Reviews as Matchmakers: Dynamic Learning of Product Fit

Tommaso Bondi

Cornell University

February 20, 2026

Abstract

Consumer reviews are predominantly conceptualized as quality signals: tools for *sorting* products from best to worst. But they also serve a complementary *routing* function: helping consumers determine not just whether a product is good, but whether it is *good for them*. We model review text as a signal that different readers decode heterogeneously, generating personalized inferences without personalized delivery. As reviews route consumers toward fitting products, the buyer pool homogenizes and the text signal degrades. In equilibrium, matching is therefore partial, sustained by a diversity dividend from mismatched consumers whose reviews cover features that fans' reviews miss. The key constraint is feature overlap between consumer types, which platforms can influence through review design. The model yields implications for platform design: informational barriers to entry, a non-monotone welfare effect of score-following, and information loss from AI summaries that falls disproportionately on minority consumer types. Last, we compare reviews with algorithmic recommenders, the standard tool for personalized routing, and study complementarities between the two.

1 Introduction

A neighborhood has two Thai restaurants. Restaurant A makes fiercely spicy Northern Thai food; Restaurant B makes mild, fusion-influenced Thai. Both are well-run. The difference is style, not quality.

When a review platform launches, both attract a diverse mix of diners. Restaurant A draws "incredible larb, real Isaan heat" (★★★★★) alongside "way too spicy, couldn't finish" (★★). A spice-lover reads these and knows A is her restaurant. A family diner reads the same words and reaches the opposite conclusion: different readers extract different signals from the same text, each correct for her own tastes.

Now suppose the reviews work. Heat-seekers go to A; family diners go to B. Restaurant A's reviews become a fan-only chorus: "best som tum in the city," "never disappoints," all five stars. The two-star reviews from mismatched diners are gone. The reviews are narrower: they cover spice and authenticity but say nothing about parking, noise, or kids' menus. A new diner can tell A is well-liked, but may struggle to determine whether it fits *her*.

Some family diners try A and have bad experiences. Their negative reviews restore the information the system lost. The system converges to a balance: enough *routing* to improve consumer-product *matches*, but not so much that the review base loses the diversity it needs to sustain further routing.[1] This balance is the central object of the paper.

The mechanism extends beyond restaurants. On Goodreads, the most upvoted reviews often position a book rather than score it: "If you loved *Piranesi* for the atmosphere but found *House of Leaves* too demanding, this is your book" is a routing device, not a quality signal. Letterboxd, a popular movie ratings platform, allows its users to dispense with numeric scores altogether: its diary format relies entirely on text; its most popular reviews position films along dimensions like tone, casting, and emotional register rather than ranking them.[2]

Platform design choices shape how much of this fit information reaches consumers. Amazon's "Customers say" AI summary distills hundreds of reviews into a paragraph highlighting what the majority mentions most; for a niche hot sauce, this means "great flavor, arrives well-packaged" while the fit-relevant "extremely hot, not for beginners" disappears. Conversely, programs like Amazon Vine and Yelp's Elite Squad recruit diverse early reviewers for new products, interventions whose value our model quantifies.

[1] We use the two terms interchangeably throughout the paper: routing is the process (directing consumers toward products), matching is the outcome (consumers end up with fitting products).

[2] The same happens on Goodreads, where many reviewers "refuse to reduce a complex book to a simplistic score", instead opting to leave lengthy, nuanced, and highly subjective textual reviews.

The theoretical literature on reviews has primarily studied their *sorting* role: helping consumers identify high-quality products by aggregating dispersed information about a common-value attribute. Learning about *fit*, that is, which product matches which consumer, has been predominantly treated as the domain of recommender systems. There are good reasons for this: from a purely theoretical standpoint, it is not immediate how to model *social* learning about an *idiosyncratic* preference. And yet, the examples above suggest consumers routinely learn about fit from their peers.

We develop a dynamic model with microfounded text accuracy. Each consumer type attends to a different subset of the feature space; the accuracy with which a reader identifies her match depends on how well the reviews cover *her* features. A diverse reviewer pool covers both subsets; a fan-only pool covers only the fans' features. We study a regime where *quantity* of reviews is not the constraint; what matters is the *variety* of the reviewer pool.[3]

The degree of feature overlap between consumer types determines how informative a homogeneous review base is for non-fans. A family diner and a food enthusiast both care about service and cleanliness (shared features), but the family diner also attends to noise level and kids' menus while the enthusiast attends to spice sourcing and ingredient provenance (exclusive features).

Three sets of results emerge. First, review systems cannot sustain perfect matching: the diversity that makes reviews informative is what successful matching eliminates. The system converges to a unique equilibrium with partial matching, characterized by a *diversity equation* that decomposes the mismatch rate into three forces: the information deficit of a fan-only review base, the majority's tendency to self-sort, and the diversity value of a mixed reviewer pool. A positive fraction of mismatched consumers persist in every period, not because the system fails but because their reviews cover features that fans' reviews miss; this is the *diversity dividend*.

Second, the binding constraint on matching is the feature overlap between consumer types. Markets where types attend to similar features (hotels: cleanliness, location, and noise are near-universal) sustain high matching even with homogeneous reviewers; markets with disjoint features (books: literary fiction and thriller readers attend to entirely different dimensions) depend critically on diversity. The feature overlap governs the return on platform investments in reviewer recruitment, structured review templates, and stratified summaries, and we derive a shadow value formula that quantifies the return per marginal underrepresented reviewer.

Third, the equilibrium generates implications for information design. AI-generated review

[3] The dynamics operate on the cumulative share of majority-type buyers in a product's review base. In the large-sample regime, the cumulative share approximately equals the recent flow composition.

summaries erase the diversity dividend by surfacing only the features the majority mentions most. Numerical scores, which we introduce in Section 4, interact with the text channel in a non-monotone way: consumers who follow aggregate ratings rather than reading text can either improve or degrade matching, depending on the demographic composition of the market. The informational barrier to entry that emerges when platforms rank by scores is compositional rather than quality-based, and absent these external forces does not persist: the unique equilibrium is globally stable, so a new product's review base converges to the same steady state as the incumbent's.

Fourth, we study the interaction between reviews and algorithmic recommender systems. The two systems serve a similar purpose – routing consumers to fitting products – but through different channels. We model a setting in which each consumer uses one system or the other for her purchase decision. A well-matched recommender user generates a homogeneous review that degrades the text signal for others. The optimal platform nevertheless deploys both systems, because each fails along dimensions the other covers.

Our contribution. To summarize, we propose a novel theory of review-driven consumer matching that generates four results: (i) an impossibility – perfect matching is inconsistent with the data-generating process it creates; (ii) a diversity dividend – persistent mismatch is welfare-enhancing because it sustains informative text; (iii) an asymmetric distortion – AI summaries and score-based ranking disproportionately harm minority-type consumers through feature-frequency selection; and (iv) a complementarity – reviews and recommender systems fail along different dimensions, making the optimal information architecture a hybrid.

Roadmap. Section 1.1 reviews the literature. Section 2 develops the model. Section 3 characterizes the equilibrium and derives comparative statics. Section 4 introduces numerical scores and analyzes score-following, AI-generated summaries, and informational barriers to entry. Section 5 studies the interaction between reviews and algorithmic recommenders. Section 6 discusses extensions and Section 7 concludes. All proofs are in the Appendix.

1.1 Related Literature

Online reviews have been shown to affect sales (Chevalier and Mayzlin, 2006; Luca, 2016), pricing (Sun, 2012), market structure (Godes and Mayzlin, 2004; Forman et al., 2008), and strategic manipulation (Mayzlin et al., 2014); see Pocchiari et al. (2025) and Bondi and Rossi (2026) for recent surveys. De Langhe et al. (2016) find that consumers treat average ratings as face-value quality signals rather than noisy composites of heterogeneous experiences, even when the rating distribution is visibly polarized. This makes understanding the informational content of reviews (what they actually convey, and to whom) a first-order question for platform design and consumer welfare. Survey evidence confirms consumers value text: 56% of shoppers say that the length, depth, and detail of review content influences their purchase decisions.[4]

A large theoretical literature studies the average rating as a quality signal. Acemoglu et al. (2022) characterize learning speed under endogenous selection where taste differences slow convergence to the true quality. Bondi (2025) shows that, when consumers are naive, reviewer self-selection can advantage polarizing products precisely because fit variation distorts quality signals. Chen et al. (2024) analyze the dual roles of reviews and prices in conveying quality under selection bias. In these models, fit is a complication for quality inference; in ours, fit is the object of learning. The distinction matters for the information mechanism: in quality-learning models, endogenous selection *slows convergence* to a fixed true state (Acemoglu et al., 2022); in our model, successful matching does not slow convergence but *degrades the signal itself*, because the diversity that makes text informative is what matching eliminates. The signal is not noisy – it is shrinking.

Because we study taste heterogeneity, we naturally relate to the literature on *variance*, not just averages, of ratings. Sun (2012) shows that high variance signals horizontal differentiation; Zimmermann et al. (2018) and Bollinger et al. (2023) show that variance need not reflect just taste mismatch, and decompose it into taste and quality components. In our model, score variance is an equilibrium outcome of the routing process.

A separate literature studies the competitive impact of fit information, but this information is not socially learned: it is either exogenous or designer-controlled. Ellison and Ellison (2014) show that improved search raises prices and welfare through better matching; Armstrong and Zhou (2022) find that consumer match information has first-order consequences for market structure; Guo and Zhang (2012) show that consumers' costly deliberation to uncover their fit has first-order effects on product line design. On the supply side, Hummel and Morgan (2014) show how taste aggregation failures produce product flops; our model identifies an informational barrier that is compositional rather than driven by survey design.

The intersection of "learning from reviews" and "fit learning" is recent and predominantly empirical. Wang (2021) finds that reviews mentioning fit context reduce purchase uncertainty more effectively than quality-focused reviews, providing direct evidence for the channel we model. Lei et al. (2022) show experimentally that a few top-ranked text reviews can overturn the influence of average ratings on purchase decisions, and that this swaying effect operates entirely through text content; the result supports our modeling choice of consumers who decode review text rather than follow aggregate scores. Hong and Pavlou (2014) define and

[4] The Ever-Growing Power of Reviews, 2021.

measure fit uncertainty empirically, and Chen et al. (2021) propose methods for separating quality and fit components in rating data.

