise, free information is optimal.

In the interior equilibrium, a higher k intensifies price competition, lowering prices enough to raise consumer surplus despite the higher information cost. In the quasi-monopoly regime, consumer surplus falls as Firm 1 raises its price. When $\bar{w} > 0$, it is therefore uniquely maximized at $k = k^\dagger$.

We summarize these results in the following lemma.

Lemma 9. *Under the uniform prior,*

(i) Even when $k = 0$, the equilibrium fails to achieve the first best when $\bar{w} > 0$.

(ii) Total surplus is uniquely maximized at $k = 0$ when $\bar{w} < 3/5$. When $\bar{w} \geq 3/5$, it is maximized for every $k \geq k^\dagger$.

(iii) Consumer surplus is uniquely maximized at $k = k^\dagger$ when $\bar{w} > 0$.

8 Concluding Remarks

This paper studies price competition when the consumer observes prices before deciding whether to acquire costly information about product fit. When the information cost is below

[37] In the baseline model, the matching effect need not decrease with k because the probability of learning can increase. Here, the consumer always learns in the interior equilibrium, but a higher k shifts purchases toward Firm 1, making product choices less correlated with realized match values.

[38] When $\bar{w} = 0$, the allocation effect vanishes. For any $k > 0$, the matching effect is weakly lower and learning incurs a positive information cost, so free information maximizes total surplus.

a threshold, we show that no pure-strategy pricing equilibrium exists and fully characterize the unique mixed-strategy equilibrium. An important implication is that price dispersion arising from firms' strategic interaction can lead to inefficient product choices even when information is free, unlike in the monopoly benchmark.

The welfare analysis shows that the effects of information costs depend on firms' pricing responses. While both firms are active, higher information costs intensify price competition and raise consumer surplus. However, the effect on total surplus depends on market asymmetry. Free information is optimal when Firm 1's advantage is relatively small, because learning improves product matching. With sufficiently strong asymmetry, no learning can instead be optimal because the gain from shifting demand toward Firm 1 is large. Once information costs become high enough, the consumer stops learning and purchases from Firm 1. Further increases allow the firm to raise its price and reduce consumer surplus. Thus, easier access to information need not improve consumer or total surplus.

Allowing Firm 1 to choose transparency further shows that its private incentives need not align with welfare. It values learning because it can support higher prices, but does not internalize its social value for product matching. Firm 1 may therefore choose full transparency even when full opacity maximizes total surplus. Depending on the source of market asymmetry, it may also provide either more or less transparency than the consumer would prefer.

There are two natural directions for future research. One is to extend our endogenous transparency model by introducing an information design problem in which the advantaged firm designs a signal about product fit before prices are set. After observing the signal and prices, the consumer could still acquire additional information at a cost. The firm's disclosure can therefore affect both the consumer's learning incentives and price competition. This would allow us to study how market asymmetry shapes optimal disclosure and when profit-driven information provision improves product matching and welfare. Another direction is to extend our analysis of general match values and flexible learning beyond quadratic information costs. Under suitable assumptions on the cost function, it may be possible to establish the existence and uniqueness of a pure-strategy equilibrium for sufficiently high information costs. We leave these extensions for future work.

References

Albrecht, Brian C., and Mark Whitmeyer (2026), "Efficient Learning under Competition," *working paper*.

Anderson, Simon P., and Régis Renault (1999), "Pricing, Product Diversity, and Search Costs: A Bertrand-Chamberlin-Diamond Model," *RAND Journal of Economics 30 (4)*, 719-735.

Anderson, Simon P., and Régis Renault (2009), "Comparative Advertising: Disclosing Horizontal Match Information," *RAND Journal of Economics 40 (3)*, 558-581.

Armstrong, Mark, and John Vickers (2019), "Discriminating against Captive Customers," *American Economic Review: Insights 1 (3)*, 257-272.

Armstrong, Mark, and John Vickers (2022), "Patterns of Competitive Interaction," *Econometrica 90 (1)*, 153-191.

Armstrong, Mark, and Jidong Zhou (2011), "Paying for Prominence," *The Economic Journal 121 (556)*, 368-395.

Armstrong, Mark, and Jidong Zhou (2022), "Consumer Information and the Limits to Competition," *American Economic Review 112 (2)*, 534-577.

Au, Pak Hung, and Mark Whitmeyer (2023), "Attraction versus Persuasion: Information Provision in Search Markets," *Journal of Political Economy 131 (1)*, 202-245.

Bar-Isaac, Heski, Guillermo Caruana, and Vicente Cuñat (2010), "Information Gathering and Marketing," *Journal of Economics & Management Strategy 19 (2)*, 375-401.

Baye, Michael R., Dan Kovenock, and Casper G. de Vries (1992), "It Takes Two to Tango: Equilibria in a Model of Sales," *Games and Economic Behavior 4 (4)*, 493-510.

Biglaiser, Gary, Jiadong Gu, and Fei Li (2025), "Information Acquisition and Product Differentiation Perception," *American Economic Journal: Microeconomics 17 (3)*, 1-34.

Blume, Andreas (2003), "Bertrand without Fudge," *Economics Letters 78 (2)*, 167-168.

Board, Simon, and Jay Lu (2018), "Competitive Information Disclosure in Search Markets," *Journal of Political Economy 126 (5)*, 1965-2010.

Branco, Fernando, Monic Sun, and J. Miguel Villas-Boas (2012), "Optimal Search for Product Information," *Management Science 58 (11)*, 2037-2056.

Caplin, Andrew, Mark Dean, and John Leahy (2022), "Rationally Inattentive Behavior: Characterizing and Generalizing Shannon Entropy," *Journal of Political Economy 130 (6)*, 1676-1715.

Chen, Yanbin, Sanxi Li, Kai Lin, and Jun Yu (2021), "Consumer Search with Blind Buying," *Games and Economic Behavior 126*, 402-427.

Choi, Michael, Anovia Yifan Dai, and Kyungmin Kim (2018), "Consumer Search and Price Competition," *Econometrica 86 (4)*, 1257-1281.

Elliott, Matthew, Andrea Galeotti, Andrew Koh, and Wenhao Li (2026), "Market Segmentation Through Information," *The Review of Economic Studies*, forthcoming.

Gentzkow, Matthew, and Emir Kamenica (2016), "A Rothschild-Stiglitz Approach to Bayesian Persuasion," *American Economic Review 106 (5)*, 597-601.

Gu, Zheyin (Jane), and Ying Xie (2013), "Facilitating Fit Revelation in the Competitive Market," *Management Science 59 (5)*, 1196-1212.

Haan, Marco A., José L. Moraga-González, and Vaiva Petrikaitė (2018), "A Model of Directed Consumer Search," *International Journal of Industrial Organization 61*, 223-255.

Hwang, Ilwoo, Kyungmin Kim, and Raphael Boleslavsky (2026), "Competitive Advertising and Pricing," *The Review of Economic Studies*, forthcoming.

Jain, Vasudha, and Mark Whitmeyer (2021), "Search and Competition with Flexible Investigations," *working paper (arXiv:2104.13159)*.

Jain, Vasudha, and Mark Whitmeyer (2026), "Competitive Disclosure of Information to a Rationally Inattentive Agent," *Journal of Economic Behavior & Organization 242*, Article 107410.

Jing, Bing, and Zhong Wen (2008), "Finitely Loyal Customers, Switchers, and Equilibrium Price Promotion," *Journal of Economics & Management Strategy 17 (3)*, 683-707.

Johnen, Johannes, and David Ronayne (2021), "The Only Dance in Town: Unique Equilibrium in a Generalized Model of Price Competition," *Journal of Industrial Economics 69 (3)*, 595-614.

Kaye, Aaron P. (2026), "The Personalization Paradox: Welfare Effects of Personalized Recommendations in Two-Sided Digital Markets," *working paper*.

Li, Yupeng, Ben Chen, Mingyue Cheng, Qingpeng Cai, Zhiding Liu, Xuxin Zhang, Zhentao Song, Chenyi Lei, Qi Liu, and Peng Jiang (2026), "KuaiSearch: An E-Commerce Search Dataset with Authentic Queries and Product Texts for Recall, Ranking, and Relevance," *arXiv:2602.11518*.

Mensch, Jeffrey, and Komal Malik (2024), "Posterior-Mean Separable Costs of Information Acquisition," *working paper (arXiv:2311.09496)*.

Mensch, Jeffrey, Doron Ravid, and João Thereze (2026), "Monopoly, Quality, and Costly Information," *working paper*.

Minaev, Andrey (2021), "Consumer Data and Consumer Welfare: Evidence from the Hotel Booking Market," *working paper*.

Moscarini, Giuseppe, and Marco Ottaviani (2001), "Price Competition for an Informed Buyer," *Journal of Economic Theory 101 (2)*, 457-493.

Myatt, David P., and David Ronayne (2026), "Asymmetric Models of Sales," *American Economic Journal: Microeconomics 18 (1)*, 281-311.

Narasimhan, Chakravarthi (1988), "Competitive Promotional Strategies," *The Journal of Business 61 (4)*, 427-449.

Obradovits, Martin, and Philipp Plaickner (2023), "Price-Directed Search, Product Differentiation and Competition," *Review of Industrial Organization 63 (3)*, 317-348.

Raju, Jagmohan S., V. Srinivasan, and Rajiv Lal (1990), "The Effects of Brand Loyalty on Competitive Price Promotional Strategies," *Management Science 36 (3)*, 276-304.

Ravid, Doron, Anne-Katrin Roesler, and Balázs Szentes (2022), "Learning before Trading: On the Inefficiency of Ignoring Free Information," *Journal of Political Economy 130 (2)*, 346-387.

Roesler, Anne-Katrin, and Balázs Szentes (2017), "Buyer-Optimal Learning and Monopoly Pricing," *American Economic Review 107 (7)*, 2072-2080.

Rosenthal, Robert W. (1980), "A Model in Which an Increase in the Number of Sellers Leads to a Higher Price," *Econometrica 48 (6)*, 1575-1579.

Shilony, Yuval (1977), "Mixed Pricing in Oligopoly," *Journal of Economic Theory 14 (2)*, 373-388.

Stahl, Dale O., II (1989), "Oligopolistic Pricing with Sequential Consumer Search," *American Economic Review 79 (4)*, 700-712.

Strassen, Volker (1965), "The Existence of Probability Measures with Given Marginals," *The Annals of Mathematical Statistics 36 (2)*, 423-439.

Varian, Hal R. (1980), "A Model of Sales," *American Economic Review 70 (4)*, 651-659.

Wolinsky, Asher (1986), "True Monopolistic Competition as a Result of Imperfect Information," *The Quarterly Journal of Economics 101 (3)*, 493-511.

Appendix A Proofs

We first establish a lemma that will be used throughout the proofs.

Let $\alpha = -1 + k\mu(1+s)^2$, $\beta = s - k(1-\mu)(1+s)^2$, and $L = (1+s)[1-k(1+s)]$. Define

$$\tilde{x} := \frac{b+\beta-a}{L}, \qquad \tilde{y} := \frac{b+\beta}{L}, \qquad \tilde{b} := \frac{b}{L}, \qquad \tilde{\beta} := \frac{\beta}{L}.$$

Thus $a = L(\tilde{y}-\tilde{x})$, $b = L\tilde{b}$, and $a - \alpha = L(\tilde{b}+1-\tilde{x})$.

Under this normalization, the endpoint equations (3)-(4) become

$$(1-\mu)\tilde{y}^2 = \tilde{x}\tilde{y} + \mu\tilde{x}(1-\tilde{x}), \tag{9}$$

$$\mu\tilde{b}^2 = \big[(1-\mu) - (2\mu-1)\tilde{b}\big](1-\tilde{x}), \tag{10}$$

together with $\tilde{y} - \tilde{b} = \tilde{\beta}$.

Lemma A.1. *The normalized endpoint equations have the following properties.*

(i) For every $\mu \in (0,1)$ and $\tilde{x} \in (0,1)$, (9)-(10) have unique positive solutions $\tilde{y} = \tilde{y}(\tilde{x})$ and $\tilde{b} = \tilde{b}(\tilde{x})$. These solutions are continuously differentiable and satisfy $\tilde{y}'(\tilde{x}) > 1$ and $\tilde{b}'(\tilde{x}) < 0$. Hence $\tilde{y}(\tilde{x}) - \tilde{b}(\tilde{x})$ is strictly increasing from $-(1-\mu)/\mu$ to $1/(1-\mu)$. In particular, $0 < \tilde{y}(\tilde{x}) - \tilde{x} < \mu/(1-\mu)$ and $0 < \tilde{b}(\tilde{x}) < (1-\mu)/\mu$.

(ii) Suppose that $\bar{w} \geq 0$. For each $k \in [0, \bar{k})$, $\tilde{y}(\tilde{x}) - \tilde{b}(\tilde{x}) = \tilde{\beta}$ has a unique solution $\tilde{x}(k) \in (0,1)$ if and only if $k < \hat{k}$. For every such k, (3)-(4) have a unique strictly positive interior solution $(a(k), b(k))$. Along this solution, $\tilde{x}(k)$ is strictly increasing in k if $\bar{w} > 0$ and constant in k if $\bar{w} = 0$. Moreover, if $\bar{w} > 0$, then $\tilde{x}(k) \to 1$, $\tilde{b}(\tilde{x}(k)) \to 0$, and $\tilde{y}(\tilde{x}(k)) \to 1/(1-\mu)$ as $k \uparrow \hat{k}$.

Proof. Part (i). Fix $\tilde{x} \in (0,1)$. As equations (9) and (10) are quadratic in $\tilde{y}$ and $\tilde{b}$, respectively, with positive leading coefficients and negative constant terms, each has exactly one positive solution. Since the derivative with respect to $\tilde{y}$ or $\tilde{b}$ is nonzero at the positive root, the implicit function theorem gives continuously differentiable $\tilde{y}(\tilde{x})$ and $\tilde{b}(\tilde{x})$.

Equation (9) implies $(1-\mu)\tilde{y} > \tilde{x}$, and hence $\tilde{y} > \tilde{x}/(1-\mu) > \tilde{x}$. Differentiating (9)-(10) with respect to $\tilde{x}$ gives

$$[2(1-\mu)\tilde{y} - \tilde{x}](\tilde{y}' - 1) = \mu + (2\mu-1)(\tilde{y}-\tilde{x}),$$
$$[2\mu\tilde{b} + (2\mu-1)(1-\tilde{x})]\tilde{b}' = (2\mu-1)\tilde{b} - (1-\mu).$$

The coefficient multiplying $\tilde{y}' - 1$ is positive. If $\mu \geq 1/2$, the right-hand side of the first equation is positive. If $\mu < 1/2$, (9) implies $(1-2\mu)\tilde{x}(\tilde{y}-\tilde{x}) + (1-\mu)(\tilde{y}-\tilde{x})^2 = \mu\tilde{x}$, and hence $(1-2\mu)(\tilde{y}-\tilde{x}) < \mu$. Thus the right-hand side is positive in either case, so $\tilde{y}' > 1$. (10) implies

$(1-\mu)-(2\mu-1)\tilde{b}>0$. Also, dividing that equation by $\tilde{b}>0$ gives $\mu\tilde{b}+(2\mu-1)(1-\tilde{x})=(1-\mu)(1-\tilde{x})/\tilde{b}>0$. Thus the coefficient multiplying $\tilde{b}'$ is positive, while its right-hand side is negative, so $\tilde{b}'<0$.

As $\tilde{x}\downarrow 0$, equations (9)-(10) give $\tilde{y}\to 0$ and $\tilde{b}\to(1-\mu)/\mu$. As $\tilde{x}\uparrow 1$, they give $\tilde{y}\to 1/(1-\mu)$ and $\tilde{b}\to 0$. Since $\tilde{y}'-\tilde{b}'>0$, $\tilde{y}-\tilde{b}$ is strictly increasing from $-(1-\mu)/\mu$ to $1/(1-\mu)$. Since $\tilde{y}'-1>0$, $\tilde{y}-\tilde{x}$ is strictly increasing from 0 to $\mu/(1-\mu)$.

Part (ii). Suppose that $\bar{w}\geq 0$ and fix $k\in[0,\bar{k})$. Since $\beta=\bar{w}+(1-\mu)L$, we have $\tilde{\beta}\geq 1-\mu>0$. By part (i), $\tilde{y}(\tilde{x})-\tilde{b}(\tilde{x})=\tilde{\beta}$ has a unique solution in $(0,1)$ if and only if $\tilde{\beta}<1/(1-\mu)$. Equivalently, $L-(1-\mu)\beta>0$. Since $L-(1-\mu)\beta=1+\mu s-k\mu(2-\mu)(1+s)^2$, this holds if and only if $k<\hat{k}$. For such k, the endpoints are $a=L(\tilde{y}-\tilde{x})$ and $b=L\tilde{b}$, and hence are strictly positive by part (i). Moreover, $b+\beta-a=L\tilde{x}\in(0,L)$ and $a-\alpha-b=L(1-\tilde{x})\in(0,L)$, so $a<b+\beta<a+L$ and $b<a-\alpha<b+L$.

Differentiating $\tilde{y}(\tilde{x}(k))-\tilde{b}(\tilde{x}(k))=\tilde{\beta}$ with respect to k gives $(\tilde{y}'-\tilde{b}')\tilde{x}'(k)=(1+s)^2\bar{w}/L^2$. Since $\tilde{y}'-\tilde{b}'>0$, $\tilde{x}'(k)>0$ if $\bar{w}>0$, whereas $\tilde{x}'(k)=0$ if $\bar{w}=0$. Finally, suppose that $\bar{w}>0$. As $k\uparrow\hat{k}$, $\tilde{\beta}\to 1/(1-\mu)$. By part (i), $\tilde{y}(\tilde{x})-\tilde{b}(\tilde{x})$ is strictly increasing from $-(1-\mu)/\mu$ to $1/(1-\mu)$, so $\tilde{x}(k)\to 1$. Part (i) then gives $\tilde{b}(\tilde{x}(k))\to 0$ and $\tilde{y}(\tilde{x}(k))\to 1/(1-\mu)$. □

A.1 Equilibrium Characterization

Proof of Proposition 1. Since $\bar{k}-\hat{k}=(1-\mu)\bar{w}/[\mu(2-\mu)(1+s)^2]\geq 0$, we have $k<\hat{k}\leq\bar{k}$. Thus the demand characterization applies throughout the proof. In particular, $\beta>0$ and $L=\beta-\alpha>0$.

Step 1: Positive profits and price bounds. Consider any equilibrium in price distributions. Let F_i denote Firm i's equilibrium price distribution and Π_i its equilibrium profit. Also, let $\pi_i(p_i)$ denote its expected profit from charging p_i. Firm 1 can charge any $p_1\in(0,\beta)$. Since $p_1-p_2<\beta$ for every $p_2\geq 0$, its demand is at least μ, thus $\Pi_1\geq\mu p_1>0$.