The closest paper to ours is Fainmesser et al. (2021), who study a Hotelling model in which user-generated content conveys positioning information and show that the type of UGC (average ratings versus detailed reviews) has first-order effects on firm advertising, pricing, and profits. Their central insight, that detailed text reviews enable consumers to assess fit while numerical ratings do not, is shared by our model. The key difference is that in their framework the informational content of each UGC type is exogenous: detailed reviews are simply assumed to be more informative about fit. In our model, this informativeness is an equilibrium object that degrades endogenously as the review system's own success homogenizes the reviewer pool. This endogeneity is what generates the impossibility result, the diversity dividend, and the self-undermining dynamics that are central to our analysis.

Because fit has been primarily the realm of recommender systems rather than reviews, we also relate to that literature. Che and Hörner (2018) show that noise injection can improve information aggregation in quality learning; our non-monotonicity result echoes this in a fit-learning context. Dzyabura and Hauser (2019) show that consumers learn preference weights during search, creating a moving target for recommendation algorithms. Aridor et al. (2024), in a field experiment on MovieLens, find that the informational channel of recommendations dominates the consideration-set channel. Feldman et al. (2019) show that early adopters generate information externalities whose value parallels our diversity dividend. Jerath and Ren (2021) show that consumers rationally allocate attention across favorable and unfavorable product information, and Chakraborty et al. (2024) study endogenous WOM generation; both model selection into reviewing, while our mechanism operates through selection into *buying*. Our paper studies a decentralized text-based mechanism and crucially endogenizes the quality of match information as a function of the matching outcome.

2 Model

2.1 Environment

Two products $j \in \{A, B\}$ have styles $\sigma_A = a$ and $\sigma_B = b$. Each period $t = 1, 2, \ldots$, a cohort of N consumers arrives, with N large.[5] Each consumer has type $\tau \in \{a, b\}$, drawn i.i.d. with $\mathbb{P}(\tau = a) = \lambda \in (1/2, 1)$; type a is the majority. The binary type represents two latent taste clusters: a consumer knows her own taste direction but not which product matches it, and

[5] We normalize the cohort to a unit mass and treat proportions as continuous when deriving equilibrium. The formal stability proof (Proposition 1) uses the discrete stochastic framework.

holds a flat prior: $\mathbb{P}(\sigma_A = a) = \mathbb{P}(\sigma_B = b) = 1/2$.

Assumption 1 (Payoffs)**.** Consumer of type τ purchasing product j receives $u(\tau, \sigma_j) = \mathbf{1}\{\tau = \sigma_j\}$: utility 1 from a match, 0 from a mismatch. Each consumer buys exactly one product. Quality is known and equal across products.

The assumption that quality is known and equal is a key modeling choice. It isolates the fit channel, which is the object of study. It captures settings where quality is established – a Michelin-starred restaurant, a Pulitzer-winning novelist's next book, a sequel from a trusted director – and consumers care primarily about whether the product matches *their* tastes. The entire quality-learning literature (e.g., Acemoglu et al., 2022) studies the complementary case; we study the one it abstracts from. The interaction between quality and fit uncertainty is discussed in Section 6, where we argue that the feedback loop we identify persists when quality varies, because text conveys fit information regardless of quality level.[6]

2.2 Scores

Each buyer observes her payoff (match or mismatch) and posts a public review consisting of a numerical score and text.

Assumption 2 (Scores)**.** Each buyer posts a score $s = \mathbf{1}\{\text{match}\}$. Product j's average score at time t is $\bar{s}_j(t) = m_j(t)/n_j(t)$, where $m_j(t)$ is the number of matched buyers through period t and $n_j(t)$ is the total.

Since $\sigma_A = a$, a buyer of product A is matched if and only if she is type a, so the average score equals the fraction of A's buyers who are type a: $\bar{s}_A(t) = \phi_A(t)$, where $\phi_A(t) \equiv n_a(t)/n_A(t)$ denotes the type-a share of A's buyer pool through period t.

We focus on the text channel: consumers choose which product to buy based on their text signal, without conditioning on aggregate scores. This is the relevant model for a consumer whose goal is fit rather than quality. A consumer who already knows the restaurant is well-run, the author talented, or the director skilled does not need the average rating; she needs to know whether the spice level, the pacing, or the tone matches *her*. In a Yelp survey, 90% of consumers reported trusting reviews with text more than star ratings alone, and 59% said a rating without text should not count as a review.[7] Lei et al. (2022) show experimentally that a few top-ranked text reviews can overturn the influence of average ratings on purchase decisions, and that this swaying effect operates entirely through text:

[6] For example, a Goodreads review describing a novel as "beautifully written but glacially paced" is a fit signal whether the book is a masterpiece or mediocre.

[7] Yelp Trust Survey, 2022.

without written content, consumers are not swayed. The PowerReviews 2023 survey confirms that 56% of shoppers do not trust star ratings alone as much as ratings accompanied by written reviews.[8] Under our equal-quality assumption, scores are artificially clean signals of pool composition; when quality varies (Section 6), they confound quality and fit, making text the primary fit channel. We return to scores explicitly in Section 4, where we study how aggregate ratings interact with the text-based routing mechanism, platform design choices such as score-based ranking, and consideration set formation.

2.3 Text as Feature Coverage

We now microfound the text channel. The standard approach models each review as a public signal with a common interpretation (Acemoglu et al., 2022). We instead model *heterogeneous decoding*: products have features, reviewers describe some of them, and different consumer types attend to different features.

Assumption 3 (Feature Attention)**.** Each product has K features. Type τ attends to a subset $\mathcal{F}_\tau \subset \{1, \dots, K\}$, with $|\mathcal{F}_a| = |\mathcal{F}_b| = L$. The overlap is $\mu = |\mathcal{F}_a \cap \mathcal{F}_b|$, with $0 \leq \mu < L$.

The overlap μ measures how much the types have in common: at $\mu = 0$, they attend to entirely disjoint features; near $\mu = L$, they attend to mostly the same features but disagree on preferred valences. The parameter μ is the key structural primitive governing equilibrium matching, and one that platforms can influence through review design (Section 3.4).

Each feature k carries a *style valence* $v_k \in \{a, b\}$, indicating which consumer type it appeals to.[9] The total K plays no further role beyond the constraint $K \geq 2L - \mu$; only L, μ, and ζ (defined below) affect the dynamics.

Assumption 4 (Review Content and Signals)**.** Each reviewer mentions one feature drawn uniformly from her type's feature set $\mathcal{F}_\tau$. Each mentioned feature, when observed by a reader who attends to it, generates a signal about the product's style with log-likelihood ratio $\ln(\zeta/(1-\zeta))$, where $\zeta \in (1/2, 1)$.

The "one feature per review" assumption generates the coupon-collector structure that makes diversity valuable: the marginal information from a review is highest when it covers a previously uncovered feature, which happens more often when the reviewer pool is diverse. Allowing multiple features per review would accelerate coverage but preserve the core trade-off; the asymmetric-verbosity extension in Section 6 shows that the equilibrium structure is

[8] The Ever-Growing Power of Reviews, 2023.

[9] Style valences are properties of the product ("very spicy"), not the consumer ("likes spice"). All features of a given product share the same underlying style; we assume each covered feature provides a conditionally independent signal about this common style.

unchanged when mismatched reviewers mention $r > 1$ features. The uniform draw from $\mathcal{F}_\tau$ captures the empirical regularity that reviewers write about what they care about, which is precisely what generates the coverage asymmetry: fans cover fan features, and non-fans cover the rest.[10]

A consumer of type τ reads the reviews and extracts information from features that are both mentioned in at least one review and in her attention set $\mathcal{F}_\tau$. Features outside her attention set describe aspects of the product she cannot map to her own preferences. This is *heterogeneous decoding*: the same collection of reviews, read by different types, yields different observed feature sets and therefore different inferences about the product's style.[11]

Heterogeneous decoding draws a distinction between *personalized information* and *personalized decoding of information*. Recommender systems personalize by showing different content to different users; reviews personalize by showing the *same* content that different users read differently. A platform like Goodreads, which displays identical review pages and "Best Books of the Year" lists to every visitor, nevertheless functions as a personalization device: the literary fiction reader and the thriller reader see the same reviews and reach opposite conclusions, each correctly for her own tastes. This means that platforms can contribute to variety, support niches, and serve the long tail even without explicit algorithmic personalization – a point we develop in Section 5.

2.4 Coverage, Accuracy, and the Role of Diversity

Consider product A (style a), which has accumulated n_a reviews from type-a buyers and n_b from type-b buyers. Feature $k \in \mathcal{F}_a$ is covered if at least one reviewer mentioned it. Features exclusive to $\mathcal{F}_a$ can only be covered by type-a reviewers; shared features ($k \in \mathcal{F}_a \cap \mathcal{F}_b$) can be covered by either type. The structure is a coupon-collector problem.[12]

Lemma 1 (Expected Coverage)**.** *The expected number of features in $\mathcal{F}_a$ covered by the reviews is*

$$C_a(n_a, n_b) = (L - \mu)\Big[1 - \big(1 - \tfrac{1}{L}\big)^{n_a}\Big] + \mu\Big[1 - \big(1 - \tfrac{1}{L}\big)^{n_a+n_b}\Big]. \tag{1}$$

By symmetry, $C_b(n_a, n_b) = C_a(n_b, n_a)$.

[10] Strategic review writing would add a signaling game orthogonal to our mechanism. Empirical evidence suggests most reviewers write spontaneously (Wang, 2021).

[11] Could a reviewer simply disclose her type? In practice, types are multidimensional and unlabeled. The Goodreads review "if you loved *Piranesi* for the atmosphere but found *House of Leaves* too demanding" routes by similarity without ever stating a type. Text conveys fit information by describing features rather than declaring types.

[12] The coupon-collector structure gives diminishing returns to same-type reviews, making the first reviews from a new type highly informative.

The first term counts *exclusive* features: those in $\mathcal{F}_a$ but not $\mathcal{F}_b$. Only type-a reviewers contribute to these. The second term counts *shared* features: those in $\mathcal{F}_a \cap \mathcal{F}_b$. Both types contribute, which is why diversity helps even for shared features. This is the key asymmetry: a fan-only review base ($n_b = 0$) provides complete coverage for the fans but only partial coverage for non-fans (only the μ shared features in $\mathcal{F}_b$).