Suppose that $\Pi_2=0$. If $\Pr(p_1>\alpha)>0$, there is a sufficiently small $q>0$ such that $\Pr(p_1>\alpha+q)>0$. Firm 2 could then charge $p_2=q$ and obtain demand at least $1-\mu$ on that event, earning positive profit, a contradiction. Hence $p_1\leq\alpha$ almost surely. Since profit cannot exceed price, $\Pi_1\leq\mathbb{E}[p_1]\leq\alpha$. On the other hand, letting $p_1\uparrow\beta$ in $\Pi_1\geq\mu p_1$ gives $\Pi_1\geq\mu\beta$. This contradicts

$$\alpha-\mu\beta=\mu(2-\mu)(1+s)^2\big(k-\hat{k}\big)<0.$$

Therefore $\Pi_2>0$.

Let $S_i:=\operatorname{supp}(F_i)$. Since S_i is a closed subset of the compact price set $[0,\bar{p}]$, it is compact. Write

$$a:=\min S_1,\quad A:=\max S_1,\quad b:=\min S_2,\quad B:=\max S_2.$$

A support point need not yet be a pure best response because profits may be discontinuous. However, each firm's equilibrium strategy assigns probability one to its pure best responses. Every neighborhood of a support point has positive probability and therefore contains a pure best response. Thus, for every $p_i \in S_i$, there is a sequence of pure best responses $p_i^n \to p_i$. Since $\Pi_i = \pi_i(p_i^n) \le p_i^n$, taking limits gives $p_i \ge \Pi_i > 0$. In particular, $a, b > 0$. This also allows us to choose pure best responses converging to A and B below.

We next show that the firms cannot both have atoms at prices whose difference is α or β. Suppose Firm 1 has an atom of size $\gamma_1 > 0$ at p_1 and Firm 2 has an atom of size $\gamma_2 > 0$ at p_2. If $p_1 - p_2 = \alpha$, a small price cut by Firm 1 raises its demand on Firm 2's atom by $(1 - d_\alpha)\gamma_2$, while a small price cut by Firm 2 raises its demand on Firm 1's atom by $(d_\alpha - \mu)\gamma_1$. Both gains are nonnegative, and at least one is positive because $(1 - d_\alpha) + (d_\alpha - \mu) = 1 - \mu > 0$. If $p_1 - p_2 = \beta$, the corresponding gains are $(\mu - d_\beta)\gamma_2$ and $d_\beta\gamma_1$; at least one is positive because their coefficients sum to $\mu > 0$. At all other rival prices, a price cut weakly increases the deviating firm's demand. Since $p_1, p_2 > 0$ and the revenue loss on existing sales can be made arbitrarily small by choosing a sufficiently small price cut, at least one firm has a profitable deviation.

Positive profits imply $\alpha \le A - B \le \beta$. If $A - B > \beta$, pure best responses sufficiently close to A give Firm 1 zero demand, contradicting $\Pi_1 > 0$. If $A - B < \alpha$, pure best responses sufficiently close to B give Firm 2 zero demand, contradicting $\Pi_2 > 0$. If $A - B = \beta$, Firm 1 cannot have an atom at A. Any price played with positive probability must yield the positive equilibrium profit, and hence positive demand. But at A, Firm 1 has positive demand only when Firm 2 charges B, so Firm 2 would also have an atom at B, contradicting the preceding argument. Similarly, if $A - B = \alpha$, Firm 2 cannot have an atom at B.

Choose pure best responses $p_2^n \to B$, with $p_2^n \le B$. Firm 2's expected demand at p_2^n is Π_2/p_2^n. Its demand is zero when $p_1 < p_2^n + \alpha$, at most $1 - \mu$ when $p_2^n + \alpha \le p_1 < p_2^n + \beta$, and at most one otherwise. Hence

$$\frac{\Pi_2}{p_2^n} \le (1 - \mu)\Pr(p_1 \ge p_2^n + \alpha) + \mu \Pr(p_1 \ge p_2^n + \beta).$$

We have $\Pr(p_1 \ge p_2^n + \beta) \to 0$: if $A - B < \beta$, the event is empty for all large n; if $A - B = \beta$, the limit is zero because Firm 1 has no atom at A. Since $p_2^n \le B$ and $p_2^n \to B$, we also have $\Pr(p_1 \ge p_2^n + \alpha) \to \Pr(p_1 \ge B + \alpha)$. Taking limits therefore gives $\Pi_2 \le (1 - \mu)B \Pr(p_1 \ge B + \alpha)$, thus $\Pr(p_1 \ge B + \alpha) > 0$ because $\Pi_2 > 0$.

If $B > L$, Firm 2 can charge $B - L - \varepsilon > 0$ for sufficiently small $\varepsilon > 0$. Whenever $p_1 \ge B + \alpha$, the price gap then exceeds β, and Firm 2 wins the whole market. The deviation therefore yields profit at least $(B - L - \varepsilon)\Pr(p_1 \ge B + \alpha)$, which cannot exceed Π_2. Combining this with $\Pi_2 \le (1 - \mu)B \Pr(p_1 \ge B + \alpha)$ and dividing by $\Pr(p_1 \ge B + \alpha) > 0$ gives $B - L - \varepsilon \le (1 - \mu)B$. Letting $\varepsilon \downarrow 0$ gives $B \le L/\mu$. If $B \le L$, the same bound holds immediately. The same argument for Firm 1 gives $A \le L/(1 - \mu)$. These bounds show that the price cap does not

bind.[39]

Step 2: No atoms. Suppose Firm 1 has an atom of size $\gamma > 0$ at a price p_1, and let $\hat{D}_2(p_2)$ denote Firm 2's expected demand at price p_2. For sufficiently small $h > 0$, Firm 2's demand when Firm 1 charges p_1 is one at $p_1 - \beta - h$ and $1 - \mu$ at $p_1 - \beta + h$. At all other Firm 1 prices, raising Firm 2's price weakly reduces its demand. Hence $\hat{D}_2(p_1 - \beta + h) \le \hat{D}_2(p_1 - \beta - h) - \mu\gamma$.

If $p_1 - \beta < \Pi_2$, there exists $\varepsilon > 0$ such that $(p_1 - \beta, p_1 - \beta + \varepsilon) \cap S_2 = \varnothing$, since every support price is at least Π_2. If $p_1 - \beta \ge \Pi_2$, then for all sufficiently small $h > 0$, no profitable deviation to $p_1 - \beta - h$ implies $\hat{D}_2(p_1 - \beta - h) \le \Pi_2/(p_1 - \beta - h)$. Hence

$$\pi_2(p_1 - \beta + h) \le (p_1 - \beta + h)\left(\frac{\Pi_2}{p_1 - \beta - h} - \mu\gamma\right) < \Pi_2.$$

The strict inequality holds for sufficiently small h because the middle expression converges to $\Pi_2 - (p_1 - \beta)\mu\gamma < \Pi_2$ as $h \downarrow 0$. Thus, in either case, there exists $\varepsilon > 0$ such that $(p_1 - \beta, p_1 - \beta + \varepsilon) \cap S_2 = \varnothing$. The same argument gives $\varepsilon > 0$ such that $(p_1 - \alpha, p_1 - \alpha + \varepsilon) \cap S_2 = \varnothing$.

By Step 1, Firm 2 has no atom at either $p_1 - \beta$ or $p_1 - \alpha$. Since the price cap does not bind, we can choose $\varepsilon > 0$ small enough that $p_1 + \varepsilon < \bar{p}$, $F_2(p_1 + \varepsilon - \beta) = F_2(p_1 - \beta)$, and $F_2(p_1 + \varepsilon - \alpha) = F_2(p_1 - \alpha)$. Firm 1 then faces the same demand at $p_1 + \varepsilon$ as at p_1 but charges a higher price. Since p_1 is played with positive probability and yields positive profit, this is a profitable deviation. Hence Firm 1 has no atoms. The same argument applies to Firm 2.

Because both price distributions are atomless, the demand cutoffs $\Delta = \alpha$ and $\Delta = \beta$ occur with probability zero. Hence profits are

$$\begin{aligned} \pi_1(p_1) &= p_1\big(1 - (1 - \mu)F_2(p_1 - \alpha) - \mu F_2(p_1 - \beta)\big), \\ \pi_2(p_2) &= p_2\big(1 - \mu F_1(p_2 + \beta) - (1 - \mu)F_1(p_2 + \alpha)\big). \end{aligned} \tag{A.1}$$

These profit functions are continuous. Since every support point is a limit of pure best responses, continuity implies that every support point itself yields the equilibrium profit.

Step 3: Endpoint inequalities. Every $p_1 \in S_1$ satisfies $p_1 - \alpha \in S_2$ or $p_1 - \beta \in S_2$. Otherwise both CDF terms in (A.1) would be locally constant, and Firm 1 could raise its price slightly without changing demand. Similarly, every $p_2 \in S_2$ satisfies $p_2 + \alpha \in S_1$ or $p_2 + \beta \in S_1$. At a and b, these conditions and $\beta > \alpha$ imply $a - \alpha \ge b$ and $b + \beta \ge a$, respectively, hence $\alpha \le a - b \le \beta$. At the upper endpoints, if $A - B \ge \beta$, Firm 1 has zero demand at A, and if $A - B \le \alpha$, Firm 2 has zero demand at B. This includes equality because neither distribution has an atom at

[39] Recall that $\alpha > -1$, $\beta < s$, and $L < 1 + s$. If $\mu \ge 1/2$, then $B \le 2L$ and $A \le B + \beta \le 2L + \beta$, so $a + L \le 3L + \beta < 3 + 4s$ and $b + L \le 3L < 3 + 4s$. If $\mu < 1/2$, then $A \le 2L$ and $B \le A - \alpha \le 2L - \alpha$, so $a + L \le 3L < 3 + 4s$ and $b + L \le 3L - \alpha < 3(1 + s) + 1 < 3 + 4s$. The last inequality uses $\mu \ge 1/(1 + s)$, which implies $s > 1$. Thus, in either case, $A, B, a + L, b + L < \bar{p}$. Hence the price cap does not bind, and all upward deviations considered below are feasible.

its upper endpoint. Hence

$$\alpha < A - B < \beta. \tag{A.2}$$

Using the support conditions above, $B + \beta > A$ implies $B + \alpha \in S_1$, while $A - \alpha > B$ implies $A - \beta \in S_2$.

For the remainder of this step and Step 4, suppose that $\mu \geq 1/2$. If $\mu < 1/2$, interchange the firms and use $-\Delta$ as the price gap. The relabeled demand has the same form, with demand $1 - \mu > 1/2$ between the cutoffs $-\beta$ and $-\alpha$, and the corresponding admissible demands at the cutoffs. Applying the same argument and restoring the original labels gives the same conclusions.

Define $\phi_1(t)$ for $t > \beta$ and $\phi_2(t)$ for $t > -\alpha$ by

$$\phi_1(t) := \frac{1 - \Pi_2/(t-\beta)}{\mu} - \left(1 - \frac{\Pi_2}{(1-\mu)(t-\alpha)}\right),$$
$$\phi_2(t) := \frac{1 - \Pi_1/(t+\alpha)}{1-\mu} - \left(1 - \frac{\Pi_1}{\mu(t+\beta)}\right).$$

Using $r = t - \beta$ for ϕ_1 and $r = t + \alpha$ for ϕ_2, we obtain

$$\mu(1-\mu)r(r+L)\phi_1(t) = (1-\mu)^2 r(r+L) + (2\mu - 1)\Pi_2 r - (1-\mu)\Pi_2 L,$$
$$\mu(1-\mu)r(r+L)\phi_2(t) = \mu^2 r(r+L) - (2\mu - 1)\Pi_1 r - \mu\Pi_1 L.$$

Each quadratic has a positive leading coefficient and a negative constant term, so it has exactly one positive root. Since $\mu(1-\mu)r(r+L) > 0$ for $r > 0$, each ϕ_i changes sign exactly once, from negative to positive. Consequently,

$$\phi_i(t') \geq 0, \quad \phi_i(t) \leq 0 \quad \Longrightarrow \quad t \leq t'. \tag{A.3}$$

Suppose that $F_2(t - L) = 0$. At $t + \alpha$, (A.1) then gives Firm 1 demand $1 - (1-\mu)F_2(t)$. Hence Firm 1's no-deviation condition at $t + \alpha$ implies $F_2(t) \geq [1 - \Pi_1/(t+\alpha)]/(1-\mu)$, with equality if $t + \alpha \in S_1$. At $t + \beta$, (A.1) and $F_2(t + L) \leq 1$ imply that Firm 1's demand is at least $\mu[1 - F_2(t)]$. Hence its no-deviation condition at $t + \beta$ implies $F_2(t) \geq 1 - \Pi_1/[\mu(t+\beta)]$. Equality holds if $t + \beta \in S_1$ and $F_2(t + L) = 1$. Therefore, if $t + \alpha \in S_1$, then $\phi_2(t) \geq 0$, whereas if $t + \beta \in S_1$ and $F_2(t + L) = 1$, then $\phi_2(t) \leq 0$. Similarly, if $F_1(t - L) = 0$, then $t - \beta \in S_2$ implies $\phi_1(t) \geq 0$, while $t - \alpha \in S_2$ and $F_1(t + L) = 1$ imply $\phi_1(t) \leq 0$.

Step 4: Support lengths. Let $W_1 := A - a$ and $W_2 := B - b$. By Step 1, $B \leq L/\mu \leq 2L$, so $W_2 < 2L$ because $b > 0$. We first show that $W_1, W_2 \geq L$. If $W_2 < L$, then $B - L < b$ and $b + L > B$, so (A.1) gives Firm 1 demand μ at both $B + \alpha$ and $b + \beta$. Since $B + \alpha \in S_1$ and $B + \alpha < b + \beta$, a deviation to $b + \beta$ is profitable. Hence $W_2 \geq L$. The same argument for Firm 2, using the support price $A - \beta$ and the higher price $a - \alpha > A - \beta$ when $W_1 < L$, gives

$W_1 \geq L$.

Since $a-\beta \leq b$ and $b+\alpha \leq a$, atomlessness implies $F_2(a-\beta) = F_1(b+\alpha) = 0$. Using (A.1) at a and b and the no-deviation conditions at $a+L$ and $b+L$ gives $F_2(a-\alpha) = [1-\Pi_1/a]/(1-\mu) \geq 1-\Pi_1/[\mu(a+L)]$ and $F_1(b+\beta) = [1-\Pi_2/b]/\mu \geq 1-\Pi_2/[(1-\mu)(b+L)]$. Hence $\phi_2(a-\alpha) \geq 0$ and $\phi_1(b+\beta) \geq 0$.

We next show that $B-L \leq a-\alpha$. Suppose instead that $B-L > a-\alpha$. Then $B-L+\alpha > a > 0$, and $W_2 < 2L$ implies $B-2L < b$, so $F_2(B-2L) = 0$. Also, $B+\alpha \in S_1$ and $F_2(B) = 1$. Thus the sign implication in Step 3, applied at $t = B-L$, gives $\phi_2(B-L) \leq 0$. Together with $\phi_2(a-\alpha) \geq 0$, (A.3) gives $B-L \leq a-\alpha$, a contradiction. Hence $B-L \leq a-\alpha$.

We now show that $W_2 = L$. Since $W_2 \geq L$, we have $B-L \geq b > 0$. $B-L \leq a-\alpha$ implies $B+\alpha-L \leq a$, so $F_1(B+\alpha-L) = 0$. Also, $B \in S_2$, while (A.2) implies $B+\beta > A$ and hence $F_1(B+\beta) = 1$. Thus the sign implication in Step 3 applied at $t = B+\alpha$ gives $\phi_1(B+\alpha) \leq 0$. Together with $\phi_1(b+\beta) \geq 0$, (A.3) gives $B+\alpha \leq b+\beta$, or $W_2 \leq L$. Therefore $W_2 = L$.

Since $B = b+L$, (A.2) implies $A-\beta-L < b$ and $A-\alpha > B$, so $F_2(A-\beta-L) = 0$ and $F_2(A-\alpha) = 1$. Also, $A \in S_1$ and $A-L \geq a > 0$ because $W_1 \geq L$. Thus the sign implication in Step 3 applied at $t = A-\beta$ gives $\phi_2(A-\beta) \leq 0$. Together with $\phi_2(a-\alpha) \geq 0$, (A.3) gives $A-\beta \leq a-\alpha$, or $W_1 \leq L$. Therefore $W_1 = L$.

Since $W_1 = W_2 = L$, we have $a-b = A-B$. Together with (A.2), this establishes

$$\alpha < a-b < \beta, \quad \alpha < A-B < \beta, \quad A-a = B-b = L. \tag{A.4}$$

The relabeling argument in Step 3 shows that (A.4) also holds when $\mu < 1/2$.

Step 5: Connected supports. Suppose that S_1 has a gap $(x,y) \subset (a, a+L)$ with $x, y \in S_1$. If $p_2 \in (x-\alpha, y-\alpha)$, then $p_2+\alpha \in (x,y)$ and $p_2+\beta > x+L > a+L = A$. From the support argument in Step 3, any $p_2 \in S_2$ must satisfy $p_2+\alpha \in S_1$ or $p_2+\beta \in S_1$. Hence $p_2 \notin S_2$. Likewise, if $p_2 \in (x-\beta, y-\beta)$, then $p_2+\beta \in (x,y)$ and $p_2+\alpha < y-L < a$, so again $p_2 \notin S_2$. Since F_2 is atomless, $F_2(x-\alpha) = F_2(y-\alpha)$ and $F_2(x-\beta) = F_2(y-\beta)$. Firm 1 therefore faces the same positive demand at the two support prices $x < y$, which contradicts indifference. Thus $S_1 = [a, a+L]$. The same argument gives $S_2 = [b, b+L]$.

Step 6: Equilibrium CDFs and endpoints. By (A.4), the prices $b+\beta = B+\alpha$ and $a-\alpha = A-\beta$ lie strictly inside Firm 1's and Firm 2's supports, respectively. For $p_1 \in [a, b+\beta]$, we have $F_2(p_1-\beta) = 0$, and Firm 1's indifference condition is $\Pi_1 = p_1\big(1-(1-\mu)F_2(p_1-\alpha)\big)$. Setting $p_2 = p_1-\alpha$ gives

$$F_2(p_2) = \frac{1-\Pi_1/(p_2+\alpha)}{1-\mu} \quad \text{for } p_2 \in [a-\alpha, b+L].$$

For $p_1 \in [b+\beta, a+L]$, we have $F_2(p_1-\alpha) = 1$, and $\Pi_1 = \mu p_1\big(1-F_2(p_1-\beta)\big)$. Setting $p_2 = p_1-\beta$ gives

$$F_2(p_2) = 1 - \frac{\Pi_1}{\mu(p_2+\beta)} \quad \text{for } p_2 \in [b, a-\alpha].$$

Using $F_2(b) = 0$ gives $\Pi_1 = \mu(b+\beta)$, and substitution gives the stated CDF F_2. The analogous calculation for Firm 2 gives $\Pi_2 = (1-\mu)(a-\alpha)$ and the stated CDF F_1. Continuity of F_2 at $a-\alpha$ and of F_1 at $b+\beta$ give (3)-(4). By Lemma A.1, these endpoint equations have a unique interior solution whenever $k < \hat{k}$.

Step 7: Best responses. The endpoint equations make the two pieces of each CDF meet continuously, and the formulas are strictly increasing from zero to one on their respective supports. Thus they define valid atomless CDFs. By construction, every support price gives the stated equilibrium profit.

It remains to rule out profitable deviations outside the supports. Consider first Firm 1 charging a price $p_1 \in [0, a)$. If $p_1 < b+\alpha$, Firm 1 sells for sure, and $\pi_1(p_1) = p_1$. If $p_1 \geq b+\alpha$, then $p_1 - \alpha \in [b, a-\alpha)$, while $p_1 - \beta < b$ because $p_1 < a$ and $a - \beta < b$. Hence $\pi_1(p_1) = p_1\big(\mu + (1-\mu)(b+\beta)/(p_1+L)\big)$. The two profit expressions coincide at $p_1 = b+\alpha$, and the second expression is strictly increasing since its derivative is $\mu + (1-\mu)(b+\beta)L/(p_1+L)^2 > 0$. Thus every price below a yields less than Π_1. For $p_1 > a+L$, Firm 1's profit is zero when $p_1 \geq b+\beta+L$ and otherwise equals $\frac{\mu^2}{1-\mu} p_1(b+\beta+L-p_1)/(p_1-L)$. The latter is strictly decreasing in p_1, since its derivative has the sign of $-[(p_1-L)^2 + L(b+\beta)] < 0$. Hence Firm 1 has no profitable deviation outside $[a, a+L]$. The same argument for Firm 2 shows that it has no profitable deviation outside $[b, b+L]$.

It remains to verify that the candidate supports lie below the price cap. By Lemma A.1, $\tilde{y} - \tilde{x} < \mu/(1-\mu)$ and $\tilde{b} < (1-\mu)/\mu$. Since $A = L(\tilde{y}-\tilde{x}+1)$ and $B = L(\tilde{b}+1)$, it follows that $A < L/(1-\mu)$ and $B < L/\mu$. The same bounds as in Step 1 then imply $A, B < 3 + 4s < \bar{p}$. Hence the candidate distributions are feasible.

Steps 1-6 show that any equilibrium must be atomless, have supports $[a, a+L]$ and $[b, b+L]$, and have the CDFs derived above, where (a, b) solves (3)-(4). By Lemma A.1, these endpoint equations have a unique interior solution. Step 7 verifies that the resulting distributions are mutual best responses. Therefore the equilibrium in price distributions is unique.

Finally, because both price distributions are atomless, the events $p_1 - p_2 = \alpha$ and $p_1 - p_2 = \beta$ occur with probability zero under the rival's mixed strategy. Hence the equilibrium price distributions, profits, and outcomes do not depend on the admissible tie-breaking rule. □

Proof of Lemma 2. Write $a = a(k)$ and $b = b(k)$, and let primes denote derivatives with respect to k. We first show that $a' < 0$ and $b' < 0$.

Since $b = L\tilde{b}$, Lemma A.1 gives

$$b' = -(1+s)^2\tilde{b} + L\frac{d\tilde{b}}{d\tilde{x}}\frac{d\tilde{x}}{dk} < 0,$$

because $\tilde{b} > 0$, $d\tilde{b}/d\tilde{x} < 0$, and $d\tilde{x}/dk \geq 0$.

The first endpoint equation can be written as $b = \mu a(a+L)/[(2\mu-1)a+\mu L] - \beta$. Its denominator is positive because the endpoint equation implies $b+\beta = L\tilde{y} > 0$. Differentiating gives

$$b' = \frac{\mu\big((2\mu-1)a^2 + 2\mu aL + \mu L^2\big)}{\big((2\mu-1)a+\mu L\big)^2}a' + (1-\mu)(1+s)^2\left(1 + \frac{\mu a^2}{\big[(2\mu-1)a+\mu L\big]^2}\right).$$

The coefficient of a' is strictly larger than one, since its difference from one is

$$\frac{(1-\mu)a\big((2\mu-1)a+2\mu L\big)}{\big((2\mu-1)a+\mu L\big)^2} > 0.$$

The second term in the expression for b' is strictly positive. Since $b' < 0$, it follows that $a' < 0$.

Also, $\alpha' = \mu(1+s)^2 > 0$, $\beta' = -(1-\mu)(1+s)^2 < 0$, and $L' = -(1+s)^2 < 0$. Hence each of the six prices a, $b+\beta$, $a+L$, b, $a-\alpha$, and $b+L$ is strictly decreasing in k.

Fix p. On each nonconstant part of the equilibrium CDFs, differentiation gives

$$\begin{aligned}
&\frac{\partial F_1(p;k)}{\partial k} = \frac{-a'(p-\alpha)+\alpha'(p-a)}{(p-\alpha)^2} > 0, && a < p < b+\beta,\\
&\frac{\partial F_1(p;k)}{\partial k} = \frac{1-\mu}{\mu}\frac{-(a'-\alpha')(p-\beta)-(a-\alpha)\beta'}{(p-\beta)^2} > 0, && b+\beta < p < a+L,\\
&\frac{\partial F_2(p;k)}{\partial k} = \frac{-b'(p+\beta)-\beta'(p-b)}{(p+\beta)^2} > 0, && b < p < a-\alpha,\\
&\frac{\partial F_2(p;k)}{\partial k} = \frac{\mu}{1-\mu}\frac{-(b'+\beta')(p+\alpha)+(b+\beta)\alpha'}{(p+\alpha)^2} > 0, && a-\alpha < p < b+L.
\end{aligned}$$

The signs follow from $a' < 0$, $b' < 0$, $\alpha' > 0$, $\beta' < 0$, and the inequalities defining the corresponding CDF parts.

For a fixed p, away from the values of k at which p equals one of the six decreasing prices above, $F_i(p;k)$ is either constant or strictly increasing in k. Each of the six prices crosses a fixed p at most once, and at every crossing the adjacent CDF formulas agree. Therefore $F_i(p;k)$ is weakly increasing in k for every fixed p and $i \in \{1,2\}$. This proves the result. □

Proof of Lemma 3. Let (a_0, b_0) denote the lower endpoints of the equilibrium supports at $k = 0$. When $\bar{w} = 0$, substituting $\alpha = -\lambda$, $\beta = s\lambda$, and $L = (1+s)\lambda$ into the endpoint equations shows that $(\lambda a_0, \lambda b_0)$ is a solution for every $k < 1/(1+s)$. Since the endpoint solution is

unique, $a = \lambda a_0$ and $b = \lambda b_0$.

Substituting these expressions into the equilibrium CDFs gives $F_i(p;k) = F_i(p/\lambda;0)$. The profit formulas $\Pi_1 = \mu(b+\beta)$ and $\Pi_2 = (1-\mu)(a-\alpha)$ similarly give $\Pi_i(k) = \lambda\Pi_i(0)$. Since the firms randomize independently, the induced price gap satisfies $\Delta(k) \overset{d}{=} \lambda\Delta(0)$. □

Lemma A.2. *For $\bar{w} \geq 0$, the equilibrium support of $\Delta(k)$ shrinks strictly with k:* $\operatorname{supp}(\Delta(k_2)) \subset \operatorname{supp}(\Delta(k_1))$ *for $k_2 > k_1$.*

Proof. The price supports $[a, a+L]$ and $[b, b+L]$ imply that the support of Δ is $[a-b-L, a-b+L]$. Since $\tilde{x} = (b+\beta-a)/L$, we have $a-b = \beta - L\tilde{x}$. Using $L = \beta - \alpha$, its lower and upper endpoints can be written as $\alpha - L\tilde{x}$ and $\beta + L(1-\tilde{x})$.

By Lemma A.1, $d\tilde{x}/dk \geq 0$ when $\bar{w} \geq 0$. Since $\beta' = -(1-\mu)(1+s)^2$ and $L' = -(1+s)^2$,

$$\frac{d}{dk}\big[\beta + L(1-\tilde{x})\big] = -(1+s)^2(2-\mu-\tilde{x}) - L\frac{d\tilde{x}}{dk} < 0,$$

where the inequality follows from $\tilde{x} < 1$ and $\mu < 1$. Thus the upper endpoint is strictly decreasing.

For the lower endpoint, let primes on $\tilde{y}$ and $\tilde{b}$ denote derivatives with respect to $\tilde{x}$. Differentiating $\tilde{y} - \tilde{b} = \beta/L$ with respect to k gives $(\tilde{y}' - \tilde{b}')d\tilde{x}/dk = (1+s)^2\bar{w}/L^2$. Moreover, $\beta - (1-\mu)L = \bar{w}$, so $\bar{w}/L = \tilde{y} - \tilde{b} + \mu - 1$. Hence

$$\frac{d}{dk}\big[\alpha - L\tilde{x}\big] = (1+s)^2\left[\tilde{x} + \mu - \frac{\tilde{y} - \tilde{b} + \mu - 1}{\tilde{y}' - \tilde{b}'}\right].$$

It remains to show that the term in brackets is positive. Differentiating (9) with respect to $\tilde{x}$ gives $[2(1-\mu)\tilde{y} - \tilde{x}]\tilde{y}' = \tilde{y} + \mu(1-2\tilde{x})$. Multiplying by $\tilde{x}+\mu$ and using (9) yields

$$\big(2(1-\mu)\tilde{y} - \tilde{x}\big)\big((\tilde{x}+\mu)\tilde{y}' - \tilde{y}\big) = \mu\big(\tilde{y} - (2\mu+1)\tilde{x} + \mu\big).$$

Since $\tilde{y} > \tilde{x}/(1-\mu)$,

$$\tilde{y} - (2\mu+1)\tilde{x} + \mu > \frac{\mu\big((2\mu-1)\tilde{x} + 1 - \mu\big)}{1-\mu} > 0.$$

Also, $2(1-\mu)\tilde{y} - \tilde{x} > 0$, so $(\tilde{x}+\mu)\tilde{y}' > \tilde{y}$. Since $\tilde{b}' < 0$, we also have $(\tilde{x}+\mu)\tilde{b}' < \tilde{b} + 1 - \mu$. Therefore $(\tilde{x}+\mu)(\tilde{y}' - \tilde{b}') > \tilde{y} - \tilde{b} + \mu - 1$. Because $\tilde{y}' - \tilde{b}' > 0$, the term in brackets above is positive. Thus the lower endpoint is strictly increasing. □

Proof of Lemma 4. The equilibrium CDFs give

$$m_1 = \frac{\tilde{x}}{1+\tilde{b}}, \qquad m_2 = \frac{1-\tilde{x}}{1+\tilde{y}-\tilde{x}}.$$

By Lemma A.1, $\tilde{x}$ is strictly increasing in k, $\tilde{b}$ is strictly decreasing in $\tilde{x}$, and $\tilde{y}-\tilde{x}$ is strictly increasing in $\tilde{x}$. Hence the numerator of m_1 increases and its denominator decreases. Thus m_1 is strictly increasing. For m_2, the numerator decreases while the denominator $1+\tilde{y}-\tilde{x}$ increases, so m_2 is strictly decreasing.

Finally, as $k \uparrow \hat{k}$, Lemma A.1 gives $\tilde{x} \to 1$, $\tilde{b} \to 0$, and $\tilde{y} \to 1/(1-\mu)$. Substituting these limits yields $m_1 \to 1$ and $m_2 \to 0$. □

Proof of Lemma 5. Use the normalized variables in Lemma A.1, and let primes denote derivatives with respect to $\tilde{x}$. Since $\bar{w} > 0$, that lemma implies that $\tilde{x}$ is strictly increasing in $k \in [0, \hat{k})$, with $\tilde{y}' > 1$ and $\tilde{b}' < 0$. Differentiating (9)-(10) gives

$$\tilde{y} - \tilde{x}\tilde{y}' = \frac{\mu\tilde{x}}{2(1-\mu)\tilde{y}-\tilde{x}} > 0, \tag{11}$$

$$\tilde{b} + (1-\tilde{x})\tilde{b}' = \frac{(1-\mu)(1-\tilde{x})}{2\mu\tilde{b} + (2\mu-1)(1-\tilde{x})} > 0. \tag{12}$$

It is therefore enough to compare the two tail probabilities as $\tilde{x}$ increases.

For each firm, let $p_i(q)$ satisfy $F_i(p_i(q)) = q$ for $q \in (0,1)$. Since F_i is atomless, $F_i(p_i)$ is uniformly distributed on $[0,1]$.

First consider $\Pr(\Delta < \alpha)$. If $q > m_1$, then $p_1(q) > b + \beta$, so $\Delta < \alpha$ is impossible. For $q \le m_1$, conditional on $p_1(q)$, the event occurs when $p_2 > p_1(q) - \alpha$. Since the firms randomize independently, this has probability $1 - F_2(p_1(q) - \alpha)$. Hence $\Pr(\Delta < \alpha) = \int_0^{m_1} [1 - F_2(p_1(q) - \alpha)]\, dq$.

For $q \le m_1$, the lower branch of F_1 gives $p_1(q)/L = \tilde{y} - \tilde{b} - 1 + (\tilde{b} + 1 - \tilde{x})/(1-q)$. Since $p_1(q) - \alpha \in [a - \alpha, b + L]$, the upper branch of F_2 gives

$$1 - F_2(p_1(q) - \alpha) = \frac{\mu}{1-\mu}\big(\Phi_A(q,\tilde{x}) - 1\big), \quad \Phi_A(q,\tilde{x}) := \frac{\tilde{y}(1-q)}{(1-q)\tilde{y} + q(1+\tilde{b}) - \tilde{x}}. \tag{13}$$

Holding q fixed and differentiating with respect to $\tilde{x}$ gives

$$\frac{\partial \Phi_A}{\partial \tilde{x}} = \frac{(1-q)\left[\tilde{y} - \tilde{x}\tilde{y}' + q\big((1+\tilde{b})\tilde{y}' - \tilde{y}\tilde{b}'\big)\right]}{\big((1-q)\tilde{y} + q(1+\tilde{b}) - \tilde{x}\big)^2} > 0.$$

The sign follows from (11), $\tilde{y}' > 0$, and $\tilde{b}' < 0$. Thus the conditional probability in (13) strictly increases with $\tilde{x}$ for each fixed $q < m_1$. Lemma 4 also shows that m_1 increases. Hence the integral over $[0, m_1]$ strictly increases and thus $\Pr(\Delta < \alpha)$ is strictly increasing in $\tilde{x}$ and therefore in k.

We next consider $\Pr(\Delta > \beta)$, conditioning on Firm 2's price. If $q > m_2$, then $p_2(q) > a - \alpha$, and $\Delta > \beta$ is impossible. For $q \le m_2$, conditional on $p_2(q)$, the event occurs when $p_1 > p_2(q) + \beta$, which has probability $1 - F_1(p_2(q) + \beta)$. Thus $\Pr(\Delta > \beta) = \int_0^{m_2} [1 - F_1(p_2(q) + \beta)]\, dq$.

The lower branch of F_2 gives $p_2(q)/L = \tilde{b} + q\tilde{y}/(1-q)$. For $q \le m_2$, we have $p_2(q) + \beta \in [b+\beta, a+L]$, so the upper branch of F_1 gives

$$1 - F_1(p_2(q)+\beta) = \frac{1-\mu}{\mu}\big(\Phi_R(q,\tilde{x}) - 1\big), \quad \Phi_R(q,\tilde{x}) := \frac{(\tilde{b}+1-\tilde{x})(1-q)}{(1-q)\tilde{b} + q\tilde{y}}. \tag{14}$$

Holding q fixed and differentiating with respect to $\tilde{x}$ gives

$$\frac{\partial \Phi_R}{\partial \tilde{x}} = -\frac{(1-q)^2\big(\tilde{b} + (1-\tilde{x})\tilde{b}'\big)}{\big((1-q)\tilde{b} + q\tilde{y}\big)^2} - \frac{q(1-q)\big(\tilde{y}(1-\tilde{b}') + (\tilde{b}+1-\tilde{x})\tilde{y}'\big)}{\big((1-q)\tilde{b} + q\tilde{y}\big)^2} < 0.$$

The first term is strictly negative by (12). The second is nonpositive because $\tilde{y}' > 0$, $\tilde{b}' < 0$, and $\tilde{b} + 1 - \tilde{x} > 0$. Thus the conditional probability in (14) strictly decreases with $\tilde{x}$ for each fixed $q < m_2$. Lemma 4 also shows that m_2 decreases. Hence the integral over $[0, m_2]$ strictly decreases, so $\Pr(\Delta > \beta)$ is strictly decreasing in $\tilde{x}$ and therefore in k.

Finally, as $k \uparrow \hat{k}$, Lemma 4 gives $m_1 \to 1$ and $m_2 \to 0$. Since $\Delta > \beta$ requires $p_2 \le a - \alpha$, we have $0 \le \Pr(\Delta > \beta) \le m_2 \to 0$.

For the left tail, Lemma A.1 gives $\tilde{x} \to 1$, $\tilde{b} \to 0$, and $\tilde{y} \to 1/(1-\mu)$. For every fixed $q \in (0,1)$, we have $q < m_1$ when k is sufficiently close to $\hat{k}$. Equation (13) then gives $\Phi_A(q,\tilde{x}) \to 1/\mu$, and hence $1 - F_2(p_1(q) - \alpha) \to 1$. This probability is zero for $q > m_1$ and always lies between zero and one. By the dominated convergence theorem,

$$\Pr(\Delta < \alpha) = \int_0^1 [1 - F_2(p_1(q) - \alpha)]\, dq \longrightarrow 1.$$

□

A.2 Welfare Analysis

Proof of Proposition 2. Note that $\operatorname{supp}(\Delta) = [a-b-L, a-b+L]$. The equilibrium conditions $b < a - \alpha$ and $a < b + \beta$ imply $\alpha < a - b < \beta$. Since $L = \beta - \alpha$, it follows that $a - b - L < \alpha < \beta < a - b + L$. Hence $[\alpha, \beta] \subset \operatorname{int}(\operatorname{supp}(\Delta))$. □

Lemma A.3. *Fix $\mu \in (0,1)$. For each $\tilde{\beta} \in (1-\mu, 1)$, let $\tilde{x}(\tilde{\beta}) \in (0,1)$ denote the unique solution to $\tilde{y}(\tilde{x}) - \tilde{b}(\tilde{x}) = \tilde{\beta}$ characterized in Lemma A.1. Under the normalized equilibrium price distributions corresponding to $\tilde{x}(\tilde{\beta})$, define $A(\tilde{\beta}) := \Pr(\Delta/L < \tilde{\beta} - 1)$ and $R(\tilde{\beta}) := \Pr(\Delta/L > \tilde{\beta})$. Then*

$$(1-\mu)A(\tilde{\beta}) + \mu\frac{\tilde{\beta}}{1-\tilde{\beta}}R(\tilde{\beta})$$

has a strictly positive derivative with respect to $\tilde{\beta}$ on $(1-\mu, 1)$.