Each covered feature carries a style valence that, when observed by an attending reader, generates a signal with log-likelihood ratio $\ln(\zeta/(1-\zeta))$ in favor of the true style, where $\zeta \in (1/2, 1)$. A reader who observes m covered features accumulates a total log-likelihood ratio of $m \ln(\zeta/(1-\zeta))$, yielding the posterior

$$\gamma(m) = \frac{\zeta^m}{\zeta^m + (1-\zeta)^m}. \tag{2}$$

The formula is strictly increasing and strictly concave in m for $m \geq 0$, with $\gamma(0) = 1/2$ (no information) and $\gamma(m) \to 1$ (perfect identification).[13] We adopt $\gamma(m)$ as the probability that the consumer correctly identifies the product's style, and therefore buys the matching product.

Definition 1 (Aggregate Text Accuracy)**.** The aggregate text accuracy is the population-weighted average: $\Gamma(n_a, n_b) = \lambda\, \gamma(C_a) + (1-\lambda)\, \gamma(C_b)$.[14]

We adopt the population-weighted average as the welfare-relevant measure of text informativeness. The weighting reflects the composition of the *reading* population: in a cohort of new arrivals, a fraction λ are type a and a fraction $1-\lambda$ are type b, each decoding the same review base differently. The aggregate Γ is therefore the expected match probability for a randomly drawn consumer. This is the natural welfare criterion when the platform cares about the average consumer's match quality, and it is the quantity that the equilibrium endogenously determines.

The following result characterizes how text accuracy depends on the primitives.

Lemma 2 (Structural Determinants of Text Accuracy)**.** *Fix total reviews n and let $\phi = n_a/n$ denote the fraction from type-a reviewers.*

(i) Γ *is increasing in n (quantity). For n sufficiently large, Γ is decreasing in ϕ for $\phi > 1/2$ (increasing in variety).*[15]

[13] Concavity follows from $\gamma'(m) = \gamma(m)(1-\gamma(m))\ln(\zeta/(1-\zeta))$, which is maximized at $m = 0$ and decreasing. All formal results depend only on the qualitative properties: γ is increasing, concave, and bounded between 1/2 and 1.

[14] Since γ is concave, Jensen's inequality implies $\mathbb{E}[\gamma(C)] \leq \gamma(\mathbb{E}[C])$, so our formula slightly overstates accuracy. For large n, coverage concentrates and the approximation is tight.

[15] The variety result requires the exponential advantage of the minority type's marginal coverage to dom-

(ii) *As* $n \to \infty$ *with* $\phi \in (0,1)$*:* $\Gamma \to \hat{\gamma} \equiv \gamma(L)$*.*

(iii) *As* $n \to \infty$ *with* $\phi = 1$*:* $\Gamma \to \bar{\gamma} \equiv \lambda\,\gamma(L) + (1-\lambda)\,\gamma(\mu)$*.*

(iv) $\bar{\gamma}$ *is strictly increasing in* μ*,* $\hat{\gamma}$ *is strictly increasing in* L *and* ζ*, and* $\hat{\gamma} > \bar{\gamma}$ *whenever* $\mu < L$*.*

The lemma reveals that the structural parameters (K, L, μ, ζ) reduce to two summary statistics: a ceiling $\hat{\gamma}$ (the best the system can do with a perfectly diverse pool) and a floor $\bar{\gamma}$ (the worst it can do with a completely homogeneous pool). The gap $\hat{\gamma} - \bar{\gamma}$ measures how much reviewer diversity matters. When μ is close to L, the gap is small and diversity is a luxury. When μ is close to 0, the gap is large and diversity is essential.

The binding constraint comes from the floor $\bar{\gamma}$: when types attend to disjoint features ($\mu = 0$), a fan-only review base is nearly useless for non-fans and the floor is low; even a few shared features raise it sharply. This drives a dramatic difference in equilibrium outcomes (Figure 1).

2.5 Reduced Form for the Dynamics

As the review base grows, *quantity* ceases to be a constraint: coverage of any feature saturates. But *variety* remains binding because the state variable ϕ_t tracks the cumulative composition, dominated by recent flows.[16]

We adopt the following maintained specification for text accuracy as a function of the buyer-pool composition $\phi \equiv \phi_A(t)$:

$$\gamma(\phi) = \bar{\gamma} + (\hat{\gamma} - \bar{\gamma}) \cdot 4\phi(1-\phi). \tag{3}$$

The specification is adopted for tractability and is pinned down by three economically motivated requirements. First, $\gamma(1) = \bar{\gamma}$: a fully homogeneous pool achieves only the fan-only baseline. Second, $\gamma(1/2) = \hat{\gamma}$: a perfectly balanced pool achieves the maximum. Third, γ is symmetric around $\phi = 1/2$ and monotonically decreasing on $[1/2, 1]$, since increasing homogeneity reduces the pool's feature coverage regardless of which type dominates. The unique degree-two polynomial satisfying these conditions is (3); the coefficient 4 is pinned down by the boundary values.[17]

inate the population weight ratio $\lambda/(1-\lambda)$. The required n grows with λ; for the large-sample regime we study (Proposition 1 onward), the condition is satisfied.

[16] Along the equilibrium path, the minority flow grows as $(1-\lambda)(1-\gamma^*)$ per period while the majority's grows as $\lambda\gamma^*$, so effective accuracy depends on flow composition even in the large-n limit.

[17] The qualitative results (uniqueness, stability, direction of comparative statics) hold for any continuous, strictly decreasing γ on $(\lambda, 1)$ with $\gamma(\lambda) > 1/2$ and $\gamma(1) < 1$. The quadratic form delivers closed-form

Economically, $4\phi(1-\phi)$ is proportional to the probability that two randomly sampled reviews come from different types, so text accuracy is a linear function of reviewer-pool diversity as measured by this index. The term $\hat{\gamma} - \bar{\gamma}$ is the return to moving from zero diversity to full diversity; $4\phi(1-\phi)$ measures how much of that return the current pool captures.

All formal results from Proposition 1 onward are derived from (3).

Each consumer reads the reviews for both products and buys the one she believes matches her type.[18] The type-a fraction among new A-buyers at time t is

$$\psi(\gamma) = \frac{\lambda\gamma}{\lambda\gamma + (1-\lambda)(1-\gamma)}, \tag{4}$$

with $\psi(1/2) = \lambda$ (random assignment when text is uninformative) and $\psi(1) = 1$ (perfect matching).[19] The cumulative composition updates as

$$\phi_{t+1} = \phi_t + \frac{1}{n_t + 1}\big[\psi(\gamma_t) - \phi_t\big], \tag{5}$$

where n_t is A's total buyers through period t and $\gamma_t = \gamma(\phi_t)$. Expression (4) maps text accuracy to buyer composition; the composite map $F(\phi) = \psi(\gamma(\phi))$ gives the composition of new buyers as a function of the current stock, and a steady state satisfies $F(\phi^*) = \phi^*$.

3 Equilibrium

3.1 Partial Matching and the Diversity Dividend

Proposition 1 (Generic Impossibility). *Let $\gamma : [\lambda, 1] \to (0, 1)$ be any text accuracy function satisfying:*

(i) *γ is continuous on $[\lambda, 1]$ and strictly decreasing on $(\lambda, 1)$;*

(ii) *$\gamma(\lambda) > 1/2$;*

(iii) *$\gamma(1) < 1$.*

comparative statics and the diversity equation (7).

[18] Since the model has two products and a flat prior, each consumer compares her posterior for A against her posterior for B. By symmetry of the text signal across products, the choice is deterministic given the text realization.

[19] The same selection function appears in Acemoglu et al. (2022), but there γ is exogenous. Here, γ depends on buyer-pool composition, creating the feedback loop.

Then the updating map $F = \psi \circ \gamma$ has a unique fixed point $\phi^* \in (\lambda, 1)$. The fixed point is globally stable. Perfect matching ($\phi = 1$) is not a steady state.

The conditions require only that a diverse pool is informative (ii) and that a homogeneous pool cannot sustain perfect matching (iii); both are properties of the information technology, not of the equilibrium. For a platform like Yelp or Amazon, the result says that no amount of review solicitation, algorithmic curation, or interface redesign can eliminate mismatch entirely. Persistent mismatch is not a platform failure but a structural feature of fit learning.

The productive question is not "how to achieve perfect matching" but "how to shift the equilibrium balance" – an actionable question for platform managers. The comparative statics in Section 3.4 show that the answer lies in the feature overlap μ, a structural property of the product category that platforms can influence through review design.

The results that follow vary in their dependence on the quadratic specification (3). The impossibility result (Proposition 1), uniqueness, and global stability hold for *any* text accuracy function satisfying conditions (i)–(iii); they are properties of the feedback loop, not the functional form. The closed-form diversity equation (Proposition 2), the shadow value formula (Proposition 3), and the specific thresholds in Propositions 6 and 8 use the quadratic form for tractability, but the underlying comparative statics – that higher μ improves matching, that summaries erase the diversity dividend, that recommender interaction is non-trivial – hold generically.

The quadratic specification (3) satisfies conditions (i)–(iii) and yields closed-form solutions for the equilibrium. In particular, the unique steady state has $\phi^* \in (\lambda, 1)$, matching is partial ($\gamma^* = \gamma(\phi^*) \in (\bar{\gamma}, \hat{\gamma})$), and under the stochastic dynamics (5) the system converges to ϕ^* from any initial condition.[20]

The contrast with herding models is instructive. In Banerjee (1992) and Bikhchandani et al. (1992), the equilibrium is fragile: a cascade on the wrong action can lock in permanently. Here, the fixed point ϕ^* is unique and globally attracting. The difference is that successful routing is *self-undermining*: it eliminates the diversity that made routing possible, creating a restoring force. If the pool becomes too homogeneous, text accuracy drops, mismatch rises, and diversity is restored. The self-correcting dynamics ensure convergence from any starting point, so the steady state is not an artifact of favorable initial conditions.