Proof. Fix μ, and regard $\tilde{b}$, $\tilde{y}$, and $\tilde{\beta} = \tilde{y} - \tilde{b}$ as functions of $\tilde{x}$. Throughout the proof, primes denote derivatives with respect to $\tilde{x}$. By Lemma A.1, $\tilde{\beta}' = \tilde{y}' - \tilde{b}' > 0$. It is therefore enough to show that the expression in the statement has a strictly positive derivative with respect to $\tilde{x}$.

Differentiating the normalized endpoint equations gives $\tilde{y}' = [\tilde{y} + \mu(1-2\tilde{x})]/[2(1-\mu)\tilde{y} - \tilde{x}]$ and $\tilde{b}' = [(2\mu-1)\tilde{b} - (1-\mu)]/[2\mu\tilde{b} + (2\mu-1)(1-\tilde{x})]$. Both denominators are positive at the positive endpoint solutions.

Define $D_A(\tilde{x}, z) := (1-z)(1+\tilde{b})(\tilde{y}-\tilde{x}) + z\tilde{y}(\tilde{b}+1-\tilde{x})$ and $D_R(\tilde{x}, z) := (1-z)\tilde{b}(1+\tilde{y}-\tilde{x}) + z\tilde{y}(\tilde{b}+1-\tilde{x})$. Conditioning on Firm 1's price for A and Firm 2's price for R, and changing variables so that each integral is over $z \in [0,1]$, the equilibrium CDFs give

$$A(\tilde{\beta}) = \frac{\mu}{1-\mu}\int_0^1 \frac{\tilde{x}^2(1-z)}{D_A(\tilde{x}, z)}\,dz, \qquad R(\tilde{\beta}) = \frac{1-\mu}{\mu}\int_0^1 \frac{(1-\tilde{x})^2(1-z)}{D_R(\tilde{x}, z)}\,dz,$$

$$(1-\mu)A(\tilde{\beta}) + \mu\frac{\tilde{\beta}}{1-\tilde{\beta}}R(\tilde{\beta}) = \int_0^1 (1-z)\Psi(\tilde{x}, z)\,dz,$$

where $\Psi(\tilde{x}, z) := \mu\tilde{x}^2/D_A(\tilde{x}, z) + (1-\mu)\tilde{\beta}(1-\tilde{x})^2/[(1-\tilde{\beta})D_R(\tilde{x}, z)]$. Since $\tilde{y} > \tilde{x}$, $\tilde{b} > 0$, and $0 < \tilde{x} < 1$, both $D_A(\tilde{x}, z)$ and $D_R(\tilde{x}, z)$ are strictly positive for every $z \in [0,1]$.

Differentiating Ψ and substituting the expressions for $\tilde{y}'$ and $\tilde{b}'$, we can write

$$\frac{\partial\Psi(\tilde{x}, z)}{\partial\tilde{x}} = \frac{N_0(1-z)^3 + 3N_1 z(1-z)^2 + 3N_2 z^2(1-z) + N_3 z^3}{\mathcal{D}}, \tag{15}$$

where the coefficients $N_0, \ldots, N_3$ do not depend on z, and

$$\mathcal{D} := \big[2(1-\mu)\tilde{y} - \tilde{x}\big]\big[2\mu\tilde{b} + (2\mu-1)(1-\tilde{x})\big](1-\tilde{\beta})^2 D_A(\tilde{x}, z)^2 D_R(\tilde{x}, z)^2 > 0.$$

The first two factors are the positive denominators in $\tilde{y}'$ and $\tilde{b}'$. The remaining factors are positive because $\tilde{\beta} < 1$ and $D_A, D_R > 0$.

We use Mathematica to check the signs of the four coefficients. Under the normalized endpoint equations (9)-(10), $0 < \mu < 1$, $0 < \tilde{x} < 1$, $\tilde{b} > 0$, $\tilde{y} > 0$, and $1 - \mu < \tilde{y} - \tilde{b} < 1$, we obtain $N_j > 0$ for $j = 0, 1, 2, 3$.[40]

Since the four weights multiplying $N_0, \ldots, N_3$ in (15) are nonnegative and sum to one for every $z \in [0,1]$, it follows that $\partial\Psi(\tilde{x}, z)/\partial\tilde{x} > 0$ for every $z \in [0,1]$. Differentiation gives

$$\frac{d}{d\tilde{x}}\left[(1-\mu)A(\tilde{\beta}) + \mu\frac{\tilde{\beta}}{1-\tilde{\beta}}R(\tilde{\beta})\right] = \int_0^1 (1-z)\frac{\partial\Psi(\tilde{x}, z)}{\partial\tilde{x}}\,dz > 0.$$

Since $d\tilde{\beta}/d\tilde{x} > 0$, the derivative with respect to $\tilde{\beta}$ is also strictly positive. □

[40] The accompanying code derives the four coefficients from Ψ and the formulas for $\tilde{y}'$ and $\tilde{b}'$. For each coefficient, it checks whether the endpoint equations and the stated restrictions can hold together with $N_j \le 0$.

Lemma A.4. *Suppose that $\bar{w} > 0$. Define $TS(\hat{k}) := \lim_{k \uparrow \hat{k}} TS(k) = V + \bar{w}/2$. Then, for every $k \in (0, \hat{k})$,*

$$TS(k) < \max\{TS(0), TS(\hat{k})\}.$$

Proof. For $k \in [0, \hat{k})$, define $A_k := \Pr(\Delta(k) < \alpha(k))$ and $R_k := \Pr(\Delta(k) > \beta(k))$. Let $G(k) := V + [\mu s + (1-\mu)]/2 - TS(k)$ denote the welfare gap relative to the first-best benchmark. The welfare formula gives

$$G(k) = (1-\mu)A_k + \mu s R_k + k\mu(1-\mu)(1+s)^2(1 - A_k - R_k). \tag{16}$$

By Lemma 5, as $k \uparrow \hat{k}$, we have $A_k \to 1$, $R_k \to 0$, and $1 - A_k - R_k \to 0$. Hence $G(\hat{k}) = 1 - \mu$, which gives $TS(\hat{k}) = V + \bar{w}/2$.

Suppose first that $1/[\mu(1+s)^2] \le k < \hat{k}$. Conditional on $\Delta < \alpha$, the corresponding welfare loss is $1 - \mu$; conditional on $\alpha \le \Delta \le \beta$, it is the information cost $k\mu(1-\mu)(1+s)^2 \ge 1 - \mu$; and conditional on $\Delta > \beta$, it is μs. All three losses are at least $1 - \mu$. Moreover, $b < a - \alpha$ implies $(a + L) - b > \beta$, and hence $R_k > 0$. Since $\mu s > 1 - \mu$, the expected loss is therefore strictly larger than $1 - \mu$. Thus $G(k) > G(\hat{k})$ and $TS(k) < TS(\hat{k})$.

It remains to consider $0 < k < 1/[\mu(1+s)^2]$. Set $\tau := k\mu(1+s)^2 \in (0, 1)$. Then $\alpha = \tau - 1$, $-\alpha = 1 - \tau$, and $\mu\beta = \mu s - \tau(1-\mu)$. Since $G(0) = (1-\mu)A_0 + \mu s R_0$ and $G(\hat{k}) = 1 - \mu$, rearranging (16) gives

$$G(k) - \big[(1-\tau)G(0) + \tau G(\hat{k})\big] = (1-\tau)\left[(1-\mu)(A_k - A_0) + \mu\left(-\frac{\beta}{\alpha}R_k - sR_0\right)\right].$$

By Lemma A.3, the expression in braces on the right-hand side is strictly positive.[41] Therefore $G(k) > (1-\tau)G(0) + \tau G(\hat{k})$. Equivalently, $TS(k) < (1-\tau)TS(0) + \tau TS(\hat{k}) \le \max\{TS(0), TS(\hat{k})\}$. This proves the result for all $k \in (0, \hat{k})$. □

Lemma A.5. *Let $A_0 = \Pr(\Delta(0) < -1)$ and $R_0 = \Pr(\Delta(0) > s)$. Define $\Lambda(\mu, s) := A_0 + \mu s R_0/(1-\mu)$. For every $s > 0$ and $\mu \in [1/(1+s), 1)$,*

$$TS(0) - TS(\hat{k}) = (1-\mu)[1 - \Lambda(\mu, s)]. \tag{17}$$

Moreover, $\Lambda(1/(1+s), s) = 1 - P_0 < 1$. On the region $s > 0$ and $\mu > 1/(1+s)$, Λ is continuously differentiable and the following properties hold.

(i) Fix $\bar{w} > 0$. Along the constant-$\bar{w}$ curve $s = (1+\bar{w})/\mu - 1$, for $\mu \in (0, 1)$, at every point satisfying $\Lambda(\mu, (1+\bar{w})/\mu - 1) = 1$, $\frac{d}{d\mu}\Lambda(\mu, (1+\bar{w})/\mu - 1) > 0$.

[41] Indeed, $\tilde{\beta} = \beta/L = 1 - \mu + \bar{w}/L$ is strictly increasing in k. Since $0 < k < 1/[\mu(1+s)^2]$, we have $\alpha < 0$, and thus $\tilde{\beta} \in (1-\mu, 1)$ and $\alpha/L = \tilde{\beta} - 1$. Hence $A_k = A(\tilde{\beta})$, $R_k = R(\tilde{\beta})$, and $-\beta/\alpha = \tilde{\beta}/(1-\tilde{\beta})$. At $k = 0$, $\tilde{\beta} = s/(1+s)$, so $\tilde{\beta}/(1-\tilde{\beta}) = s$. Thus the term in braces is precisely the increase in the expression of Lemma A.3 from $s/(1+s)$ to $\tilde{\beta}$.

(ii) For every fixed $\bar{w} > 0$, $\lim_{\mu \downarrow 0} \Lambda(\mu, (1+\bar{w})/\mu - 1) < 1/2$.

(iii) At every point satisfying $\Lambda(\mu, s) = 1$, $\Lambda_\mu(\mu, s) > 0$ *and* $\Lambda_s(\mu, s) > 0$.

(iv) For every fixed $s > 0$, $\Lambda(\mu, s) \to \infty$ *as* $\mu \uparrow 1$.

Proof. The welfare losses relative to TS^{FB} at $k = 0$ and $k = \hat{k}$ are $(1-\mu)A_0 + \mu s R_0$ and $1-\mu$, respectively. This gives (17). When $\bar{w} = 0$, we have $\mu s = 1 - \mu$, and $\Lambda = A_0 + R_0 = 1 - P_0 < 1$.

For the remaining calculations, suppose that $\mu > 1/(1+s)$ and set $k = 0$, so $L = 1 + s$, $\tilde{\beta} = s/(1+s)$, and $1 - \mu < \tilde{\beta} < 1$. Write $a_0 = a(0)$, $b_0 = b(0)$, and $\rho = \mu/(1-\mu)$. The endpoint equations (3)-(4) give, respectively,

$$\rho = \frac{a_0(b_0 + s)}{(s - a_0 + b_0)(a_0 + L)} = \frac{(1 + a_0 - b_0)(b_0 + L)}{(a_0 + 1)b_0}. \tag{18}$$

Let $g(z) := -\log(1-z) - z$, and define

$$u := \frac{s - a_0 + b_0}{(a_0 + 1)(b_0 + s)}, \quad v := \frac{s(1 + a_0 - b_0)}{(a_0 + 1)(b_0 + s)}. \tag{19}$$

Since $a_0 < b_0 + s$ and $b_0 < a_0 + 1$, we have $u, v \in (0, 1)$. Then we have

$$\begin{aligned} A_0 &= \int_{a_0}^{b_0+s} [1 - F_2(p+1; 0)]\, dF_1(p; 0) = \rho(a_0 + 1)(b_0 + s) g(u), \\ R_0 &= \int_{b_0}^{a_0+1} [1 - F_1(p+s; 0)]\, dF_2(p; 0) = \frac{(a_0 + 1)(b_0 + s)}{\rho s^2} g(v). \end{aligned}$$

Substituting into $\Lambda = A_0 + \rho s R_0$ and using $su + v = s(1+s)/[(a_0+1)(b_0+s)]$ yields

$$\Lambda = \frac{(1+s)\big(\rho s g(u) + g(v)\big)}{su + v} = \frac{\rho s g(u) + g(v)}{\tilde{\beta} u + (1 - \tilde{\beta}) v}. \tag{20}$$

Lemma A.1 implies that a_0 and b_0 are continuously differentiable in (μ, s). Then Λ is also continuously differentiable by (20).

Part (i). Fix $\bar{w} > 0$ and vary μ along $s = (1+\bar{w})/\mu - 1$. In this part, primes on equilibrium variables denote total derivatives with respect to μ with $\bar{w}$ fixed. We first prove $u' < 0$ and $v' > v/\mu$. We will use the following bounds:

$$0 < \tilde{b} < \frac{1 - \mu}{\mu}, \qquad \tilde{b}^2 < 1 - \mu. \tag{21}$$

Lemma A.1 gives $0 < \tilde{b} < (1-\mu)/\mu$. To prove $\tilde{b}^2 < 1 - \mu$, let $q = \sqrt{1-\mu}$. If $(1-\mu)/\mu \le q$, then $\tilde{b} < (1-\mu)/\mu \le q$, so $\tilde{b}^2 < 1 - \mu$. Otherwise, $(1-\mu)/\mu > q$. Since $1 - \mu = q^2$, this implies $q > \mu = 1 - q^2$, and hence $q > 1/2$. Therefore, $\tilde{x}_q = (q^2 + q - 1)/(2q - 1) \in (0, 1)$. Substituting $\tilde{x}_q$

into (10) gives $\tilde{b}(\tilde{x}_q) = q$. At $\tilde{x} = \tilde{x}_q$, subtracting the right-hand side of (9) from the left and evaluating at $\tilde{y} = q + q^2$ gives

$$(1-\mu)(q+q^2)^2 - \tilde{x}_q(q+q^2) - \mu\tilde{x}_q(1-\tilde{x}_q) = \frac{q^3(q+1)(2q^2-1)^2}{(2q-1)^2} \geq 0.$$

The quadratic has positive leading coefficient and negative constant term, so it has one positive and one negative root. Since its valeu at $\tilde{y} = q + q^2 > 0$ is nonnegative, its positive root satisfies $\tilde{y}(\tilde{x}_q) \leq q + q^2$. Hence $\tilde{y}(\tilde{x}_q) - \tilde{b}(\tilde{x}_q) \leq q^2 = 1 - \mu < \tilde{\beta}$. By Lemma A.1, $\tilde{y} - \tilde{b}$ increases and $\tilde{b}$ decreases with $\tilde{x}$. Thus $\tilde{y} - \tilde{b} = \tilde{\beta}$ requires $\tilde{x} > \tilde{x}_q$, implying $\tilde{b} < q \Leftrightarrow \tilde{b}^2 < 1 - \mu$.

We next derive a lower bound on $\tilde{b}'$. Let $\ell := (1-\tilde{x})/\tilde{y} > 0$. After substituting $1 - \tilde{x} = \ell\tilde{y}$ into (9), we have

$$\tilde{y} = \frac{1+\mu\ell}{(1+\ell)(1-\mu+\mu\ell)}, \qquad \tilde{x} = \frac{1-\mu}{(1+\ell)(1-\mu+\mu\ell)}. \tag{22}$$

Regard $\tilde{y}$ and $\tilde{b}$ as functions of (ℓ, μ), with subscripts denoting partial derivatives. At fixed ℓ, it follows that $\tilde{y}_\mu = 1/[(1+\ell)(1-\mu+\mu\ell)^2] > 0$ and $\tilde{x}_\mu = -\ell/[(1+\ell)(1-\mu+\mu\ell)^2]$. Differentiating (10) with respect to μ,

$$\big(2\mu\tilde{b} + (2\mu-1)(1-\tilde{x})\big)\tilde{b}_\mu = (1-\tilde{x})\left[-\frac{1+\tilde{b}}{\mu} + \frac{(1-\mu)-(2\mu-1)\tilde{b}}{(1-\mu+\mu\ell)(1+\mu\ell)}\right] < 0.$$

The coefficient on the left-hand side equals $\mu\tilde{b} + (1-\mu)(1-\tilde{x})/\tilde{b} > 0$ by (10). The second term in brackets on the right-hand side is less than $1 + \tilde{b}$: its numerator is $(1-\mu)(1+\tilde{b}) - \mu\tilde{b}$ and its denominator exceeds $1 - \mu$. Since $1 + \tilde{b} < (1+\tilde{b})/\mu$, the bracket is negative. Thus $\tilde{b}_\mu < 0$, and $\tilde{y}_\mu - \tilde{b}_\mu > 0$. At fixed μ, $\tilde{y}$ and $\tilde{y} - \tilde{b}$ increase with $\tilde{x}$ by Lemma A.1, whereas $\ell = (1-\tilde{x})/\tilde{y}$ decreases. Hence $\tilde{y}_\ell - \tilde{b}_\ell < 0$.

Along the constant-$\bar{w}$ curve, ℓ changes with μ, so $\tilde{b}' = \tilde{b}_\mu + \tilde{b}_\ell\ell'$, and similarly for $\tilde{y}'$. Since $\tilde{\beta} = 1 - \mu/(1+\bar{w})$, $\tilde{\beta}' = -(1-\tilde{\beta})/\mu < 0$. Applying the chain rule to $\tilde{y} - \tilde{b} = \tilde{\beta}$,

$$(\tilde{y}_\ell - \tilde{b}_\ell)\ell' = \tilde{\beta}' - (\tilde{y}_\mu - \tilde{b}_\mu) < 0.$$

The coefficient of ℓ' is negative, so $\ell' > 0$. Rearranging (10), and then using (22) and $1 + \tilde{b} - \tilde{x} = \tilde{y} - \tilde{x} + 1 - \tilde{\beta}$, gives

$$\frac{1}{\tilde{b}} - \rho = \frac{\tilde{x}}{1+\tilde{b}-\tilde{x}} = \frac{1-\mu}{(1+\ell)\big(\mu + (1-\tilde{\beta})(1-\mu+\mu\ell)\big)}.$$

In the last expression, the numerator decreases with μ, while the denominator increases. Indeed, $1 + \ell$ increases since $\ell' > 0$, and $\mu + (1-\tilde{\beta})(1-\mu+\mu\ell)$ is also increasing, since its derivative is $\tilde{\beta} + (1-\tilde{\beta})(\ell + \mu\ell') - \tilde{\beta}'(1-\mu+\mu\ell) > 0$. Hence $(1/\tilde{b} - \rho)' < 0$. Using $\rho' = 1/(1-\mu)^2$

and $\mu\tilde{b} < 1-\mu$ from (21), we obtain

$$\tilde{b}' > -\frac{\tilde{b}^2}{(1-\mu)^2} > -\frac{\tilde{b}(1+\tilde{b})}{1-\mu}. \tag{23}$$

Using the definitions of u and v, together with (18) and $1+\tilde{y}-\tilde{x} = (1+\ell)\tilde{y}$, we have

$$\frac{u}{1-u} = \frac{1-\tilde{\beta}}{\rho(1+\ell)(1+\tilde{b})}, \qquad \frac{v}{\mu} = \frac{\tilde{\beta}\tilde{b}}{(1-\mu)(1+\tilde{b})(\tilde{b}+\tilde{\beta})}.$$

For u, using $\rho'/\rho = 1/\mu + 1/(1-\mu)$ and $-\tilde{\beta}'/(1-\tilde{\beta}) = 1/\mu$, logarithmic differentiation yields

$$\frac{u'}{u(1-u)} = -\frac{1}{1-\mu} - \frac{\ell'}{1+\ell} - \frac{\tilde{b}'}{1+\tilde{b}} < -\frac{1}{1-\mu} + \frac{\tilde{b}^2}{(1-\mu)^2(1+\tilde{b})} < 0.$$

The first inequality follows from $\ell' > 0$ and $\tilde{b}' > -\tilde{b}^2/(1-\mu)^2$ in (23). The last follows from $\tilde{b}^2 < 1-\mu$ in (21). Hence $u' < 0$.

For v/μ, logarithmic differentiation gives

$$\frac{(v/\mu)'}{v/\mu} = \frac{1}{1-\mu} + \frac{\tilde{\beta}'\tilde{b}}{\tilde{\beta}(\tilde{b}+\tilde{\beta})} + \frac{\tilde{\beta}-\tilde{b}^2}{\tilde{b}(1+\tilde{b})(\tilde{b}+\tilde{\beta})}\tilde{b}' > \frac{\tilde{b}}{\tilde{b}+\tilde{\beta}}\left(\frac{1+\tilde{b}}{1-\mu} + \frac{\tilde{\beta}'}{\tilde{\beta}}\right) > 0.$$

Since $\tilde{b}^2 < 1-\mu < \tilde{\beta}$, the coefficient of $\tilde{b}'$ is positive. The first inequality follows from $\tilde{b}' > -\tilde{b}(1+\tilde{b})/(1-\mu)$ in (23). For the last inequality, $-\tilde{\beta}'/\tilde{\beta} = (1-\tilde{\beta})/(\mu\tilde{\beta}) < 1/(1-\mu)$, where the inequality is equivalent to $1-\tilde{\beta} < \mu$. Hence the term in parentheses is greater than $\tilde{b}/(1-\mu) > 0$, and therefore $v' > v/\mu$.

We now show that $\Lambda' > 0$ whenever $\Lambda = 1$. The definition of u implies $u/(1-u) = (s-a_0+b_0)/[a_0(b_0+L)]$. Since $g'(z) = z/(1-z)$, (18) gives

$$\rho s g'(u) = \frac{s(b_0+s)}{(a_0+L)(b_0+L)} < \frac{s}{L} = \tilde{\beta}, \tag{24}$$

where the inequality follows from $a_0 > 0$ and $b_0 + L > b_0 + s$. For every $z \in (0,1)$,

$$g(z) = \int_0^z \frac{t}{1-t}\,dt < \frac{1}{1-z}\int_0^z t\,dt = \frac{z^2}{2(1-z)} = \frac{zg'(z)}{2}. \tag{25}$$

Combining (24) and (25) implies $\rho s g(u) < \tilde{\beta}u/2$. At $\Lambda = 1$, (20) therefore gives $g(v) = \tilde{\beta}u + (1-\tilde{\beta})v - \rho s g(u) > (1-\tilde{\beta})v$. Together with $g(v) < vg'(v)/2$ from (25), this implies $g'(v) > 2(1-\tilde{\beta})$.

At $\Lambda = 1$, differentiating (20) along the constant-$\bar{w}$ curve gives

$$[\tilde{\beta}u + (1-\tilde{\beta})v]\Lambda' = \frac{\bar{w}}{(1-\mu)^2}g(u) + [\rho s g'(u) - \tilde{\beta}]u' + [g'(v) - (1-\tilde{\beta})]v' + \tilde{\beta}'(v-u),$$

where we use $(\rho s)' = \bar{w}/(1-\mu)^2$ along the constant-$\bar{w}$ curve. The first term is positive. The coefficient of u' is negative by (24), and $u' < 0$. Also, $-\tilde{\beta}'u > 0$, and $[g'(v) - (1-\tilde{\beta})]v' + \tilde{\beta}'v > v[g'(v) - 2(1-\tilde{\beta})]/\mu > 0$, using $v' > v/\mu$ and $\tilde{\beta}' = -(1-\tilde{\beta})/\mu$. Therefore, $\Lambda' > 0$.

Part (ii). Fix $\bar{w} > 0$ and let $s = (1+\bar{w})/\mu - 1$. Lemma A.1 shows that $\tilde{y} - \tilde{x}$ increases to $\mu/(1-\mu)$. Using $\mu L = 1 + \bar{w}$, we obtain $0 < a_0 = L(\tilde{y} - \tilde{x}) < (1+\bar{w})/(1-\mu)$. Also, $0 < b_0 < a_0 - \alpha = a_0 + 1$. As $\mu \downarrow 0$, both endpoints remain bounded. Using $\mu L = 1 + \bar{w}$, (3) and (4) can be written, respectively, as

$$\frac{b_0 + s}{L}\big((2\mu - 1)a_0 + 1 + \bar{w}\big) = \mu a_0\left(1 + \frac{a_0}{L}\right),$$
$$(a_0 + 1)\left(1 - \mu - (2\mu - 1)\frac{b_0}{L}\right) = (1-\mu)b_0\left(1 + \frac{b_0}{L}\right).$$

As $\mu \downarrow 0$, $s \to \infty$ and $L \to \infty$. Because a_0 and b_0 are bounded, every limit point must satisfy $-a_0 + 1 + \bar{w} = 0$ and $a_0 + 1 = b_0$, respectively. Hence $a_0 \to 1 + \bar{w}$ and $b_0 \to 2 + \bar{w}$.

Substituting into (19) gives $u \to 1/(2+\bar{w})$ and $v \to 0$. Also, $\rho s \to 1 + \bar{w}$ and $\tilde{\beta} \to 1$. Equation (20) implies $\lim_{\mu \downarrow 0} \Lambda = (1+\bar{w})(2+\bar{w})g(1/(2+\bar{w})) < 1/2$, where the inequality follows from (25) at $z = 1/(2+\bar{w})$.

Part (iii). Here subscripts on $\Lambda(\mu, s)$ denote partial derivatives. We first prove $\Lambda_s > 0$. At $k = 0$, $A_0 = A(\tilde{\beta})$ and $R_0 = R(\tilde{\beta})$ in the notation of Lemma A.3. Hence $(1-\mu)\Lambda = (1-\mu)A(\tilde{\beta}) + \mu\tilde{\beta}R(\tilde{\beta})/(1-\tilde{\beta})$. That lemma gives a strictly positive derivative of the right-hand side with respect to $\tilde{\beta} \in (1-\mu, 1)$. Since $d\tilde{\beta}/ds = 1/(1+s)^2 > 0$ at fixed μ, we obtain $\Lambda_s > 0$.

At any point with $\Lambda(\mu, s) = 1$, let $\bar{w} = \mu(1+s) - 1 > 0$. Part (i) then implies

$$0 < \frac{d}{d\mu}\Lambda\left(\mu, \frac{1+\bar{w}}{\mu} - 1\right) = \Lambda_\mu(\mu, s) - \frac{1+s}{\mu}\Lambda_s(\mu, s).$$

Therefore $\Lambda_\mu(\mu, s) > (1+s)\Lambda_s(\mu, s)/\mu > 0$.

Part (iv). Fix s and let $\mu \uparrow 1$. Equation (21) gives $\tilde{b} \to 0$, so $\tilde{y} \to s/L > 0$. By (9), $\tilde{x} \le (1-\mu)\tilde{y} \to 0$. Thus $b_0 \to 0$, $a_0 \to s$, and (19) gives $v \to 1$. Since $g(v) \to \infty$, we have $\Lambda \to \infty$ from (20). □

Proof of Proposition 3. The case $\bar{w} = 0$ was established above. Suppose that $\bar{w} > 0$, or equivalently, $\mu > 1/(1+s)$. By Lemma A.4, it is enough to compare total surplus at $k = 0$ and $k = \hat{k}$.

By Lemma A.5,

$$TS(0) - TS(\hat{k}) = (1-\mu)[1-\Lambda(\mu, s)].$$

For fixed s,

$$\Lambda\left(\frac{1}{1+s}, s\right) = 1 - P_0 < 1,$$

whereas $\Lambda(\mu, s) \to \infty$ as $\mu \uparrow 1$. Continuity therefore implies that $\Lambda(\mu, s) = 1$ has at least one solution in $(1/(1+s), 1)$.

At every such solution, Lemma A.5 gives $\Lambda_\mu(\mu, s) > 0$. Hence the solution is unique. To see this, note that every solution crosses the level curve from below. If there were two solutions, continuity would require another solution between them at which Λ crosses from above, a contradiction. Denote the unique solution by $\mu^*(s)$. It follows that $\Lambda(\mu, s) < 1$ for $\mu < \mu^*(s)$ and $\Lambda(\mu, s) > 1$ for $\mu > \mu^*(s)$. Therefore, total surplus is maximized at $k = 0$ if $\mu < \mu^*(s)$ and at $k = \hat{k}$ if $\mu \geq \mu^*(s)$.

At $\mu = \mu^*(s)$, Lemma A.5 gives $\Lambda_\mu > 0$ and $\Lambda_s > 0$. The implicit function theorem therefore implies that $\mu^*(s)$ is continuously differentiable and

$$\mu^{*\prime}(s) = -\frac{\Lambda_s(\mu^*(s), s)}{\Lambda_\mu(\mu^*(s), s)} < 0.$$

To establish the limit, fix any $\mu \in (0, 1)$ and let $s \to \infty$. At $k = 0$, $L = 1 + s$ and $\tilde{\beta} = s/(1+s) \to 1$. Since 1 lies strictly inside the range of $\tilde{y} - \tilde{b}$ in Lemma A.1, $\tilde{x}$ converges to some $\tilde{x}_\infty \in (0, 1)$. Hence the upper endpoint of the support of Δ/L, $\tilde{\beta} + 1 - \tilde{x}$, converges to $2 - \tilde{x}_\infty > 1$. The normalized equilibrium price distributions converge to nondegenerate limits, and thus $R_0 = \Pr(\Delta(0) > s)$ converges to a strictly positive value, since $s/L \to 1$. Therefore,

$$TS(\hat{k}) - TS(0) = \mu s R_0 - (1-\mu)(1 - A_0) \to \infty.$$

Thus, for every fixed $\mu > 0$, $\Lambda(\mu, s) > 1$ for all sufficiently large s, and hence $\mu > \mu^*(s)$. Since $\mu > 0$ is arbitrary, $\lim_{s\to\infty} \mu^*(s) = 0$. □

Proof of Corollary 3. Fix $\bar{w} > 0$ and vary μ along the curve $\bar{w} = \mu s - (1-\mu)$, so that $s = (1+\bar{w})/\mu - 1$. Note that

$$TS(0) - TS(\hat{k}) = (1-\mu)[1-\Lambda(\mu, s)].$$

By part (ii) of Lemma A.5, $\Lambda(\mu, s) < 1$ for all sufficiently small $\mu > 0$, and hence $TS(0) > TS(\hat{k})$.

As $\mu \uparrow 1$, we have $s \downarrow \bar{w}$. By Proposition 3, $\mu^*(s)$ is continuous and $\mu^*(\bar{w}) < 1$. Hence $\mu - \mu^*(s) \to 1 - \mu^*(\bar{w}) > 0$. Thus, for μ sufficiently close to one, $\mu > \mu^*(s)$, and Proposition 3 implies $TS(\hat{k}) > TS(0)$, or equivalently, $\Lambda(\mu, s) > 1$.

Continuity therefore implies that $\Lambda(\mu, s) = 1$ has at least one solution along the constant-$\bar{w}$ curve. At every such solution, part (i) of Lemma A.5 gives $d\Lambda(\mu, (1+\bar{w})/\mu - 1)/d\mu > 0$. By the same argument as in the proof of Proposition 3, the solution is unique. Denote it by $\mu^\dagger(\bar{w})$. It follows that $\Lambda(\mu, s) < 1$ for $\mu < \mu^\dagger(\bar{w})$ and $\Lambda(\mu, s) > 1$ for $\mu > \mu^\dagger(\bar{w})$. Together with Lemma A.4, this proves the result. □

Proof of Proposition 4. Let primes denote derivatives with respect to k. Equations (3)-(4) can be written as $b = \psi_1(a, k) = \psi_2(a, k)$, where

$$\psi_1(a, k) := \frac{\mu a(a + L)}{(2\mu - 1)a + \mu L} - \beta,$$

while $\psi_2(a, k)$ is the positive solution of $b^2 + [(2\mu - 1)(a - \alpha)/(1 - \mu) + L]b - (a - \alpha)L = 0$. Direct differentiation gives

$$\psi_{1a} - 1 = \frac{(1-\mu)a[(2\mu-1)a + 2\mu L]}{[(2\mu-1)a + \mu L]^2} > 0, \quad \psi_{2a} = \frac{L - \frac{2\mu-1}{1-\mu}b}{2b + \frac{2\mu-1}{1-\mu}(a-\alpha) + L} \in (0, 1).$$

For the first expression, rearranging (3) gives $b + \beta = \mu a(a + L)/[(2\mu - 1)a + \mu L]$. Since $b + \beta > 0$, $a > 0$, and $L > 0$, we have $(2\mu - 1)a + \mu L > 0$, and hence $(2\mu - 1)a + 2\mu L > 0$. For the second expression, rearranging (4) gives $(a - \alpha)[L - (2\mu - 1)b/(1 - \mu)] = b(b + L) > 0$, so its numerator is positive. Moreover, since $a - \alpha < b + L$, (4) implies $b > (1 - \mu)(a - \alpha)$, and hence $b + (2\mu - 1)(a - \alpha) > 0$. The denominator is therefore strictly larger than the numerator. Hence $\psi_{1a} > 1 > \psi_{2a} > 0$.

Along $b = \psi_j(a, k)$, consider a change in k that keeps $a + C$ fixed. Since $da/dk = -C'$, the corresponding change in $b + C$ is $\psi_{jk} + C'(1 - \psi_{ja})$.

Using $\tilde{b} = b/L$, differentiation and equations (3)-(4) give

$$\psi_{1k} + C'(1 - \psi_{1a}) = \frac{(1+s)^2(1-\mu)\mu^2\left[\left(L + \frac{a(\mu^2+\mu-1)}{\mu}\right)^2 + a^2(2-\mu^2)\right]}{[(2\mu-1)a + \mu L]^2} > 0,$$

$$\psi_{2k} + C'(1 - \psi_{2a}) = -\frac{(1+s)^2 L\mu\left[\left(\mu\tilde{b} - (1-\mu)\right)^2 + 2(1-\mu)^3\tilde{b}(\tilde{b}+1)\right]}{\left[2b + \frac{2\mu-1}{1-\mu}(a-\alpha) + L\right](1-\mu)\left[(1-\mu) - (2\mu-1)\tilde{b}\right]} < 0.$$

The denominator in the second expression is positive because (10) implies $(1-\mu) - (2\mu-1)\tilde{b} > 0$.

Differentiating $\psi_1(a, k) = \psi_2(a, k)$ and rearranging yields

$$(\psi_{1a} - \psi_{2a})(a' + C') = \left[\psi_{2k} + C'(1 - \psi_{2a})\right] - \left[\psi_{1k} + C'(1 - \psi_{1a})\right] < 0.$$

Since $\psi_{1a} > \psi_{2a}$, it follows that $a' + C' < 0$. Moreover, since $b = \psi_2(a, k)$,

$$b' + C' = \psi_{2a}(a' + C') + \psi_{2k} + C'(1 - \psi_{2a}) < 0.$$

Thus $b' + C' < 0$.

For $u \in (0,1)$, let $Q_i(u,k) := F_i^{-1}(u;k)$, and recall that $m_1(k) := F_1(b+\beta;k)$ and $m_2(k) := F_2(a-\alpha;k)$. Inverting equations (1)-(2) and differentiating gives

$$\begin{aligned}
\frac{\partial}{\partial k}(Q_1 + C) &= \frac{a' + C' - u(\alpha' + C')}{1-u} < 0, && u < m_1(k),\\
\frac{\partial}{\partial k}(Q_1 + C) &= \beta' + C' + \frac{(1-\mu)(a' - \alpha')}{1-\mu u} < 0, && u > m_1(k),\\
\frac{\partial}{\partial k}(Q_2 + C) &= \frac{b' + C' + u(\beta' - C')}{1-u} < 0, && u < m_2(k),\\
\frac{\partial}{\partial k}(Q_2 + C) &= -\alpha' + C' + \frac{\mu(b' + \beta')}{1-(1-\mu)u} < 0, && u > m_2(k).
\end{aligned}$$

Indeed, $\alpha' = \mu(1+s)^2$, $\beta' = -(1-\mu)(1+s)^2$, and $C' = \mu(1-\mu)(1+s)^2$. Hence $\alpha' + C' > 0$, $\beta' + C' < 0$, $\beta' - C' < 0$, and $-\alpha' + C' < 0$. Moreover, $a' + C' < 0$ and $b' + C' < 0$ imply $a' - \alpha' < 0$ and $b' + \beta' < 0$.

By Lemma 4, $m_1(k)$ is strictly increasing and $m_2(k)$ is strictly decreasing when $\bar{w} > 0$. When $\bar{w} = 0$, Lemma 3 implies that both are constant. Thus, for each fixed u, $m_i(k)$ can cross u at most once, or coincide with u for all k when $\bar{w} = 0$. In the latter case, $Q_1(u,k) = b + \beta$ or $Q_2(u,k) = a - \alpha$, whose sums with $C(k)$ are also strictly decreasing since $b' + \beta' + C' < 0$ and $a' - \alpha' + C' < 0$. Since $Q_i(u,k)$ is continuous in k, it follows that $Q_i(u,k) + C(k)$ is strictly decreasing in k for every $u \in (0,1)$ and $i \in \{1,2\}$.