The balance point is maintained by what we call the *diversity dividend*: the gap $\gamma^* - \bar{\gamma}$ between the equilibrium match rate and the fan-only baseline. This gap measures the welfare contribution of reviewer-pool diversity sustained by persistent mismatch. At ϕ^*, mismatched consumers are individually worse off but serve a collective function: their reviews cover features in $\mathcal{F}_b$ that fans' reviews miss. On Goodreads, a literary fiction enthusiast who reads

[20] Convergence follows from a standard stochastic approximation argument (Appendix).

a popular thriller and writes "predictable plot but the prose is surprisingly clean" covers a dimension (prose quality) that thriller fans' reviews rarely mention.

The steady state is the composition at which the marginal mismatched buyer's contribution to text quality exactly offsets the matching loss from her mismatch. The diversity dividend is emergent, not planned: no agent deliberately explores, yet the system sustains the variety it needs.

3.2 Characterizing the Equilibrium

The equilibrium condition $\phi^* = \psi(\gamma(\phi^*))$ does not immediately reveal how matching depends on the structural primitives. Define the *mismatch rate* $m^* = 1 - \phi^*$ and the *majority odds ratio* $R = \lambda/(1-\lambda)$.

Proposition 2 (The Diversity Equation)**.** *The equilibrium mismatch rate is characterized by:*

$$\gamma^* \;=\; \psi^{-1}(\phi^*) \;\approx\; 1 - R\cdot m^*, \qquad m^* \;\approx\; \frac{1-\bar{\gamma}}{R+4\Delta}, \tag{6}$$

where $R = \lambda/(1-\lambda)$ and $\Delta = \hat{\gamma} - \bar{\gamma}$. The first expression is exact; the second and third are first-order approximations around $\phi^* = 1$.

We call (6) the *diversity equation.* A closed-form solution does not exist: ψ and γ compose into a rational function whose root cannot be expressed in elementary terms. However, the approximation is precise: with baseline parameters ($L = 10$, $\zeta = 0.56$, $\lambda = 0.55$, $\mu = 3$), the exact equilibrium is $\gamma^* = 0.8549$ while the approximation gives 0.8586, an error of 0.4%.[21]

Three structural features of the product category determine how much mismatch the market sustains:

$$m^* \;\approx\; \frac{\overbrace{1-\bar{\gamma}}^{\text{information deficit}}}{\underbrace{R}_{\text{demographic routing}} \quad + \quad \underbrace{4\Delta}_{\text{diversity value}}}. \tag{7}$$

The *numerator* is the information deficit: the gap between perfect text quality and the fan-only baseline $\bar{\gamma}$. When $\bar{\gamma}$ is close to 1 (high feature overlap, as in hotels), the fan-only review base is nearly sufficient and the equilibrium requires little mismatch. When $\bar{\gamma}$ is low (disjoint features, as in books), the system must sustain a substantial mismatch flow.

The *denominator* sums two compression forces: the majority's self-sorting tendency $R = \lambda/(1-\lambda)$, and the diversity value 4Δ, which converts mismatch into better future routing. Both the numerator and the denominator depend on μ: increasing feature overlap raises $\bar{\gamma}$

[21] Both approximations are first-order expansions in m^*; the error is quadratic in m^*. Exact bounds are $(1-\bar{\gamma})/(R+4\Delta) \le m^* \le 1 - \psi(\bar{\gamma})$, with the lower bound tight for small m^*.

(reducing the numerator) and shrinks Δ (reducing the denominator). But the numerator effect dominates, so m^* is unambiguously decreasing in μ (Proposition 4). To make this concrete: in a book market with low feature overlap, the formula predicts roughly 10% mismatch; in a hotel market with high overlap, mismatch drops below 3%. The formula identifies the highest-priority targets for platform intervention: categories with low feature overlap and balanced demand.

The identity $\gamma^* \approx 1 - Rm^*$ says that match efficiency equals one minus the mismatch rate amplified by the odds ratio. Markets with dominant consumer segments pay a steeper efficiency cost for the same level of mismatch. In the Thai restaurant example with $\lambda = 0.55$ (a moderate spice-lover majority), $R \approx 1.2$: each percentage point of mismatched diners costs 1.2 points of match efficiency. On Goodreads, where literary fiction readers might constitute 70% of a novel's audience ($\lambda = 0.7$, $R \approx 2.3$), the amplification is nearly double: a 7% mismatch rate produces roughly 15% of consumers buying the wrong product.

3.3 Welfare

Aggregate welfare per period is the match efficiency γ^*, bounded by three benchmarks: random assignment (1/2), uninformed score-following (λ), and perfect matching (1).

Corollary 1 (Welfare)**.** *The review system closes a fraction* $(\gamma^* - \lambda)/(1 - \lambda)$ *of the gap between uninformed choice and perfect matching. This fraction is increasing in* μ *and* ζ*, and approaches 1 as* $\bar{\gamma} \to 1$*.*

In practice, the match rate γ^* maps to observable platform KPIs: product return rates and "not for me" complaints (direct measures of $1-\gamma^*$), review helpfulness votes conditional on reviewer type (a proxy for coverage of reader-relevant features), and repeat purchase rates (which increase with match quality). The diversity dividend $\gamma^* - \bar{\gamma}$ predicts the incremental benefit of maintaining a diverse reviewer pool, measurable as the gap in these KPIs between products with heterogeneous vs. homogeneous review bases.

Although each type faces the same per-capita mismatch probability $1 - \gamma^*$, the *informational contribution* of mismatch is not symmetric: product A needs type-b reviewers to cover $\mathcal{F}_b$-exclusive features. Since the minority is a smaller group, each minority consumer's mismatch is more informationally pivotal. Note that "minority" here is product-specific, not demographic: type b is underrepresented on product A because routing has worked, not because of any population-level characteristic. The same type is the majority-match on product B.

To give a sense of magnitudes, we briefly return to our two opening examples. In the Thai restaurant market with $\lambda = 0.55$, $L = 10$, $\mu = 3$ (spice-lovers and family diners share three features: service, cleanliness, portion size), the equilibrium match rate is $\gamma^* \approx 0.86$; about 12% of diners end up at the wrong restaurant. The diversity dividend is $\gamma^* - \bar{\gamma} \approx 0.05$: removing all mismatched reviewers would drop accuracy from 86% to 81%, a five-percentage-point loss. In the Goodreads example with $\lambda = 0.70$, $\mu = 2$ (literary fiction and thriller readers share only "plot" and "pacing"), the equilibrium is similar at $\gamma^* \approx 0.85$, but the mismatch cost is amplified by $R = 2.3$: each mismatched buyer inflicts more than twice the welfare damage. The effects are larger in low-overlap markets. With $\mu = 0$ (no shared features), match accuracy drops to 82%, mismatch rises to 15%, and the diversity dividend reaches nearly ten percentage points (Figure 1).

Corollary 2 (Informational Pivotality)**.** *At the steady state, each minority reviewer's marginal contribution to text accuracy exceeds each majority reviewer's. The pivotality ratio is at least* $R = \lambda/(1-\lambda)$ *and is increasing in* λ*.*

The asymmetry arises because the underrepresented type's reviewer stock on its "off" product is smaller ($n_b^A/n_a^B = 1/R < 1$), and diminishing returns in the coupon-collector process make each additional review more valuable at the margin; the exact pivotality ratio also depends on L, μ, and n through the coverage marginals, but the R lower bound captures the first-order effect. The pattern echoes Waldfogel (2007)'s "tyranny of the majority" in a new setting. In his framework, high fixed costs cause markets to target the majority's preferences, leaving preference minorities underserved. In ours, the mechanism is informational: the review base tilts toward the majority not because of production costs but because successful matching homogenizes the reviewer pool.

Proposition 3 (Shadow Value of an Underrepresented Review)**.** *At the steady state, the marginal welfare gain from one additional type-b review on product A is*

$$\frac{\partial \gamma^*}{\partial n_b^A} = \underbrace{(1-\lambda)}_{\text{minority weight}} \cdot \underbrace{\gamma'(m_b^*)}_{\substack{\text{signal value per} \\ \text{covered feature}}} \cdot \underbrace{\frac{(L-\mu)}{L}\left(1-\frac{1}{L}\right)^{n_b^A}}_{\text{marginal coverage probability}}, \tag{8}$$

where $n_b^A = (1-\lambda)(1-\gamma^*)n$ *is the minority reviewer stock on A and* $m_b^* = \mu + (L-\mu)[1 - (1-1/L)^{n_b^A}]$ *is the expected type-b feature coverage. The shadow value is:*

(i) *strictly decreasing in* μ *(highest in low-overlap markets),*

(ii) *strictly decreasing in* n_b^A *(highest when minority reviews are scarce),*

(iii) *for fixed* L*, strictly increasing in* $L - \mu$ *(highest when the minority attends to many exclusive features).*

The shadow value has a clean coupon-collector structure. The first term is the underrepresented type's population weight. The second is the signal value of one additional covered feature at the current margin. The third is the probability that the marginal review covers a previously uncovered exclusive feature; this is the coupon-collector term, and it declines geometrically as the minority reviewer stock grows. In the Thai restaurant example, a family diner's review of Restaurant A that mentions noise level or parking covers a feature that no spice enthusiast's review would mention. If five such reviews exist and $L = 10$, the probability that the sixth covers a new family-relevant feature is roughly $(7/10)(9/10)^5 \approx 0.41$; by the twentieth, it drops to 0.09.

The first few underrepresented reviews are dramatically more valuable than the last. Platforms running reviewer recruitment programs should target underrepresented consumer segments, not just prolific reviewers: composition matters more than volume, echoing Godes and Mayzlin (2004)'s finding that WOM *dispersion* across communities predicts sales better than WOM volume. The formula also quantifies the hidden cost of helpfulness filters that suppress exactly the reviews with the highest shadow value (Section 6).

3.4 Comparative Statics

The comparative statics connect the equilibrium to the structural primitives of the product category. The key insight is that the binding constraint on matching is baseline text quality $\bar{\gamma}$: the informativeness of a homogeneous review base, which depends primarily on the feature overlap μ.