Fix any pair of price quantiles $(u_1, u_2) \in (0,1)^2$. For each k, let $p_i(k) = Q_i(u_i, k)$ for $i = 1, 2$. By market coverage, the consumer's equilibrium payoff at these prices is[42]

$$\max\Big\{\mathbb{E}[v_1] - p_1(k),\ \mathbb{E}[v_2] - p_2(k),\ \mathbb{E}[\max\{v_1, v_2\}] - \mu p_1(k) - (1-\mu)p_2(k) - C(k)\Big\}.$$

By the preceding argument, $Q_i(u_i, k) + C(k)$ is strictly decreasing in k for both firms. Since $C(k)$ is strictly increasing, both prices decrease, and the first two terms in the maximum strictly increase. The third term can be written as

$$\mathbb{E}[\max\{v_1, v_2\}] - \mu\big(p_1(k) + C(k)\big) - (1-\mu)\big(p_2(k) + C(k)\big),$$

and is strictly increasing as well. Thus, the consumer's equilibrium payoff strictly increases

[42] The third term is the learning payoff for $-1 < \Delta < s$. Outside this interval, learning does not affect the purchase decision, and the third term is below the relevant no-learning payoff. Hence the same expression can be used inside the maximum for all Δ.

with k for every fixed pair (u_1, u_2). Since the firms' equilibrium prices are independent, integrating over $(u_1, u_2) \in (0,1)^2$ gives $CS(k_2) > CS(k_1)$ whenever $0 < k_1 < k_2 < \hat{k}$. □

A.3 Quasi-Monopoly

Proposition A.1. *Suppose that $\bar{w} > 0$ and $k \in [\hat{k}, \bar{k})$. Under any admissible tie-breaking rule, every equilibrium has $p_1 = \alpha$ almost surely, and the consumer does not acquire information and purchases from Firm 1 almost surely. Moreover,*

$$(\Pi_1, \Pi_2) = (\alpha, 0), \qquad TS = \mathbb{E}[v_1], \qquad CS = \mathbb{E}[v_1] - \alpha.$$

Proof. Since $k \in [\hat{k}, \bar{k})$, we have $\beta > \bar{w} > 0$ and $\alpha - \mu\beta = \mu(2-\mu)(1+s)^2[k - \hat{k}] \geq 0$. Hence $\alpha \geq \mu\beta > 0$.

We first show that $\Pi_2 = 0$. Suppose instead that $\Pi_2 > 0$. Firm 1 also earns strictly positive equilibrium profit, since any price $p \in (0, \beta)$ gives it demand at least μ. Given that both firms make positive profits, the argument in Steps 1-6 of the proof of Proposition 1 implies that both price distributions are atomless and the support endpoints form an interior solution of (3)-(4). By Lemma A.1, such a solution exists only if $k < \hat{k}$, a contradiction. Hence $\Pi_2 = 0$.

We next show that $p_1 \leq \alpha$ almost surely. Otherwise, there would exist a sufficiently small $q > 0$ such that $\Pr(p_1 > \alpha + q) > 0$. If Firm 2 deviated to $p_2 = q$, its demand would be at least $1 - \mu$ whenever $p_1 > \alpha + q$. This deviation would yield strictly positive profit, contradicting $\Pi_2 = 0$.

For every $\varepsilon \in (0, \alpha)$, Firm 1 can charge $p_1 = \alpha - \varepsilon$ and sell for sure, since $p_2 \geq 0$ implies $p_1 - p_2 < \alpha$. Letting $\varepsilon \downarrow 0$ gives $\Pi_1 \geq \alpha$. Combining this with $p_1 \leq \alpha$ almost surely yields

$$\alpha \leq \Pi_1 = \mathbb{E}[p_1 D_1(p_1 - p_2)] \leq \mathbb{E}[p_1] \leq \alpha.$$

Thus $\Pi_1 = \alpha$. Since $p_1 \leq \alpha$ almost surely, $\mathbb{E}[p_1] = \alpha$ implies $p_1 = \alpha$ almost surely. As $\alpha > 0$, the equality $\mathbb{E}[p_1 D_1(p_1 - p_2)] = \alpha$ then implies $D_1(p_1 - p_2) = 1$ almost surely.

Since $p_1 = \alpha$ and $p_2 \geq 0$, we have $\Delta \leq \alpha$ almost surely. If $\Delta < \alpha$, the consumer strictly prefers not to acquire information and purchases from Firm 1. If $\Delta = \alpha$ occurs with positive probability, then $D_1(p_1 - p_2) = 1$ almost surely implies $d_\alpha = 1$. Since no learning gives Firm 1 demand one at $\Delta = \alpha$, whereas learning gives demand $\mu < 1$, the consumer must choose no learning. Thus the consumer does not acquire information and purchases from Firm 1 almost surely.

No information cost is incurred, so $TS = V + \bar{w}/2$, the expected value of Firm 1's product. Subtracting the firms' profits gives $CS = V + \bar{w}/2 - \alpha$. □

Proof of Proposition 5. Since $k \in [\hat{k}, \bar{k})$, we have $\beta > \bar{w} > 0$ and $\alpha \geq \mu\beta > 0$. Note that F_2 satisfies $F_2(0) = 0$ and is strictly increasing on $[0, L)$, since $F_2'(p) = \alpha/[(1-\mu)(\alpha + p)^2] > 0$.

Since $L = \beta - \alpha \le (1-\mu)\beta$, we also have

$$F_2(L^-) = \frac{L}{(1-\mu)\beta} \le 1.$$

Thus F_2 is well-defined as a CDF on $[0, L]$. It has an atom at L when $k > \hat{k}$.

Given $p_1 = \alpha$, every feasible price $p_2 > 0$ gives $\Delta = \alpha - p_2 < \alpha$, and Firm 2 receives zero demand. At $p_2 = 0$, its profit is zero regardless of the tie-breaking rule. Thus every feasible price gives Firm 2 zero profit, and F_2 is a best response.

We next consider Firm 1's deviations given F_2. If $p_1 < \alpha$, then $p_1 - p_2 < \alpha$ for every $p_2 \in [0, L]$, so Firm 1 sells for sure and earns $p_1 < \alpha$. If $p_1 = \alpha + t$ for some $t \in [0, L)$, its demand is one when $p_2 > t$ and μ when $p_2 < t$. Since F_2 has no atom at $t < L$,

$$\pi_1(\alpha + t) = (\alpha + t)\big[1 - (1-\mu)F_2(t)\big] = \alpha.$$

Thus every price in $[\alpha, \beta)$ yields profit α.

For $p_1 = \beta + t$ with $t \in (0, L)$, Firm 1 has demand μ when $p_2 > t$ and zero demand when $p_2 < t$. Hence

$$\begin{aligned}\pi_1(\beta + t) &= \mu(\beta + t)\big[1 - F_2(t)\big] \\ &= \mu\beta - \frac{\mu t(L + \mu\alpha + \mu t)}{(1-\mu)(\alpha + t)} < \mu\beta \le \alpha.\end{aligned}$$

Thus every price in $(\beta, \beta + L)$ yields profit less than α.

It remains to check the two boundary prices at which the atom at L may matter. Using $d_\alpha \le 1$ and $d_\beta \le \mu$, we obtain

$$\begin{aligned}\pi_1(\beta) &= \beta\Big[\mu F_2(L^-) + d_\alpha\big(1 - F_2(L^-)\big)\Big] \le \beta\Big[1 - (1-\mu)F_2(L^-)\Big] = \alpha, \\ \pi_1(\beta + L) &= (\beta + L)d_\beta\big[1 - F_2(L^-)\big] \le \mu(\beta + L)\big[1 - F_2(L^-)\big] = \lim_{t \uparrow L} \pi_1(\beta + t) \le \mu\beta \le \alpha.\end{aligned}$$

Finally, if $p_1 > \beta + L$, then $p_1 - p_2 > \beta$ for every $p_2 \in [0, L]$, so Firm 1 earns zero. Thus no deviation gives Firm 1 more than α, while $p_1 = \alpha$ gives profit α. The proposed strategies are therefore mutual best responses.

Since $F_2(0) = 0$, we have $p_2 > 0$ almost surely. Consequently, $\Delta = \alpha - p_2 < \alpha$ almost surely, so the consumer does not acquire information and purchases from Firm 1 almost surely. The resulting profits are $(\Pi_1, \Pi_2) = (\alpha, 0)$.

By Proposition A.1, every equilibrium has $TS = V + \bar{w}/2$ and $CS = V + \bar{w}/2 - \alpha$. Hence TS is independent of k, while $dCS/dk = -\mu(1+s)^2 < 0$. □

Proof of Lemma 6. On the interior mixed-equilibrium branch, $\Pi_1(k) = \mu[b(k) + \beta]$ and $\Pi_2(k) = (1-\mu)[a(k) - \alpha]$. Hence $\Pi_1'(k) = \mu[b'(k) - (1-\mu)(1+s)^2] < 0$ and $\Pi_2'(k) = (1-\mu)[a'(k) - \mu(1+s)^2] < 0$, since $a'(k), b'(k) < 0$. Thus Firm 1's profit and industry profit

are both strictly decreasing in k on the mixed branch.

In the quasi-monopoly region, industry profit equals Firm 1's profit, $\Pi_1(k) = \alpha$, which is strictly increasing in k and bounded above by $\alpha(\bar{k}) = \bar{w}$. At $k = 0$, however, $\Pi_1(0) - \bar{w} = \mu b(0) + (1-\mu) > 0$. Thus Firm 1's profit at $k = 0$ alone exceeds industry profit throughout the quasi-monopoly region. Therefore both Firm 1's profit and industry profit are uniquely maximized at $k = 0$. □

A.4 Endogenous Transparency

Proof of Proposition 6. Suppose first that $\bar{w} > 0$. Let primes on $\tilde{y}$ and $\tilde{b}$ denote derivatives with respect to $\tilde{x}$. By Lemma A.1, $\tilde{y}' > 1$ and $\tilde{b}' < 0$. Twice differentiating the normalized endpoint equations (9)-(10) and substituting their first derivatives gives

$$\tilde{y}'' = -\frac{2\mu^2(1-\mu)}{[2(1-\mu)\tilde{y} - \tilde{x}]^3} < 0, \quad \tilde{b}'' = -\frac{2\mu(1-\mu)^2}{[2\mu\tilde{b} + (2\mu-1)(1-\tilde{x})]^3} < 0,$$

where both denominators are positive at the positive roots. Since $\Pi_1 = \mu L\tilde{y}$ and $\tilde{y} - \tilde{b} = 1 - \mu + \bar{w}/L$, differentiating twice with respect to k yields

$$\Pi_1''(k) = \frac{\mu\bar{w}^2(1+s)^4}{L^3(\tilde{y}' - \tilde{b}')^3}\left(\tilde{y}'\tilde{b}'' - \tilde{y}''\tilde{b}'\right) < 0.$$

When $\bar{w} = 0$, $\Pi_1(k)$ is linear below $\hat{k}$ by Lemma 3. Thus, in either case, $\tilde{\Pi}_1''(k) = \Pi_1''(k) - c''(k) < 0$ for $0 < k < \hat{k}$. Since $\tilde{\Pi}_1$ is continuous at $\hat{k}$, it has a unique maximizer on $[0, \hat{k}]$. If the maximizer is interior, it satisfies $\Pi_1'(k) = c'(k)$.

For $k \geq \hat{k}$, Firm 1's profit $\Pi_1(k)$ is weakly increasing, while $c(k)$ is strictly decreasing. Hence every $k \in [\hat{k}, K)$ gives lower net profit than $k = K$. Firm 1's choice can therefore fail to be unique only if K and the unique maximizer on $[0, \hat{k}]$ give the same net profit.

Now we prove $k^f \leq k^w$. We first show that $TS(k) - \Pi_1(k)$ is strictly increasing on $[0, \hat{k})$. If $\bar{w} = 0$, $(TS - \Pi_1)'(k) = (1+s)\Pi_1(0) - sP_0 = b(0) + s - sP_0 > 0$, since $b(0) > 0$ and $P_0 \leq 1$. Suppose now that $\bar{w} > 0$, and write $A_k = \Pr(\Delta(k) < \alpha(k))$ and $R_k = \Pr(\Delta(k) > \beta(k))$. The welfare decomposition and $\Pi_1 = \mu(b + \beta)$ imply

$$TS(k) - \Pi_1(k) = V - \frac{\bar{w}}{2} - \mu b + (1-\mu)\alpha A_k - \mu\beta R_k.$$

The proof of Lemma 2 gives $b' < 0$, while Lemma 5 shows that A_k increases and R_k decreases with k. Since $\alpha' > 0$, $\beta' < 0$, and $\beta > 0$, αA_k increases whenever $\alpha \geq 0$, and βR_k decreases. Thus $TS - \Pi_1$ strictly increases when $\alpha \geq 0$. For $\alpha < 0$, it remains to show that $\mu b - (1-\mu)\alpha A_k$ decreases.

With D_A as defined in the proof of Lemma A.3, we have $D_A(\tilde{x}, z) = (1+\tilde{b})(\tilde{y} - \tilde{x}) + z\tilde{x}(1-\tilde{\beta})$.

Since $\alpha = L(\tilde{\beta}-1) < 0$ and $\beta = \bar{w} + (1-\mu)L$, we have $1-\mu < \tilde{\beta} < 1$ and $L = \bar{w}/(\tilde{\beta}+\mu-1)$. The integral representation of A_k then gives

$$\mu b - (1-\mu)\alpha A_k = \mu\bar{w}\int_0^1 \Xi\, dz, \qquad \Xi := \frac{\tilde{b} + (1-\tilde{\beta})\tilde{x}^2(1-z)/D_A(\tilde{x},z)}{\tilde{\beta}+\mu-1}.$$

Subscripts on Ξ denote partial derivatives with respect to $\tilde{x}, \tilde{y}, \tilde{b}$, with $\tilde{\beta} = \tilde{y} - \tilde{b}$. Since $\partial D_A/\partial\tilde{y} = 1+\tilde{b}-z\tilde{x} > 0$, increasing $\tilde{y}$ decreases $1-\tilde{\beta}$ and increases D_A. Thus the numerator of Ξ weakly decreases and its denominator increases, and thus $\Xi_{\tilde{y}} < 0$. Also, $\partial D_A/\partial\tilde{b} = \tilde{y}-\tilde{x}+z\tilde{x}$, so $\partial[(1-\tilde{\beta})/D_A]/\partial\tilde{b} = \tilde{y}(\tilde{y}-\tilde{x})/D_A^2 > 0$. Increasing $\tilde{b}$ therefore increases the numerator of Ξ and decreases its denominator, giving $\Xi_{\tilde{b}} > 0$.

We next show that $\Xi_{\tilde{x}} + \Xi_{\tilde{y}} < 0$. Dividing (9) by $\tilde{y}$ gives $(1-\mu)\tilde{y} = \tilde{x} + \mu\tilde{x}(1-\tilde{x})/\tilde{y} > \tilde{x}$, and hence $\tilde{y}-\tilde{x} > \mu\tilde{x}/(1-\mu)$. Similarly, dividing (10) by $\tilde{b}$ gives $\mu\tilde{b} = [(1-\mu)/\tilde{b} - (2\mu-1)](1-\tilde{x}) > (1-\mu)(1-\tilde{x})$, where the inequality uses $\tilde{b} < (1-\mu)/\mu$. Multiplying the two bounds yields $\tilde{b}(\tilde{y}-\tilde{x}) > \tilde{x}(1-\tilde{x})$. Direct differentiation gives

$$\begin{aligned}(\tilde{\beta}+\mu-1)^2 D_A^2(\Xi_{\tilde{x}} + \Xi_{\tilde{y}}) = (1-z)\tilde{x}\big[2(\tilde{\beta}+\mu-1)(1-\tilde{\beta}) - \mu\tilde{x}\big]D_A(\tilde{x},0) \\ + z(1-z)\tilde{x}^2(1-\tilde{\beta})^2(\tilde{\beta}+\mu-1-\tilde{x}) - \tilde{b}D_A^2.\end{aligned}$$

Since $\tilde{b}(\tilde{y}-\tilde{x}) > \tilde{x}(1-\tilde{x})$ and $1+\tilde{b} > \mu$, we have $\tilde{b}D_A(\tilde{x},0) = (1+\tilde{b})\tilde{b}(\tilde{y}-\tilde{x}) > \mu\tilde{x}(1-\tilde{x})$. Moreover, $\tilde{\beta}+\mu-1$ and $1-\tilde{\beta}$ sum to μ, so $2(\tilde{\beta}+\mu-1)(1-\tilde{\beta}) \le \mu^2/2 < \mu$. Therefore, $\tilde{x}[2(\tilde{\beta}+\mu-1)(1-\tilde{\beta}) - \mu\tilde{x}]D_A(\tilde{x},0) < \mu\tilde{x}(1-\tilde{x})D_A(\tilde{x},0) < \tilde{b}D_A(\tilde{x},0)^2$.

If $\tilde{\beta}+\mu-1 \le \tilde{x}$, then $\tilde{x}^2(1-\tilde{\beta})^2(\tilde{\beta}+\mu-1-\tilde{x}) \le 0 < \tilde{b}D_A(\tilde{x},0)^2$. Otherwise, $\tilde{\beta} > \tilde{x}+1-\mu > \tilde{x}$, so $\tilde{x}(1-\tilde{\beta}) < \tilde{x}(1-\tilde{x}) < \tilde{b}(\tilde{y}-\tilde{x})$. Also, $\tilde{\beta} < 1$ and $\tilde{x} > 0$ imply $0 < \tilde{\beta}+\mu-1-\tilde{x} < \mu < 1+\tilde{b}$. Squaring the bound on $\tilde{x}(1-\tilde{\beta})$ and multiplying by $\tilde{\beta}+\mu-1-\tilde{x}$ gives $\tilde{x}^2(1-\tilde{\beta})^2(\tilde{\beta}+\mu-1-\tilde{x}) < \tilde{b}^2(\tilde{y}-\tilde{x})^2(1+\tilde{b}) < \tilde{b}(1+\tilde{b})^2(\tilde{y}-\tilde{x})^2 = \tilde{b}D_A(\tilde{x},0)^2$, where the second inequality uses $\tilde{b} < 1+\tilde{b}$. Since $(1-z)+z(1-z) = 1-z^2$ and $D_A \ge D_A(\tilde{x},0)$, for $0 \le z < 1$ we obtain

$$(\tilde{\beta}+\mu-1)^2 D_A^2(\Xi_{\tilde{x}} + \Xi_{\tilde{y}}) < (1-z^2)\tilde{b}D_A(\tilde{x},0)^2 - \tilde{b}D_A^2 \le 0.$$

Thus $\Xi_{\tilde{x}} + \Xi_{\tilde{y}} < 0$ for $z < 1$; at $z = 1$, $\Xi_{\tilde{x}} + \Xi_{\tilde{y}} = -\tilde{b}/(\tilde{\beta}+\mu-1)^2 < 0$.

By Lemma A.1, $\tilde{y}' > 1$ and $\tilde{b}' < 0$, where these primes denote derivatives with respect to $\tilde{x}$. The chain rule gives $d\Xi/d\tilde{x} = (\Xi_{\tilde{x}} + \Xi_{\tilde{y}}) + (\tilde{y}'-1)\Xi_{\tilde{y}} + \tilde{b}'\Xi_{\tilde{b}} < 0$. Since $\tilde{x}$ increases with k, Ξ strictly decreases with k for every $z \in [0,1]$. Hence $\mu b - (1-\mu)\alpha A_k$ strictly decreases. As βR_k also decreases, $TS - \Pi_1$ strictly increases when $\alpha < 0$. By continuity, $TS - \Pi_1$ is strictly increasing on $[0, \hat{k})$.