Proposition 4 (Feature Overlap as the Binding Constraint)**.** *At the interior steady state:*

(i) $\partial\gamma^*/\partial\mu > 0$*: greater feature overlap improves matching.*

(ii) $\partial^2\bar{\gamma}/\partial\mu^2 < 0$*: the returns to overlap are diminishing.*

(iii) *When the pool is sufficiently homogeneous,* $\partial\gamma^*/\partial\bar{\gamma} > \partial\gamma^*/\partial\hat{\gamma}$*: raising the floor matters more than raising the ceiling.*

Proposition 5 (Signal Quality and Matching)**.** *More features per type and stronger per-feature signals also improve matching:* $\partial\gamma^*/\partial L > 0$ *and* $\partial\gamma^*/\partial\zeta > 0$*.*

Increasing μ raises $\bar{\gamma}$, shifting the equilibrium toward better matching; diminishing returns follow from the concavity of $\gamma(m)$. Part (iii) of Proposition 4 says that for most empirically relevant parameter values, the platform's highest-leverage intervention is raising

Table 1: Comparative statics of the equilibrium. Signs indicate the effect of increasing each parameter on the equilibrium outcome. Baseline values: $L = 10$, $\zeta = 0.56$, $\lambda = 0.55$.

Parameter	Interpretation	γ^*	m^*	V_A	Shadow value	$\bar{\lambda}$
μ	Feature overlap between types	+	−	−	−	+
ζ	Per-feature signal accuracy	+	−	−	−	+
L	Features per type	+	−	−	ambig.	ambig.
λ	Majority share	+	−	−	+	n/a
r	Mismatched reviewer verbosity	+	−	−	−	+

Notes: γ^* = match efficiency; $m^* = 1 - \phi^*$ = mismatch rate; $V_A = \phi^*(1 - \phi^*)$ = score variance; shadow value = marginal welfare gain from one underrepresented review (Proposition 3); $\bar{\lambda}$ = threshold above which score-following improves welfare (Proposition 6). Signs for μ, L, and ζ are proved in Proposition 4; signs for λ follow from the diversity equation (7), noting that R and $\bar{\gamma}$ are both increasing in λ while Δ is decreasing, and verified numerically. The effect of λ on $\bar{\lambda}$ is not defined (it *is* the threshold variable). The effect of L on shadow value and $\bar{\lambda}$ is ambiguous because L enters through both $\hat{\gamma}$ and the coupon-collector marginals. The r row refers to the asymmetric-verbosity extension (Section 6).

the floor: making fan-only reviews more informative for non-fans, rather than making diverse reviews even better.[22]

Platforms can influence μ through review design. Structured templates that ask about multiple dimensions (food, service, atmosphere, value) ensure that even a fan reviewer covers features outside her core interest, effectively increasing μ. Prompts like "Who would you *not* recommend this to?" serve the same function. Displaying reviewer profiles helps readers map features to their own preferences, increasing ζ. A fourth lever is *segmented ratings*: TripAdvisor's separate scores for families, couples, and solo travelers help each consumer type bypass the diversity constraint. All four interventions are most valuable in low-μ categories.

Cross-category predictions. The model predicts that platforms should invest most in text-based features (detailed reviews, structured templates, stratified summaries) for low-μ categories, and in score-based features for high-μ categories. This is broadly consistent with observed design choices: Amazon's "Customers say" AI summaries are most useful for commodity products with high feature overlap, but most distortionary for niche products where minority perspectives are critical.

Each lever connects to testable predictions. Structured templates should increase topic entropy in reviews, with the largest gain in low-μ categories. Encouraging *detailed negative reviews* increases the effective verbosity of mismatched reviewers ($r > 1$), reducing the mismatch needed to sustain a given accuracy level.

Remark 1 (Asymmetric Feature Coverage). The symmetric case $|\mathcal{F}_a| = |\mathcal{F}_b| = L$ is adopted

[22] The condition for (iii) is $8\phi^*(1-\phi^*) < 1$, which holds whenever the equilibrium pool is more than about 85% majority-type. This covers the empirically relevant range; the condition fails only when signals are very weak or the majority is barely dominant.

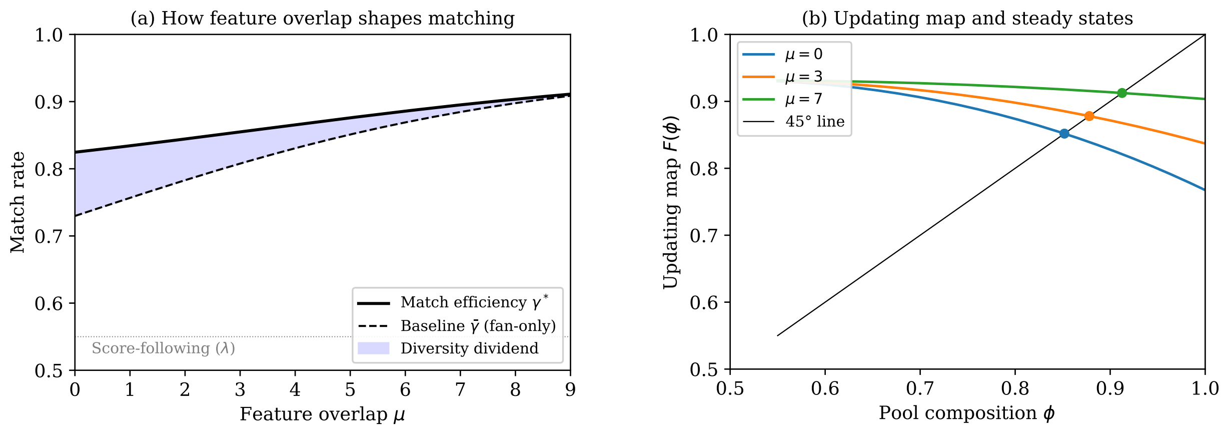

Figure 1: $L = 10$, $\zeta = 0.56$, $\lambda = 0.55$. Panel (a): Equilibrium match efficiency γ^* (solid) and fan-only baseline $\bar{\gamma}$ (dashed) as functions of feature overlap μ. The shaded region is the diversity dividend. At $\mu = 0$, about 15% of buyers are mismatched; removing them would drop accuracy from 82% to 73%. Panel (b): Updating map $F(\phi) = \psi(\gamma(\phi))$ for three values of μ. Each curve crosses the diagonal exactly once; higher μ shifts the steady state toward better matching.

for tractability, but the model extends naturally to the case where one type's feature set is nearly a subset of the other's: $\mathcal{F}_b \subset \mathcal{F}_a$ with $|\mathcal{F}_b| = L' < L$. In this case, type-a reviewers cover most of type-b's features, so the diversity problem is one-sided: type b is informationally "sheltered" by the majority's reviews, while type a bears the full cost of homogenization. The impossibility result still holds ($\bar{\gamma} < 1$ as long as some exclusive features exist), but the welfare loss is concentrated among the type with exclusive features.

The implication for platform design is that the relevant margin is *feature* diversity, not *reviewer* diversity per se. A review base dominated by one type can still be feature-diverse if that type writes expansively about dimensions outside its core interest; the platform's goal should be to elicit broad feature coverage from whatever reviewers it has.

The equilibrium also generates predictions about observable patterns in review data. Although consumers in the baseline model do not condition on scores, the matching equilibrium determines score distributions, which matter for platforms that use scores for ranking and display. Score variance $V_A = \phi^*(1-\phi^*)$ is decreasing in γ^* and hence in μ (by Proposition 4): better-matched products have lower score variance. The model offers a structural interpretation of the empirical findings discussed in Section 1.1: low variance reflects successful *routing*, not consensus about quality, and its level is pinned down by the feature overlap μ.[23] A related pattern appears in Holtz et al. (2020), who find that Spotify's recommendation

[23] Along the convergence path ϕ_t oscillates, so score variance need not decline monotonically. The prediction is cleanest as a cross-sectional comparative static.

algorithm reduces consumption diversity; in our framework, this corresponds to a more homogeneous buyer pool and lower score variance as the recommender routes more consumers to their match.

3.5 Path Dependence and Informational Barriers to Entry

A new product enters at composition $\phi_0 = \lambda$ with average score $\bar{s} = \lambda$. An established competitor at ϕ^* has score $\phi^* > \lambda$. The gap is compositional, not quality-based: the same product generates lower scores simply because its buyer pool has not yet self-selected.[24]

An important distinction arises between the text channel and the score channel. In our baseline model, where consumers read text, the barrier is *transient*: the entrant's diverse early buyer pool is actually more informative for text readers than the incumbent's homogeneous pool, and the dynamics converge to ϕ^* from any initial condition (Proposition 1(ii)). Formally, the stochastic dynamics $\phi_{t+1} = \phi_t + (n_t + 1)^{-1}[F(\phi_t) - \phi_t]$ converge to ϕ^* almost surely from $\phi_0 = \lambda$. The text-based system is self-correcting: poor initial matches generate informative reviews, which improve future matches.

The barrier becomes persistent when platforms rank products by score. Score-based ranking diverts traffic away from the low-scoring entrant, starving it of the buyer flow needed for convergence. The "flop" is then an informational trap created by the interaction between the text-based routing mechanism and the score-based display mechanism.

Corollary 3 (Informational Barrier to Entry)**.** *The score gap between an entrant at $\phi_0 = \lambda$ and an incumbent at ϕ^* is $\phi^* - \lambda$. The gap is increasing in μ: the barrier is strongest in high-overlap markets where the incumbent's fan base is highly homogeneous.*

On Amazon, a new specialty product can be buried in search rankings not because customers dislike it but because it has not yet found its audience. Hummel and Morgan (2014) identify a related mechanism: when firms cannot correctly identify consumer preferences, many products "flop." In our model, the mismatch is endogenous. Programs like Amazon's Vine, which recruit a broad cross-section of reviewers, directly address this.[25]

The barrier differs from Vellodi (2022), where reviews reveal *quality* and the cold start reflects posterior uncertainty; his design response is upper censorship of incumbent reviews.

[24] The score gap can be substantial: with a moderate majority ($\lambda = 0.55$) and low feature overlap, it exceeds one star on a five-star scale.

[25] Yelp's Elite Squad serves a similar function. De-emphasizing scores for products with few reviews also reduces the barrier. Note that the score penalty from a diverse buyer pool interacts with other score-based penalties imposed by platform design – such as rewarding products with high review *volume* – creating a compounding disadvantage for entrants. Vellodi (2022) studies a related design problem where the platform can censor incumbent reviews to reduce the barrier.

In our model, the barrier arises because the entrant's *diverse* buyer pool generates low scores even though its review base is more informative for text readers. The design response targets reviewer diversity rather than review suppression.

4 Scores, Summaries, and Information Design

The baseline model studies consumers who learn about fit from review text alone. In practice, review platforms present both text and numerical scores, and platforms increasingly layer AI-generated summaries on top of the raw review data. This section studies how these additional information channels interact with the text-based routing mechanism.

We first introduce *score-following*: consumers who buy the higher-rated product without reading text. Score-followers are uninformed about fit: they buy the product that the majority prefers, matching at rate λ. They still purchase, experience the product, and write reviews; what differs is only their purchase decision, not their reviewing behavior.[26] This creates two competing effects on the consumers who *do* read text.

Proposition 6 (The Non-Monotone Effect of Score-Following)**.** *If a fraction α of buyers follow average scores rather than reading text, the aggregate match rate is $M(\alpha) = \alpha\lambda + (1-\alpha)\gamma^*(\alpha)$. There exists a threshold $\bar{\lambda}(\mu, \zeta, L) < 1$ such that:*

(i) *When $\lambda < \bar{\lambda}$: score-following reduces welfare at the margin: $dM/d\alpha\big|_{\alpha=0} < 0$.*

(ii) *When $\lambda > \bar{\lambda}$: score-following* improves *welfare at the margin: $dM/d\alpha\big|_{\alpha=0} > 0$.*

The threshold $\bar{\lambda}$ is increasing in μ and in ζ.

Applied to the Thai restaurant example: when the market has a moderate mix (λ is moderate), score-followers match worse than text-readers and degrade matching for everyone. But when spice-lovers overwhelmingly dominate (λ close to 1), the review base is a fan-only chorus and text is nearly useless. Score-followers who wander into Restaurant A despite being family diners diversify the review base and improve the text signal; even uninformed participation helps when the system is starved for diversity.

The tradeoff has two sides. The *direct* effect is signal loss: uninformed consumers match at rate λ rather than γ^*. The *indirect* effect is diversity gain.[27] When λ is moderate, the direct loss dominates; when extreme, the diversity gain swamps it.

[26] A score-follower who buys product A contributes to A's review base with the same text process as any other buyer. The distinction is that her purchase was not guided by text, so her type is drawn from the population (λ probability of being type a) rather than being routed by the text signal.

[27] This connects to Che and Hörner (2018)'s result that noise injection can improve information aggregation, but in their setting the underlying attribute is common-value; in ours, the fit-learning problem is permanent.

A rapidly spreading instantiation is the AI-generated review summary. Amazon's "Customers say" feature distills hundreds of reviews into a paragraph highlighting common themes; Google Maps, Yelp, and TripAdvisor deploy similar tools. By construction, these summaries foreground themes mentioned most frequently, which are those mentioned by the majority. A seller of artisanal hot sauce may receive dozens of rave reviews from spice enthusiasts alongside a handful of complaints about excessive heat; the AI summary will report "great flavor, arrives well-packaged" while erasing the fit-relevant "extremely hot, not for beginners."

Suppose the platform generates a summary by selecting the k most frequently mentioned features ($k \leq L$). Since shared features are mentioned by both types while exclusive features are mentioned only by one, the ranking by frequency places shared features first, majority-exclusive second, and minority-exclusive last.

Proposition 7 (Summarization Erases the Diversity Dividend). *Suppose the platform generates a k-feature summary ($k \leq L$) by selecting the most frequently mentioned features.*

(i) *For $k > \mu$, the summary includes zero features from $\mathcal{F}_b \setminus \mathcal{F}_a$. A type-$b$ summary-reader achieves accuracy $\gamma(\mu)$: the fan-only baseline.*

(ii) *The effective text accuracy for summary-readers is $\Gamma_{\text{sum}}(k) = \lambda \cdot \gamma\big(\min(k, L)\big) + (1 - \lambda) \cdot \gamma\big(\min(k, \mu)\big)$. For $k \geq \mu$, $\Gamma_{\text{sum}}(k) \leq \bar{\gamma}$: the diversity dividend is entirely erased.*

The proposition says the summary selects features that the majority already covers and excludes the minority's exclusive features entirely. The welfare implication follows directly.

Corollary 4 (Welfare Loss from Summaries). *If a fraction α of consumers read only the summary, the steady-state match rate $M(\alpha, k) = \alpha\, \Gamma_{\text{sum}}(k) + (1-\alpha)\gamma^*(\alpha)$ satisfies $dM/d\alpha\big|_{\alpha=0} < 0$: introducing summary-readers strictly reduces welfare at the margin. Moreover, $M(1, k) = \Gamma_{\text{sum}}(k) \leq \bar{\gamma} < \gamma^*$.*

The result is stark. The summary does not merely reduce information; it reduces it *asymmetrically*. The majority type loses little, because the top-k features are predominantly hers. The minority type loses almost everything, because her exclusive features are mentioned least often and summarized last.[28] The diversity dividend exists in the underlying review data, but the summarization algorithm systematically excludes it, echoing Dai et al. (2018)'s finding that aggregation method choice itself affects information content.

[28] For example, with moderate overlap ($\mu/L = 0.3$) and a summary covering 70% of features, the majority summary-reader retains about 84% accuracy while the minority drops to roughly 67%, compared to 92% for both types under full text with a diverse review base.

The natural remedy is *stratified summarization*: selecting the top features from each consumer type's perspective, rather than overall. Amazon's "Select to learn more" interface, which lets readers click on specific themes to see type-specific snippets, is a step in this direction. A *personalized* summary that conditions on the reader's inferred type would restore the diversity dividend entirely: if the platform identifies a reader as likely type b and surfaces $\mathcal{F}_b$-relevant features, the effective accuracy returns to $\gamma(L)$. Amazon and Google are both moving toward personalized summaries, though current implementations remain coarse.[29] The welfare gain from stratification is concentrated in low-μ categories where minority features are most underrepresented.

5 Interaction with Recommender Systems

The previous sections establish that review text achieves personalization through heterogeneous decoding: consumers see the same content but extract different signals, and this is enough to route them toward fitting products. This broadens the scope of platforms as personalization tools and suggests that recommender systems, while valuable, may be less necessary for consumer-product matching than commonly assumed. Even a platform that plainly ranks products – Goodreads' "Best Books of the Year," the same list for every user – need not drive all consumers to the same product, because each reader decodes the accompanying reviews through her own feature set.

In practice, however, consumers are aided by both review text and algorithmic recommendations. A reader browsing Amazon sees five recommended titles in "Customers who bought this also bought. . . " – this shapes her consideration set. She then opens Goodreads or reads Amazon reviews to determine which of those titles actually fits *her*: whether the prose style, the pacing, the thematic ambition match her sensibility. Sometimes the recommendation suffices and the consumer never reads reviews; sometimes she ignores the algorithm entirely and relies on text. The question is how the two systems interact at the market level.

We model pure substitution: each consumer uses one system or the other for her purchase decision. A recommender observes each consumer's type with precision $\rho \in [0, 1]$: with probability ρ, it correctly identifies her type and routes her to the matching product; otherwise it assigns randomly, achieving match rate $(1 + \rho)/2$. The review system outper-

[29] Google's review summaries on Maps began incorporating user-specific emphasis in 2025, e.g., highlighting accessibility features for users who previously searched for wheelchair-accessible venues. Amazon's "Customers like you say" feature, tested in select categories, represents a similar direction. The staggered rollout of AI summaries across platforms and product categories provides a natural experiment: our model predicts that summary adoption should disproportionately increase return rates and "not for me" complaints in low-μ categories, where the erased diversity dividend is largest.

forms the recommender iff $\rho < 2\gamma^* - 1$, a threshold that is increasing in μ, L, and ζ.[30] In high-μ categories (hotels), reviews already work well; in low-μ categories (books, wine), a well-specified recommender can dominate – but low μ means a rich type space, which is exactly what makes the recommender hard to specify.

What happens when both systems coexist? Suppose a fraction β of consumers follow the recommender and $1-\beta$ read reviews. Recommender-routed consumers still buy, experience, and review products, contributing to the review base. A more precise recommender routes more consumers correctly, homogenizing the buyer pool and degrading the review system's text accuracy.

Corollary 5 (Recommender-Induced Review Degradation). *Let $\gamma_H^*(\beta)$ denote the review system's equilibrium accuracy when a fraction β of consumers follow the recommender. Then $\partial\gamma_H^*/\partial\beta < 0$: a more active recommender degrades the review system.*

The corollary makes precise the tension between the two information channels. Recommender-routed consumers match at rate $(1+\rho)/2 > \lambda$, but precisely because they are better matched, they homogenize the buyer pool. This is the same self-undermining mechanism as in the baseline model, now triggered by an external routing device. Berman and Katona (2020) model an analogous channel theoretically in social networks. Yoganarasimhan (2020) and Holtz et al. (2020) provide empirical evidence: recommendation algorithms on Spotify and other platforms reduce consumption diversity precisely because accurate routing homogenizes the training data.[31]

Despite this tension, the optimal platform deploys both systems.

Proposition 8 (Optimal Information Architecture). *In a hybrid system where a fraction β of consumers follow a recommender with precision ρ and $1-\beta$ read reviews:*

(i) *The aggregate match rate $M(\beta) = \beta \cdot (1+\rho)/2 + (1-\beta) \cdot \gamma_H^*(\beta)$ is maximized at an interior $\beta^* \in (0,1)$ whenever the recommender is misspecified ($\rho < 1$) and $\rho > 2\gamma^* - 1$.*

(ii) *At the optimum, $\gamma_H^*(\beta^*) < \gamma^*$: the review system is degraded. Yet $M(\beta^*) > \max\{\gamma^*, (1+\rho)/2\}$: the combination strictly dominates either standalone system.*

The result says that deploying the recommender for *everyone* is suboptimal, even when it outperforms the standalone review system. The two systems fail along orthogonal dimensions. A misspecified recommender is bounded by the features it parameterizes: Netflix

[30] This is immediate: the review system's match rate is γ^*, the recommender's is $(1+\rho)/2$; equating gives $\rho = 2\gamma^* - 1$, which rises with γ^* and hence with μ, L, ζ (Proposition 4).