We now compare the optimal choices. On $[\hat{k}, K]$, Π_1 is weakly increasing, TS is constant, and c is strictly decreasing. Thus $\Pi_1 - c$ and $TS - c$ are strictly increasing there, so $k^f, k^w \notin$

$[\hat{k}, K)$. Suppose, to the contrary, that $k^w < k^f$. Then $k^w < \hat{k}$, and optimality implies

$$\Pi_1(k^f) - \Pi_1(k^w) \geq c(k^f) - c(k^w) \geq TS(k^f) - TS(k^w).$$

Hence $TS(k^f) - \Pi_1(k^f) \leq TS(k^w) - \Pi_1(k^w)$. If $k^f < \hat{k}$, this contradicts the strict increase of $TS - \Pi_1$ on $[0, \hat{k})$.

It remains to consider $k^f = K$. At K, $TS(K) = V + \bar{w}/2$ and $\Pi_1(K) = \bar{w}$. Optimality of k^w gives $TS(k^w) - TS(K) \geq c(k^w) - c(K) > 0$. The welfare decomposition gives $TS(k^w) - TS(K) = -(1-\mu)\alpha(k^w)(1 - A_{k^w}) - \mu\beta(k^w)R_{k^w}$. Since $\beta(k^w) > 0$, $TS(k^w) > TS(K)$ requires $\alpha(k^w) < 0$. Using $\alpha(k^w) < 0$ and $b(k^w) > 0$ in the formula for $TS - \Pi_1$ gives $TS(k^w) - \Pi_1(k^w) < V - \bar{w}/2 = TS(K) - \Pi_1(K)$, contradicting $TS(k^f) - \Pi_1(k^f) \leq TS(k^w) - \Pi_1(k^w)$. Therefore $k^f \leq k^w$ for any choice of maximizers. □

Proof of Lemma 7. Fix $s > 0$ and let $\mu \uparrow 1$. By Lemma 6, $\Pi_1(k)$ is decreasing for $k < \hat{k}$, so its gain over K is at most $\Pi_1(0) - \Pi_1(K) = \mu b(0) + (1-\mu)$. By Lemma A.1, $\tilde{b} < (1-\mu)/\mu$. Since $L = 1 + s$ at $k = 0$, it follows that $b(0) < (1+s)(1-\mu)/\mu$, so this gain converges to zero. By contrast, since s is fixed, $\bar{k} < K$ is fixed and every $k < \hat{k} \leq \bar{k}$ satisfies $c(k) - c(K) \geq c(\bar{k}) - c(K) > 0$. Thus K gives strictly higher net profit than every $k < \hat{k}$ for all μ sufficiently close to one. Every $k \in [\hat{k}, K)$ also gives lower net profit than K, so $k^f = K$. Proposition 6 then gives $k^w = K$, while $k^c = \hat{k} < K$.

Now fix $\mu \in (0,1)$ and let $s \to \infty$. For $k < \hat{k}$, $\Pi_1'(k) = \mu[b'(k) - (1-\mu)(1+s)^2] < -\mu(1-\mu)(1+s)^2$. Since c' is bounded on $[0, K]$, $\tilde{\Pi}_1'(k) = \Pi_1'(k) - c'(k) < 0$ for every $k < \hat{k}$ when s is sufficiently large. Hence $k = 0$ uniquely maximizes net profit below $\hat{k}$.

It remains to compare $k = 0$ with K. At $k = 0$, $\tilde{\beta} = s/(1+s) \to 1$. Since the range of $\tilde{y} - \tilde{b}$ contains 1 in its interior by Lemma A.1, $\tilde{x}$ converges to an interior value and $\tilde{b}$ to a strictly positive value. Therefore $b(0) = (1+s)\tilde{b} \to \infty$, and $\Pi_1(0) - \Pi_1(K) = \mu b(0) + (1-\mu)$ eventually exceeds $c(0) - c(K)$. Since every $k \in [\hat{k}, K)$ is dominated by K, Firm 1 uniquely chooses $k^f = 0$.

Finally, by Proposition 3, $\mu^*(s) \to 0$ as $s \to \infty$. Hence, for the fixed $\mu > 0$, we have $\mu > \mu^*(s)$ for all sufficiently large s, so total surplus is maximized at $k = \hat{k}$. Total surplus is constant on $[\hat{k}, K]$, while $c(k)$ is strictly decreasing. Therefore net total surplus is uniquely maximized at $k^w = K$. It is clear that $k^c = \hat{k}$, completing the proof. □

A.5 General Match Values and Flexible Learning

Proof of Proposition 7.

Step 1: Consumer demand. Under a two-signal experiment, let q be the probability of purchase from Firm 1. For $q \in (0,1)$, denote the two posterior means by m_1 and m_2, and

write $d = m_1 - m_2 \geq 0$. Since $qm_1 + (1-q)m_2 = \bar{w}$, we have $m_1 = \bar{w} + (1-q)d$ and $m_2 = \bar{w} - qd$. The variance of posterior means is $q(1-q)d^2$. At $q = 0$ or $q = 1$, we choose no learning and set $d = 0$.

We first determine which pairs (q, d) can be generated under F. Let F^{-1} be a quantile function and define $T_F(q) = \int_{1-q}^{1} [F^{-1}(u) - \bar{w}]\, du$, with $T_F(0) = T_F(1) = 0$. If $r(w)$ is the conditional probability of signal 1, then $\mathbb{E}_F r(w) = q$ and $\mathbb{E}_F[(w - \bar{w})r(w)] = q(1-q)d$.

Choose t such that $\Pr(w > t) \leq q \leq \Pr(w \geq t)$. Let r_q be the rule that sends signal 1 when $w > t$ and signal 2 when $w < t$, randomizing at $w = t$ so that signal 1 has probability q. For any other rule r with the same signal probability, $(w - t)[r(w) - r_q(w)] \leq 0$. Since $\mathbb{E}_F(r - r_q) = 0$,

$$q(1-q)d = \mathbb{E}_F[(w - \bar{w})r(w)] \leq \mathbb{E}_F[(w - \bar{w})r_q(w)] = T_F(q).$$

Conversely, taking a weighted average of $r_q(w)$ and the constant rule $r(w) = q$ preserves the signal probability and gives every value between zero and $T_F(q)$ for $\mathbb{E}_F[(w - \bar{w})r(w)]$. Thus the two posterior means are feasible exactly when $q(1-q)d \leq T_F(q)$.

Relative to buying from Firm 2 without information, the consumer's expected purchase payoff is $q(\bar{w} - \Delta) + q(1-q)d$. For now, ignore whether each signal leads to its intended purchase. After subtracting the information cost, the following problem gives an upper bound on the consumer's value in (7):

$$\max_{\substack{q \in [0,1],\ d \geq 0 \\ q(1-q)d \leq T_F(q)}} \left\{ q(\bar{w} - \Delta) + q(1-q)\left(d - \frac{k}{2}d^2\right) \right\}. \tag{26}$$

We will check that the optimal signals lead to the intended purchases.

For fixed $q \in (0,1)$, the optimal difference is $d = \min\{1/k, T_F(q)/[q(1-q)]\}$. Define $k_F = \left(\inf_{0<q<1} T_F(q)/[q(1-q)]\right)^{-1}$. Since F^{-1} is nondecreasing, T_F is continuous and concave. The upper half of the prior has mean strictly above $\bar{w}$ because F is nondegenerate, so $T_F(1/2) > 0$. Concavity and $T_F(0) = T_F(1) = 0$ therefore give $T_F(q) \geq 2T_F(1/2)\min\{q, 1-q\}$, and hence

$$0 < 2T_F(1/2) \leq \inf_{0<q<1} \frac{T_F(q)}{q(1-q)} \leq 4T_F(1/2) < \infty.$$

Thus $0 < k_F < \infty$, and $d = 1/k$ is feasible for every $q \in (0,1)$ whenever $k \geq k_F$.

For such k, substituting $d = 1/k$ into (26) gives the strictly concave objective $q(\bar{w} - \Delta) + q(1-q)/(2k)$. Its unique maximizer is the demand in (8), with $L = 1/k$. When $0 < q < 1$, the first-order condition gives $\Delta = \bar{w} + (1/2 - q)/k$, so $m_1 - \Delta = \Delta - m_2 = 1/(2k) > 0$. Both purchases are therefore strictly optimal. If $q = 1$, then $\Delta \leq \bar{w} - L/2 < \bar{w}$, and the consumer strictly prefers Firm 1 without learning. If $q = 0$, then $\Delta \geq \bar{w} + L/2 > \bar{w}$, and she strictly prefers Firm 2. Thus the bound is attained at every price pair.

Demand is unique because pooling preserves purchase probabilities and the objective

above has a unique maximizer. Moreover, when $k > 0$, an optimal experiment cannot have different posterior means among signals leading to the same purchase: pooling them would strictly lower its cost without changing the purchase payoff. The unique q and, when $0 < q < 1$, the unique $d = 1/k$ therefore determine the optimal distribution of posterior means. The experiment generating this distribution need not be unique.

Step 2: Equilibrium prices. Firm 1 can earn positive profit by charging a sufficiently small positive price, regardless of Firm 2's price. Its equilibrium profit is therefore positive.

Suppose first that both firms earn positive profits. Both prices are positive, and both demands lie strictly between zero and one. Neither firm charges $\bar{p}$, since its profit derivative there is $D_i - \bar{p}/L < 0$, using $\bar{p} > 1/k_F \geq L$. The first-order conditions are $p_i = LD_i$ for $i = 1, 2$. Together with (8), they give $p_1^* = L/2 + \bar{w}/3$ and $p_2^* = L/2 - \bar{w}/3$. Both prices and demands are positive exactly when $\bar{w} < 3L/2$, or equivalently $k < k^\dagger$. It follows that $\bar{D}_i = p_i^*/L$ and $\Pi_i = (p_i^*)^2/L$.

Suppose instead that Firm 2 earns zero profit. Then $p_1 \leq \bar{w} - L/2$; otherwise Firm 2 could charge a small positive price and obtain positive demand. Firm 1 therefore serves the entire market. If $p_1 < \bar{w} - L/2$ or $p_2 > 0$, it can raise its price slightly without losing demand. Thus $p_1 = \bar{w} - L/2$ and $p_2 = 0$. At these prices, the right derivative of Firm 1's profit is $3/2 - \bar{w}/L$. Hence this price pair can be an equilibrium only when $\bar{w} \geq 3L/2$, or $k \geq k^\dagger$.

It remains to check that these prices are best responses. Against a fixed rival price, profit rises linearly while demand is one and becomes a strictly concave quadratic when demand lies between zero and one. The derivative of profit falls at the point where demand first drops below one. Thus the conditions above rule out profitable deviations to any price with positive demand. Any price that yields zero demand cannot be a profitable deviation. In the second case, Firm 2 receives zero demand at every nonnegative price, so $p_2 = 0$ is also a best response.

The prices are feasible: when both firms are active, $p_1^* < L < \bar{p}$ and $p_2^* \leq L/2$. In the quasi-monopoly case, $p_1^* < \bar{w} < \bar{p}$. These are the only possible pure equilibrium price pairs, and each is an equilibrium in its stated range.

Step 3: Welfare. In the interior equilibrium, with X defined as in the text, $\mathrm{Cov}(X, w) = q(m_1 - \bar{w}) = q(1 - q)d$. Expected product value is therefore $V + (q - 1/2)\bar{w} + q(1 - q)d$. Subtracting the information cost and using $d = L = 1/k$ gives

$$TS = V + \left(q - \frac{1}{2}\right)\bar{w} + \frac{L}{2}q(1 - q).$$

At $q = \bar{D}_1 = 1/2 + \bar{w}/(3L)$, the allocation term is $\bar{w}^2/(3L)$ and the net information term is $L/8 - \bar{w}^2/(18L)$. Their sum gives the stated total surplus. Industry profit is $L/2 + 2\bar{w}^2/(9L)$, and subtracting it gives $CS = V - 3L/8 + \bar{w}^2/(18L)$.

When only Firm 1 sells, the consumer does not learn and pays $\bar{w}-L/2$. Hence $TS = V+\bar{w}/2$ and $CS = V - \bar{w}/2 + L/2$. The equilibrium prices and the optimal consumer choices after every price pair form a subgame perfect equilibrium. □

Proof of Lemma 8. For $w \sim U[\bar{w}-1, \bar{w}+1]$, direct integration gives $T_F(q) = q(1-q)$. Thus $k_F = 1$, and the feasible differences between posterior means are $0 \le d \le 1$ for $q \in (0,1)$. Maximizing $d - kd^2/2$ over this interval gives $d = 1$ for $0 \le k \le 1$ and $d = 1/k$ for $k > 1$. Define

$$L(k) = 2d - kd^2 = \begin{cases} 2-k, & 0 \le k \le 1, \\ 1/k, & k > 1. \end{cases}$$

This function is positive, continuously differentiable, and strictly decreasing.

Substitution into (26) gives the strictly concave objective $q(\bar{w}-\Delta) + L(k)q(1-q)/2$. Its maximizer is (8) with $L = L(k)$. When $0 < q < 1$, the first-order condition gives $\Delta = \bar{w} + (1/2 - q)L$. Since $L = 2d - kd^2$,

$$m_1 - \Delta = \frac{qL + (1-q)kd^2}{2} > 0 \quad \text{and} \quad \Delta - m_2 = \frac{(1-q)L + qkd^2}{2} > 0.$$

At $q = 0$ or $q = 1$, no learning and the corresponding purchase are optimal, as in the preceding proof. Thus demand is uniquely given by (8) for every $k \ge 0$.

At $k = 0$, an optimal experiment simply reveals whether $w \ge \Delta$. Full revelation leads to the same purchase. Because information is free, every optimal experiment must attain the same purchase payoff as under full information. Thus all induce the same allocation except at $w = \Delta$, which has probability zero.

The equilibrium price analysis above requires only $\bar{p} > \max\{\bar{w}, L(k)\}$, which follows from $\bar{p} > 2$, $\bar{w} < 1$, and $L(k) \le 2$. It therefore gives the prices, demands, and profits in Proposition 7 for every $k \ge 0$. For $k \le 1$, we have $L(k) \ge 1 > \bar{w}$, and both firms earn positive profits. For $k > 1$ and $\bar{w} > 0$, the condition $\bar{w} \ge 3L(k)/2$ is equivalent to $k \ge 3/(2\bar{w})$. Thus the cutoff remains $k^\dagger = 3/(2\bar{w})$. When $\bar{w} = 0$, both firms remain active at every k.

Total surplus depends on the information choice through $d - kd^2/2$. Since $d - kd^2/2 = L(k)/2$, the same welfare calculation applies with $L = L(k)$. □

Proof of Lemma 9. Part (i). This part is proved in the text.

Part (ii). In the interior equilibrium, differentiating total surplus gives

$$\frac{dTS}{dk} = L'(k)\left(\frac{1}{8} - \frac{5\bar{w}^2}{18L(k)^2}\right),$$

where $L'(k) < 0$. If $\bar{w} = 0$, total surplus strictly decreases for every k, so it is uniquely maximized at $k = 0$.

Suppose $\bar{w} > 0$. As k increases from zero to $k^\dagger$, $L(k)$ decreases from 2 to $2\bar{w}/3$. Since $2\sqrt{5}\bar{w}/3$ lies strictly between these values, total surplus first decreases and then increases, with its minimum at $L(k) = 2\sqrt{5}\bar{w}/3$. It is constant for $k \geq k^\dagger$. Therefore only $k = 0$ and $k \geq k^\dagger$ can maximize total surplus. Comparing them gives

$$TS(0) - TS(k^\dagger) = \frac{1}{4} + \frac{5\bar{w}^2}{36} - \frac{\bar{w}}{2} = \frac{(3-\bar{w})(3-5\bar{w})}{36}.$$

Since $0 < \bar{w} < 1$, the difference is positive when $\bar{w} < 3/5$ and negative when $\bar{w} > 3/5$. At $\bar{w} = 3/5$, both $k = 0$ and every $k \geq k^\dagger$ attain the maximum.

Part (iii). Suppose $\bar{w} > 0$. For $k < k^\dagger$, differentiating consumer surplus gives $dCS/dk = L'(k)[-3/8 - \bar{w}^2/(18L(k)^2)] > 0$. For $k \geq k^\dagger$, consumer surplus is $V - \bar{w}/2 + 1/(2k)$, which strictly decreases with k. Hence, consumer surplus is uniquely maximized at $k = k^\dagger$. □

Appendix B Monopoly Benchmark

In this section, we consider two monopoly benchmarks. In the first, as in our main model, the consumer observes the price before deciding whether to learn. In the second, following Ravid, Roesler, and Szentes (2022), she does not. We compare their efficiency implications as information costs vanish.

B.1 Learning After Pricing

A monopolist with zero marginal cost sells a single product. If the consumer buys the product at price p, her payoff is $v - p$, where $v \in \{-1, 1\}$ and $\Pr(v = 1) = \mu \in (1/2, 1)$.[43] If she does not buy, she receives zero payoff.[44]

The monopolist first posts a price $p \geq 0$. After observing the price, the consumer decides whether to learn v. We assume that the cost of learning is proportional to the uncertainty about v: $k \operatorname{Var}(v) = 4k\mu(1-\mu)$, where $k \geq 0$.[45] The state is then realized. If the consumer has chosen to learn, she observes v; otherwise, she remains uninformed. She then decides whether to buy. We use subgame perfect equilibrium. When the consumer is indifferent between learning and remaining uninformed, she learns; when she is indifferent between buying and not buying, she buys.

Without learning, the consumer obtains $U^{NL}(p) = \max\{2\mu - 1 - p, 0\}$. If she learns, then for $p \leq 1$ she buys only when $v = 1$, and her expected payoff is $U^L(p) = \mu(1-p) - 4k\mu(1-\mu)$.

[43] When $\mu \leq 1/2$, an uninformed consumer does not purchase at any price. This case does not introduce a new force for the analysis.

[44] A negative valuation means the product is a poor match, and not buying is efficient. The analysis extends to any $v_L < 0 < v_H$.

[45] We use this specification for consistency with the main model. The same qualitative results hold if the learning cost is simply k.

For $k < 1/2$, comparing these payoffs shows that the consumer learns when

$$\max\{0, 4k\mu - 1\} \le p \le 1 - 4k(1-\mu).$$

The highest price that induces learning is therefore $1 - 4k(1-\mu)$, which yields profit $\mu[1 - 4k(1-\mu)]$. If the monopolist instead induces purchase without learning, it can earn at most $4k\mu - 1$, whenever $4k\mu - 1 > 0$. Define

$$k^M := \frac{1+\mu}{4\mu(2-\mu)}.$$

Since $\mu[1 - 4k(1-\mu)] > 4k\mu - 1$ if and only if $k < k^M$, we obtain the following result.

Proposition B.1. *For $0 \le k < k^M$, the equilibrium price is $p^* = 1 - 4k(1-\mu)$, the consumer learns, and trade occurs if and only if $v = 1$. Total surplus is $\mu - 4k\mu(1-\mu)$, which converges to the first-best level μ as $k \downarrow 0$.*

Hence, when $k < k^M$, the monopolist charges the highest price at which the consumer is willing to learn, rather than lowering the price enough to sell to an uninformed buyer. Learning prevents trade in the low-value state, and total surplus approaches the first-best level as information costs vanish.