[31] See also Ke et al. (2022) on match efficiency versus surplus extraction and Liu and Liu (2025) on asymmetric effects of AI matching.

can learn from viewing history that a subscriber watches dark comedies, but not whether she prefers slow pacing over sharp dialogue. The review system avoids this ceiling because natural language makes no parametric commitment about what dimensions matter (see also Calvano et al., 2024). But the review system degrades as the buyer pool homogenizes. At β^*, the platform accepts a worse review equilibrium in exchange for the recommender's coverage of parameterized dimensions, while preserving enough review-readers to cover what the recommender is blind to.

The orthogonality is what makes the interior optimum possible: because each system fails on dimensions the other covers, reallocating consumers from one to the other always involves a tradeoff rather than a pure gain. If both systems failed on the same dimensions, the better one would dominate and the optimum would be a corner solution.[32]

A plausible criticism is that consumers often use both systems sequentially: the algorithm narrows the consideration set, then reviews do the fine-grained matching. The core mechanism applies in both cases: better individual routing homogenizes the buyer pool and degrades the review base. If anything, the sequential channel reinforces this effect. We leave formal analysis of sequential information architectures for future work.

The practical pattern is visible across platforms. Amazon recommends books via collaborative filtering while prominently displaying Goodreads reviews that convey dimensions the algorithm cannot capture. Netflix routes subscribers algorithmically while Letterboxd thrives as a review-based complement. Fleder and Hosanagar (2009) show that recommender systems can reduce sales diversity; our model identifies a structural reason why platforms preserve the review channel: reviews provide insurance against model obsolescence and transmit information along dimensions that emerge after the recommender was trained.[33]

6 Extensions and Discussion

Quality uncertainty. When quality is uncertain, scores confound quality and fit. Text retains its fit-informational content regardless: "beautifully written but glacially paced" is a fit signal whether the book is a masterpiece or mediocre. The self-undermining mechanism operates through buyer composition, not score aggregation, so the impossibility result, informational barriers, and diversity dividend all survive when quality varies. The key distinction from quality learning is that fit learning must maintain the *diversity* of the data-generating

[32] If the recommender is perfectly specified ($\rho = 1$ across all dimensions), $\beta^* = 1$ is trivially optimal. But perfect specification is the knife-edge case; for any ϵ of misspecification, the interior solution obtains.

[33] A collaborative-filtering recommender is particularly vulnerable to the degradation in Corollary 5, inheriting the non-monotonicity of Proposition 6. See also Bourreau and Gaudin (2022) and Berman et al. (2024) on the manipulation risk of centralized routing.

process, not extract a common-value signal from selected data.

Competitive implications. In a Hotelling extension, μ would determine effective differentiation: low μ implies strong fit signals, raising willingness to pay and softening price competition.[34] Niche products face a distinctive tradeoff: suppressing negative reviews may raise scores but degrades the fit signal sustaining the niche audience's willingness to pay.

Endogenous review effort and negativity bias. If dissatisfied consumers are more likely to review, the review base overrepresents mismatched buyers, which *increases* text diversity. This reverses the usual concern: in fit learning, overrepresentation of negative reviews is informationally beneficial. If mismatched reviewers also mention more features ($r > 1$), $\bar{\gamma}$ increases in r and the equilibrium requires fewer mismatched buyers. Platforms that encourage detailed negative reviews amplify the diversity dividend.

Helpfulness filtering and feature attrition. On most platforms, mismatched reviews are downvoted as idiosyncratic, reducing effective minority-reviewer stock. As the reviewer pool homogenizes, language calibration shifts toward the majority's frame, and dimensions obvious to fans stop being mentioned entirely. Bondi et al. (2024) document the numerical analogue: aggregating scores across groups can reverse rankings. All three channels amplify diversity loss beyond what the baseline model predicts.

Measuring μ. The feature overlap is the model's key structural parameter. Three empirical proxies are natural: topic-model overlap across reviewer clusters (Jaccard similarity of top-k topics as a proxy for μ/L); aspect universality (share of review aspects mentioned by all segments); and embedding overlap (cosine similarity of segment-specific review centroids). Each generates testable predictions: low-μ categories should exhibit higher score variance, larger distortion from AI summaries, and larger informational barriers to entry. The equilibrium is identified from observables: $V_A = \phi^*(1-\phi^*)$ pins down ϕ^*; the diversity equation then identifies $\bar{\gamma}$; and $\bar{\gamma} = \lambda\gamma(L) + (1-\lambda)\gamma(\mu)$ identifies μ given category-level primitives (ζ, L, λ), without requiring the econometrician to observe consumer types.

Further extensions. The model assumes symmetric products, known types, and binary type space. Niche products face qualitatively different dynamics: the informational barrier (Corollary 3) works against products whose early adopters are poorly matched, and Johnson and Myatt (2006)'s demand-rotation framework would connect review-base composition to

[34] The logic parallels Armstrong and Zhou (2022), who show that match information raises prices by increasing effective differentiation. In our model, the level of match information is itself endogenous.

equilibrium positioning. Noisy self-knowledge reduces effective ζ but leaves the impossibility result intact. With $T > 2$ types, the coupon-collector structure generalizes with a type-specific overlap matrix μ_{ij}.

7 Conclusion

The conventional division of labor assigns reviews the task of sorting products by quality and recommender systems the task of routing consumers to products. But reviews increasingly serve a routing function, and this paper studies how that decentralized routing works, how efficient it is, and what its limits are.

Our central result is that review systems cannot sustain perfect matching. Successful routing homogenizes the reviewer pool and degrades the information that enables further routing. The unique equilibrium features partial matching, sustained by a diversity dividend from mismatched consumers whose reviews cover features that fans' reviews miss. The binding constraint is the feature overlap μ between consumer types, which platforms can target through review design.

The model generates five design implications. First, AI-generated summaries erase the diversity dividend asymmetrically, because frequency-based selection systematically excludes minority features. Second, score-following has a non-monotone welfare effect: uninformed consumers generally degrade matching, but in markets dominated by a single type they improve it by injecting diversity. Third, reviews and recommender systems interact through a tension between direct matching gains and indirect review degradation; the optimal platform deploys both because each fails along dimensions the other covers. Fourth, the shadow value formula gives platforms a concrete recruitment target favoring underrepresented reviewers, with the highest returns in low-overlap categories. Fifth, informational barriers to entry create a structural disadvantage for new products, but the barrier is transient in the text-based system and becomes persistent only through score-based ranking.

The routing perspective also has distributional consequences for market structure. When reviews function primarily as quality signals – the sorting role – consumers converge on the same "best" products, reinforcing superstars and concentrating demand. When reviews function as matching devices – the routing role – consumers disperse across products that fit their respective types, supporting variety and the long tail. The same platform can have opposite effects depending on which role dominates. Goodreads, for instance, presents identical content to every user, yet heterogeneous decoding means that a literary fiction reader and a thriller reader extract different signals from the same review page and end up at different books. The more effectively a review system routes, the less it concentrates

– a prediction that distinguishes our model from the standard quality-learning framework, where better information unambiguously favors high-quality incumbents.[35]

Several limitations deserve emphasis. The binary type space captures the essential trade-off but cannot address richer preference structures. The substitution model for recommender interaction does not capture sequential use, where the recommender shapes the consideration set before reviews are consulted. And the equal-quality assumption, while isolating the fit channel cleanly, means the model does not speak directly to settings where quality and fit uncertainty coexist. Each suggests a natural extension.

The mechanism generalizes beyond reviews: any decentralized information system in which signals are generated by the audience's own consumption faces the same tension. The self-undermining feedback loop requires only that who speaks determines what can be learned.

References

Acemoglu, D., Makhdoumi, A., Malekian, A., and Ozdaglar, A. (2022). Learning from Reviews: The Selection Effect and the Speed of Learning. *Econometrica*, 90(6):2857–2899.

Aridor, G., Gonçalves, D., Kluver, D., Kong, R., and Konstan, J. (2024). The Informational Role of Online Recommendations: Evidence from a Field Experiment. Working Paper.

Armstrong, M. and Zhou, J. (2022). Consumer Information and the Limits to Competition. *American Economic Review*, 112(2):534–577.

Banerjee, A. V. (1992). A Simple Model of Herd Behavior. *Quarterly Journal of Economics*, 107(3):797–817.

Berman, R. and Katona, Z. (2020). Curation Algorithms and Filter Bubbles in Social Networks. *Marketing Science*, 39(2):296–316.

Berman, R., Zhao, H., and Zhu, Y. (2024). Strategic Design of Recommendation Algorithms. Working Paper.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy*, 100(5):992–1026.

Bollinger, B., Lee, N., and Staelin, R. (2023). Vertical Versus Horizontal Variance in Online Reviews and Their Impact on Demand. *Journal of Marketing Research*, 60(3):484–503.

Bondi, T. (2025). Alone, Together: A Model of Social (Mis)Learning from Consumer Reviews. *Marketing Science*, 44(1):1–22.

Bondi, T. and Rossi, M. (2026). Online Reviews: Information Content, Biases, and Platform Design. *Working Paper*.

Bondi, T., Rossi, M., and Stevens, R. L. (2024). The Good, the Bad and the Picky: Consumer Heterogeneity and the Reversal of Product Ratings. *Management Science*, 70(11):7469–7489.

[35] Fleder and Hosanagar (2009) document that recommender systems can increase *or* decrease sales concentration depending on design choices. Our model identifies the analogous margin for review systems: design features that strengthen routing (diverse reviewer pools, structured templates, stratified summaries) should reduce concentration, while features that strengthen sorting (prominent average scores, frequency-based AI summaries) should increase it.

Bourreau, M. and Gaudin, G. (2022). Streaming Platform and Strategic Recommendation Bias. *Journal of Economics & Management Strategy*, 31(1):25–47.

Calvano, E., Calzolari, G., Denicolò, V., and Pastorello, S. (2024). Economics of Recommender Systems. In *Proceedings of the 18th ACM Conference on Recommender Systems (RecSys '24)*, pages 1279–1280. ACM.

Chakraborty, I., Deb, J., and Öry, A. (2024). When Do Consumers Talk? Working Paper.

Che, Y.-K. and Hörner, J. (2018). Recommender Systems as Mechanisms for Social Learning. *Quarterly Journal of Economics*, 133(2):871–925.