B.2 Learning Before Pricing

We solve the monopoly model of Ravid, Roesler, and Szentes (2022) in our binary environment. The consumer first chooses whether to learn and observes v if she learns. The monopolist then sets its price without observing her learning choice or the realization of v, and the consumer decides whether to buy. We use perfect Bayesian equilibrium. An uninformed consumer uses the prior, and she buys when indifferent at the purchase stage.

Let λ denote the probability that the consumer learns. The relevant prices are $p_L = 2\mu - 1$ and $p_H = 1$. At p_L, an uninformed consumer purchases, while an informed consumer purchases only in the high state. At p_H, only an informed consumer in the high state purchases. The monopolist's profits are therefore $\pi_L = (2\mu - 1)[1 - \lambda(1-\mu)]$ and $\pi_H = \lambda\mu$.

Equating these profits gives

$$\lambda^*(\mu) = \frac{2\mu - 1}{4\mu - 2\mu^2 - 1} \in (0, 1).$$

Let q denote the probability that the monopolist charges p_H. The consumer's expected gain from learning is $(1-q)2\mu(1-\mu) - 4k\mu(1-\mu)$. Hence indifference requires $q = 1 - 2k$. For $0 < k < 1/2$, the mixed equilibrium is therefore $\lambda = \lambda^*(\mu)$ and $q = 1 - 2k$.

As $k \downarrow 0$, the probability of p_H converges to one, while the learning probability remains $\lambda^*(\mu) < 1$. Since information costs vanish, the total surplus converges to $\lambda^*(\mu)\mu < \mu$.

At $k = 0$, charging $p_H = 1$ and learning with any probability $\lambda \in [\lambda^*(\mu), 1]$ form an equilibrium. At this price, the consumer is indifferent between learning and being uninformed, while the monopolist prefers p_H whenever $\lambda \geq \lambda^*(\mu)$. Thus a full-learning equilibrium attains the first best at zero information cost, while the positive-cost equilibria converge to the lowest-learning equilibrium of the zero-cost game, with learning probability $\lambda^*(\mu)$. This limiting inefficiency reflects an informational hold-up problem: anticipating that the monopolist can appropriate part of the value of learning, the consumer underinvests in information acquisition.

Appendix C Equilibrium when $k \geq \bar{k}$

In this section, we characterize equilibrium when $k \geq \bar{k}$. In this case, no learning is optimal at every price profile. If $k > \bar{k}$, it is strictly optimal. If $k = \bar{k}$, the consumer is indifferent between learning and no learning only at $\Delta = \bar{w}$. Under no learning, she is also indifferent between the two firms at $\Delta = \bar{w}$. Accordingly, at $\Delta = \bar{w}$, any $D_1(\bar{w}) \in [0, 1]$ is admissible.

Proposition C.1. *Suppose that $k \geq \bar{k}$. Under any admissible tie-breaking rule with $D_1(\bar{w}) \in [0, 1]$, the following hold.*

1. *If $\bar{w} > 0$, every equilibrium has $p_1 = \bar{w}$ almost surely, and the consumer chooses no learning and purchases from Firm 1 almost surely. Hence*

$$(\Pi_1, \Pi_2) = (\bar{w}, 0), \qquad TS = V + \frac{\bar{w}}{2}, \qquad CS = V - \frac{\bar{w}}{2}.$$

   There exists an equilibrium in which Firm 1 sets $p_1 = \bar{w}$ and Firm 2 mixes on $[0, \bar{w}]$ according to

$$F_2(p) = \frac{p}{\bar{w} + p}, \qquad p \in [0, \bar{w}),$$

   with $F_2(\bar{w}) = 1$.

2. *If $\bar{w} = 0$, every equilibrium has $p_1 = p_2 = 0$ almost surely, and*

$$(\Pi_1, \Pi_2) = (0, 0), \qquad CS = TS = V.$$

   The pure price profile $(p_1, p_2) = (0, 0)$ is an equilibrium. If $k > \bar{k}$, no learning is strictly optimal, but the consumer may choose either firm at $\Delta = 0$. If $k = \bar{k}$, she is also indifferent between learning and no learning at $\Delta = 0$.

Proof. For $k > \bar{k}$, no learning is strictly optimal for every price gap. At $k = \bar{k}$, we have $\alpha = \beta = \bar{w}$, so no learning is strictly optimal for every $\Delta \neq \bar{w}$, while the consumer is indifferent between learning and no learning at $\Delta = \bar{w}$. Thus Firm 1 receives demand one when $\Delta < \bar{w}$, zero when $\Delta > \bar{w}$, and demand $D_1(\bar{w}) \in [0, 1]$ when $\Delta = \bar{w}$.

We first show that $\Pi_2 = 0$ in every equilibrium. Suppose instead that $\Pi_2 > 0$. Firm 2 must then charge a positive price with positive probability, so there exists $q > 0$ such that $\Pr(p_2 > q) > 0$. If Firm 1 charges $p_1 = q/2$, then $p_1 - p_2 < 0 \le \bar{w}$ whenever $p_2 > q$, and Firm 1 receives demand one on this event. Hence $\Pi_1 > 0$.

Let M_1 be the upper endpoint of the support of $p_1 - \bar{w}$ and M_2 that of p_2. These endpoints are maxima because the price supports are compact. Suppose that $M_1 > M_2$. Since M_1 belongs to the support, Firm 1 assigns positive probability to prices satisfying $p_1 - \bar{w} > M_2$. For every such price and every p_2 in Firm 2's support, $p_1 - p_2 > \bar{w}$, so Firm 1 receives zero demand and hence zero profit. This contradicts optimality because $\Pi_1 > 0$. The same argument rules out $M_2 > M_1$. Hence $M_1 = M_2$; denote their common value by M. Since $\Pi_2 > 0$, we have $M > 0$.

The distributions of $p_1 - \bar{w}$ and p_2 must both have an atom at M. Suppose that Firm 2 has no atom at M. As in Step 1 of the proof of Proposition 1, there is a sequence of Firm 1 pure best responses p_1^n such that $p_1^n - \bar{w} \to M$. The sale probability at p_1^n is at most $\Pr(p_2 \ge p_1^n - \bar{w})$, which converges to zero because M is the maximum of Firm 2's support and Firm 2 has no atom at M. Since prices are bounded, the profits from these best responses converge to zero, contradicting $\Pi_1 > 0$. The same argument shows that the distribution of $p_1 - \bar{w}$ must also have an atom at M.

Let $\gamma_i > 0$ denote Firm i's atom size at M. Since these prices are used with positive probability, $\Pi_1 = \pi_1(M + \bar{w}) = (M + \bar{w})D_1(\bar{w})\gamma_2$ and $\Pi_2 = \pi_2(M) = M[1 - D_1(\bar{w})]\gamma_1$. Hence $\Pi_2 > 0$ implies $D_1(\bar{w}) < 1$. If Firm 1 lowers its price from $M + \bar{w}$ to $M + \bar{w} - \varepsilon$, then whenever Firm 2 charges M we have $\Delta < \bar{w}$, and Firm 1 receives demand one. For sufficiently small $\varepsilon > 0$, this deviation yields at least $(M + \bar{w} - \varepsilon)\gamma_2 > (M + \bar{w})D_1(\bar{w})\gamma_2 = \Pi_1$, a contradiction. Therefore $\Pi_2 = 0$.

We next show that $p_1 \le \bar{w}$ almost surely. Otherwise, there would exist $q > 0$ such that $\Pr(p_1 > \bar{w} + q) > 0$. If Firm 2 charges $p_2 = q/2$, then $p_1 - p_2 > \bar{w}$ on this event, so Firm 2 receives demand one and earns strictly positive profit, contradicting $\Pi_2 = 0$.

Suppose first that $\bar{w} > 0$. For every $\varepsilon \in (0, \bar{w})$, Firm 1 can charge $p_1 = \bar{w} - \varepsilon$ and sell for sure. Letting $\varepsilon \downarrow 0$ gives $\Pi_1 \ge \bar{w}$. Since $p_1 \le \bar{w}$ almost surely,

$$\bar{w} \le \Pi_1 = \mathbb{E}[p_1 D_1(p_1 - p_2)] \le \mathbb{E}[p_1] \le \bar{w}.$$

Thus $\Pi_1 = \bar{w}$ and $\mathbb{E}[p_1] = \bar{w}$, which imply $p_1 = \bar{w}$ almost surely. Since $\bar{w} > 0$, $\mathbb{E}[p_1 D_1(p_1 - p_2)] = \bar{w}$ further implies $D_1(p_1 - p_2) = 1$ almost surely.

Whenever $p_2 > 0$, we have $p_1 - p_2 = \bar{w} - p_2 < \bar{w}$, so the consumer strictly chooses no learning and purchases from Firm 1. If $p_2 = 0$ occurs with positive probability, then $D_1(\bar{w}) = 1$. For $k > \bar{k}$, no learning is strictly optimal. At $k = \bar{k}$, learning gives Firm 1 demand $\mu < 1$, so demand one again requires the consumer to choose no learning and purchase from Firm 1. Thus the consumer chooses no learning and purchases from Firm 1 almost surely. No information cost

is incurred and $TS = V + \bar{w}/2$. Since $(\Pi_1, \Pi_2) = (\bar{w}, 0)$, it follows that $CS = V - \bar{w}/2$.

It remains to establish existence. Let Firm 1 set $p_1 = \bar{w}$ and let Firm 2 use the CDF stated in the proposition. This CDF is strictly increasing on $[0, \bar{w})$, satisfies $F_2(0) = 0$ and $F_2(\bar{w}^-) = 1/2$, and hence has an atom of size $1/2$ at $\bar{w}$. Given $p_1 = \bar{w}$, every $p_2 > 0$ gives Firm 2 zero demand, while $p_2 = 0$ gives zero profit. Thus F_2 is a best response.

For Firm 1, every price below $\bar{w}$ yields profit below $\bar{w}$. If $p_1 = \bar{w} + t$ for $t \in [0, \bar{w})$, then F_2 has no atom at t, so $\pi_1(\bar{w} + t) = (\bar{w} + t)[1 - F_2(t)] = \bar{w}$. At $p_1 = 2\bar{w}$, Firm 1 sells only against the atom at $p_2 = \bar{w}$, which gives $\pi_1(2\bar{w}) = \bar{w} D_1(\bar{w}) \leq \bar{w}$. Every price above $2\bar{w}$ receives zero demand. Hence $p_1 = \bar{w}$ is a best response, and the stated strategies form an equilibrium.

Finally, suppose that $\bar{w} = 0$. Since $p_1 \leq \bar{w}$ almost surely and prices are nonnegative, $p_1 = 0$ almost surely and $\Pi_1 = 0$. If Firm 2 charged a positive price with positive probability, there would exist $q > 0$ such that $\Pr(p_2 > q) > 0$. Firm 1 could then charge $p_1 = q/2$ and receive demand one whenever $p_2 > q$, which yields strictly positive profit, a contradiction. Hence $p_2 = 0$ almost surely as well.

The pure price profile $(p_1, p_2) = (0, 0)$ is an equilibrium, since every positive deviation receives zero demand. At these prices, the consumer is indifferent between the two firms under no learning, which gives consumer surplus V. If $k > \bar{k}$, no learning is strictly optimal. If $k = \bar{k}$, she is also indifferent between learning and no learning, which gives the same consumer surplus V. Since both firms earn zero profit, $TS = V$ as well. □

Appendix D Empirical Evidence on Product Inspection

Using e-commerce search data, we examine how consumers' product inspection varies with price differences among displayed products.

D.1 Data and Sample Construction

We use KuaiSearch, an e-commerce search dataset introduced by Li et al. (2026). The dataset is constructed from actual product searches on Kuaishou, a Chinese short-video and live-streaming platform with e-commerce services. It records the products displayed in response to each search and consumers' subsequent clicks and purchases. We use the publicly available Lite version, which contains 555,553 search requests involving 102,086 users and more than 6.6 million products.

We use two components of the data. In the ranking file, each observation corresponds to a product displayed in a search request and reports its posted price and indicators for whether it was clicked or purchased. The item file reports the seller and a three-level product classification for each product. We use the third-level category, denoted L3, which is the

most detailed product category available in the data.

We group product-level records into search requests. A search request consists of all records that share the same user ID, session ID, time index, query, and data-split identifier. A search may display products from several L3 categories. For each search, we identify the L3 category containing the largest number of displayed products and refer to it as the *dominant L3 category*.

We keep searches for which the dominant L3 category accounts for at least 70 percent of all displayed products. We also require at least five products with valid positive prices in the dominant category and at least three distinct sellers in that category. The five-product requirement ensures that the within-search price-difference measure is calculated from a sufficiently large set of products. We impose no upper bound on the number of products. These restrictions yield 217,570 search requests.

D.2 Measures

Let $\mathcal{I}_s$ denote the set of products in the dominant L3 category of search s. Our main outcome is

$$Y_s := \mathbf{1}\left\{\sum_{i \in \mathcal{I}_s} Click_{is} \geq 2\right\}.$$

Thus, Y_s equals one if the consumer clicks at least two products in the dominant category. A single click may simply represent the consumer opening a product that she already intends to purchase. Clicking multiple products provides a more direct, although imperfect, measure of comparison across alternatives.

To measure within-search price differences, let $\mathcal{V}_s \subseteq \mathcal{I}_s$ denote the set of products with valid positive prices and seller identifiers, and define

$$\mathcal{P}_s := \{(i,j) : i,j \in \mathcal{V}_s,\ i < j,\ Seller_i \neq Seller_j\}.$$

Our price-difference measure is

$$D_s := \frac{1}{|\mathcal{P}_s|} \sum_{(i,j) \in \mathcal{P}_s} |\log p_{is} - \log p_{js}|.$$

Thus, D_s is the average absolute log-price difference across all pairs of products sold by different sellers in the dominant category.

D.3 Empirical Specification

We estimate linear probability models. For the continuous specification, we cap D_s at its 99th percentile and standardize the resulting variable to have mean zero and standard deviation

one. Denote this variable by $\tilde{D}_s$. We estimate

$$Y_s = \beta \tilde{D}_s + \alpha_{c(s)} + \lambda_{n(s)} + \rho_{m(s)} + \tau_{J(s)} + \varepsilon_s.$$

Here, $c(s)$ is the dominant L3 category, $n(s)$ is the number of products in that category, $m(s)$ is the number of sellers in that category, and $J(s)$ is the total number of products displayed in the search. The corresponding fixed effects enter separately. The count fixed effects are defined for each exact value; for example, searches with five dominant-category products and searches with six receive different product-count fixed effects. Standard errors are clustered by dominant L3 category.

We also allow the relationship to vary across the distribution of D_s. We rank searches by the raw value of D_s and divide them into five equally sized groups. Let $Q_s \in \{1, \ldots, 5\}$ denote the resulting quintile, where $Q_s = 1$ contains the 20 percent of searches with the smallest within-search price differences. We estimate

$$Y_s = \sum_{q=2}^{5} \beta_q \mathbf{1}\{Q_s = q\} + \alpha_{c(s)} + \lambda_{n(s)} + \rho_{m(s)} + \tau_{J(s)} + \varepsilon_s.$$

The lowest price-difference quintile is the omitted group and β_q measures the difference in the probability of product inspection between quintile q and the lowest quintile, conditional on the included fixed effects.

The regressions exclude 774 searches that are the only observation in at least one of the fixed-effect groups. The resulting estimation sample contains 216,796 searches.

D.4 Regression Results

Table 2 reports the estimates. In column (1), a one-standard-deviation increase in within-search price differences is associated with a 1.66 percentage-point decrease in the probability that the consumer clicks at least two products. The clustered standard error is 0.18 percentage points, and the 95 percent confidence interval is $[-2.00, -1.31]$ percentage points.

Column (2) shows a similar pattern across the distribution of price differences. Relative to searches in the lowest price-difference quintile, the probability of product inspection is 2.17 percentage points lower in the second quintile, 3.76 percentage points lower in the third, 5.22 percentage points lower in the fourth, and 5.74 percentage points lower in the highest quintile. Figure 11 displays these estimates.

These associations are not necessarily causal, since price differences may be correlated with unobserved product characteristics, seller characteristics, or consumer preferences.

Table 2: Relative price differences and product inspection

	(1) Continuous	(2) Quintiles
Price differences, standardized	−1.656***	
	(0.177)	
Second price-difference quintile		−2.174***
		(0.387)
Third price-difference quintile		−3.763***
		(0.514)
Fourth price-difference quintile		−5.220***
		(0.656)
Fifth price-difference quintile		−5.740***
		(0.339)
Dominant-L3 fixed effects	Yes	Yes
Dominant-L3 product-count fixed effects	Yes	Yes
Dominant-L3 seller-count fixed effects	Yes	Yes
Total-displayed-product-count fixed effects	Yes	Yes
Observations	216,796	216,796
Mean dependent variable (%)	31.46	31.46

The dependent variable equals one if the consumer clicks at least two products in the dominant L3 category. The price-difference measure is the average absolute log-price difference across all pairs of products sold by different sellers in that category. The sample includes searches with at least five validly priced products and three sellers in the dominant category; no upper bound is imposed on product count. Column (1) caps the measure at its 99th percentile and standardizes it to have mean zero and standard deviation one. Column (2) divides the raw measure into quintiles; the lowest price-difference quintile is omitted. All specifications are linear probability models. Coefficients and standard errors are reported in percentage points. Standard errors, clustered by dominant L3 category, are in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.10$.

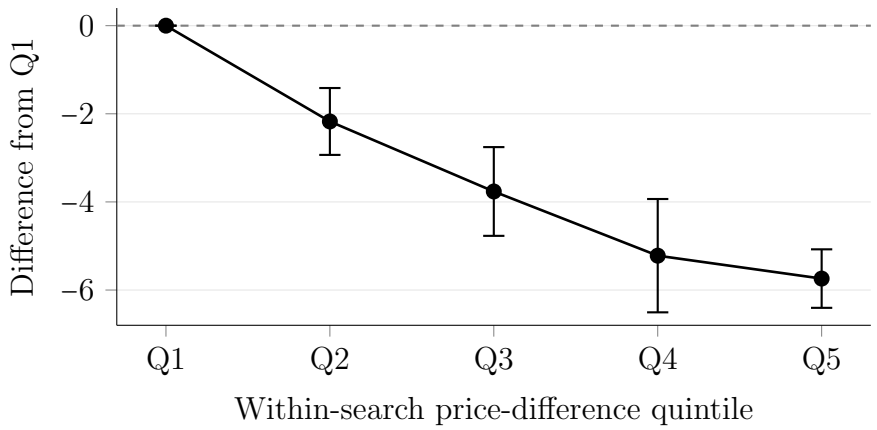


Figure 11: Relative price differences and multiple-product inspection.

Notes: The figure reports differences in the probability of clicking at least two products relative to the lowest within-search price-difference quintile, which is normalized to zero. The regression includes fixed effects for the dominant L3 category, the number of products and sellers in that category, and the total number of products displayed in the search. Vertical bars report 95 percent confidence intervals. Standard errors are clustered by dominant L3 category.