Chen, P.-Y., Hitt, L. M., Hong, Y., and Wu, S. (2021). Measuring Product Type and Purchase Uncertainty with Online Product Ratings: A Theoretical Model and Empirical Application. *Information Systems Research*, 32(4):1470–1489.

Chen, Y., Du, J., and Lei, Y. (2024). The Interactions of Customer Reviews and Price and Their Dual Roles in Conveying Quality Information. Working Paper.

Chevalier, J. A. and Mayzlin, D. (2006). The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research*, 43(3):345–354.

Dai, W., Jin, G., Lee, J., and Luca, M. (2018). Aggregation of Consumer Ratings: An Application to Yelp.com. *Quantitative Marketing and Economics*, 16(3):289–339.

De Langhe, B., Fernbach, P. M., and Lichtenstein, D. R. (2016). Navigating by the Stars: Investigating the Actual and Perceived Validity of Online User Ratings. *Journal of Consumer Research*, 42(6):817–833.

Dzyabura, D. and Hauser, J. R. (2019). Recommending Products When Consumers Learn Their Preference Weights. *Marketing Science*, 38(3):417–441.

Ellison, G. and Ellison, S. F. (2014). Match Quality, Search, and the Internet Market for Used Books. *American Economic Review*, 104(1):293–321.

Fainmesser, I. P., Lauga, D. O., and Ofek, E. (2021). Ratings, Reviews, and the Marketing of New Products. *Management Science*, 67(11):7023–7045.

Feldman, P., Papanastasiou, G., and Segev, E. (2019). Social Learning and the Design of New Experience Goods. *Management Science*, 65(9):4178–4198.

Fleder, D. and Hosanagar, K. (2009). Blockbuster Culture's Next Rise or Fall: The Impact of Recommender Systems on Sales Diversity. *Management Science*, 55(5):697–712.

Forman, C., Ghose, A., and Wiesenfeld, B. (2008). Examining the Relationship Between Reviews and Sales: The Role of Reviewer Identity Disclosure in Electronic Markets. *Information Systems Research*, 19(3):291–313.

Godes, D. and Mayzlin, D. (2004). Using Online Conversations to Study Word-of-Mouth Communication. *Marketing Science*, 23(4):545–560.

Guo, L. and Zhang, J. (2012). Consumer Deliberation and Product Line Design. *Marketing Science*, 31(6):995–1007.

Holtz, D., Carterette, B., Chandar, P., Nazari, Z., Cramer, H., and Aral, S. (2020). The Engagement-Diversity Connection: Evidence from a Field Experiment on Spotify. Working Paper, MIT Sloan School of Management.

Hong, Y. and Pavlou, P. A. (2014). Product Fit Uncertainty in Online Markets: Nature, Effects, and Antecedents. *Information Systems Research*, 25(2):328–344.

Hummel, P. and Morgan, J. (2014). A Model of Flops. *RAND Journal of Economics*, 45(1):75–96.

Jerath, K. and Ren, Q. (2021). Consumer Rational (In)Attention to Favorable and Unfavorable Product Information, and Firm Information Design. *Journal of Marketing Research*, 58(2):343–362.

Johnson, J. P. and Myatt, D. P. (2006). On the Simple Economics of Advertising, Marketing, and Product Design. *American Economic Review*, 96(3):756–784.

Ke, T. T., Lin, S., and Lu, M. Y. (2022). Information Design of Online Platforms. Working Paper.

Liu, Z. and Liu, Y. (2025). Asymmetric Impact of AI Matching on Influencer Marketing: Implications for Platform Revenue. *Marketing Science*, 44(1):65–83.

Lei, Z., Yin, D., Mitra, S., and Zhang, H. (2022). Swayed by the Reviews: Disentangling the Effects of Average Ratings and Individual Reviews in Online Word-of-Mouth. *Production and Operations Management*, 31(6):2697–2711.

Luca, M. (2016). Reviews, Reputation, and Revenue: The Case of Yelp.com. Harvard Business School Working Paper No. 12-016.

Mayzlin, D., Dover, Y., and Chevalier, J. A. (2014). Promotional Reviews: An Empirical Investigation of Online Review Manipulation. *American Economic Review*, 104(8):2421–2455.

Pocchiari, M., Proserpio, D., and Dover, Y. (2025). Online Reviews: A Literature Review and Roadmap for Future Research. *International Journal of Research in Marketing*, 42(2):275–297.

Sayedi, A. (2024). Pricing in a Duopoly with Observational Learning. Working Paper, University of Washington.

Sun, M. (2012). How Does the Variance of Product Ratings Matter? *Management Science*, 58(4):696–707.

Vellodi, N. (2022). Ratings Design and Barriers to Entry. Working Paper.

Waldfogel, J. (2007). *The Tyranny of the Market: Why You Can't Always Get What You Want*. Harvard University Press.

Wang, X. (2021). Do Fit Opinions Matter? The Impact of Fit Context on Online Product Reviews. *Information Systems Research*, 32(3):839–857.

Yoganarasimhan, H. (2020). Search Personalization Using Machine Learning. *Management Science*, 66(3):1045–1070.

Zimmermann, S., Herrmann, P., Kundisch, D., and Nault, B. R. (2018). Decomposing the Variance of Consumer Ratings and the Impact on Price and Demand. *Information Systems Research*, 29(4):984–1002.

Appendix: Proofs

Proof of Lemma 1. Partition $\mathcal{F}_a$ into exclusive features $\mathcal{F}_a \setminus \mathcal{F}_b$ (of which there are $L - \mu$) and shared features $\mathcal{F}_a \cap \mathcal{F}_b$ (of which there are μ).

Exclusive features. Consider feature $k \in \mathcal{F}_a \setminus \mathcal{F}_b$. Since $k \notin \mathcal{F}_b$, no type-b reviewer mentions k. Each type-a reviewer mentions k with probability $1/L$ (Assumption 4), independently across reviewers. Feature k is covered if at least one of the n_a type-a reviewers mentions it:

$$\mathbb{P}(k \text{ covered}) = 1 - \mathbb{P}(\text{no type-}a \text{ reviewer mentions } k) = 1 - (1 - 1/L)^{n_a}.$$

Summing over the $L - \mu$ exclusive features and using linearity of expectation gives the first term of (1).

Shared features. Consider feature $k \in \mathcal{F}_a \cap \mathcal{F}_b$. Both type-$a$ and type-b reviewers can mention k, each with probability $1/L$. Feature k is covered if at least one of the $n_a + n_b$ total reviewers mentions it:

$$\mathbb{P}(k \text{ covered}) = 1 - (1 - 1/L)^{n_a}(1 - 1/L)^{n_b} = 1 - (1 - 1/L)^{n_a + n_b}.$$

Summing over the μ shared features gives the second term.

The symmetry claim $C_b(n_a, n_b) = C_a(n_b, n_a)$ follows from the symmetric structure of the model: the roles of types a and b are interchanged, and the coverage formula depends only on how many same-type and other-type reviewers contribute. □

Proof of Lemma 2. *Part (i): Quantity.* Fix ϕ and increase n. Both $n_a = \phi n$ and $n_b = (1 - \phi)n$ increase. In (1), each term is of the form $c[1 - (1 - 1/L)^m]$ where m increases, so coverage increases. Since $\gamma(\cdot)$ is increasing, Γ increases.

Part (i): Variety. Fix n and consider $d\Gamma/d\phi$ at $\phi > 1/2$. Since $n_a = \phi n$ and $n_b = (1-\phi)n$, the chain rule gives $dC_a/d\phi = n(\partial C_a/\partial n_a - \partial C_a/\partial n_b)$ and similarly for C_b. From (1), for shared features both $\partial/\partial n_a$ and $\partial/\partial n_b$ contribute the same term (both enter through $n_a + n_b = n$, which is fixed), so these cancel in $dC_a/d\phi$. Only exclusive features remain. Hence $d\Gamma/d\phi$ depends only on the marginal coverage of exclusive features, and

$$\frac{d\Gamma}{d\phi} \propto \lambda\gamma'(C_a)\big(1 - \tfrac{1}{L}\big)^{n_a} - (1 - \lambda)\gamma'(C_b)\big(1 - \tfrac{1}{L}\big)^{n_b},$$

where the proportionality constant is $n \ln(L/(L-1))(L - \mu) > 0$. At $\phi > 1/2$: $n_a > n_b$, so $(1 - 1/L)^{n_a} < (1 - 1/L)^{n_b}$ (the minority's marginal coverage is higher). Since $C_b < C_a$ when $n_b < n_a$, the concavity of γ gives $\gamma'(C_b) \geq \gamma'(C_a)$. Both effects make the second (negative) term dominate: the ratio of the second to the first is at least $\frac{(1-\lambda)\gamma'(C_b)(1-1/L)^{n_b}}{\lambda\gamma'(C_a)(1-1/L)^{n_a}} = \frac{1-\lambda}{\lambda} \cdot \frac{\gamma'(C_b)}{\gamma'(C_a)} \cdot (1-1/L)^{n_b-n_a}$. The first factor is less than 1, but the third factor is $(1-1/L)^{-n(2\phi-1)}$, which grows exponentially in n. For any fixed $\phi > 1/2$ and $\lambda < 1$, there exists $\underline{n}(\phi, \lambda)$ such that the ratio exceeds 1 for all $n \geq \underline{n}$, giving $d\Gamma/d\phi < 0$.

Part (ii): Diverse limit. As $n \to \infty$ with $\phi \in (0, 1)$ fixed, both $n_a = \phi n \to \infty$ and $n_b = (1-\phi)n \to \infty$. Hence $(1-1/L)^{n_a} \to 0$ and $(1-1/L)^{n_b} \to 0$. From (1): $C_a \to (L-\mu)\cdot 1 + \mu \cdot 1 = L$. By symmetry, $C_b \to L$. Therefore $\Gamma \to \lambda\gamma(L) + (1-\lambda)\gamma(L) = \gamma(L) = \hat{\gamma}$.

Part (iii): Homogeneous limit. With $\phi = 1$: $n_b = 0$, so $(1-1/L)^{n_b} = 1$. From (1): $C_a \to (L-\mu) + \mu = L$. For C_b: by symmetry, $C_b(n_a, 0) = C_a(0, n_a)$. From (1) with $n_a = 0$ and $n_b = n_a$: $C_a(0, n_a) = (L-\mu)[1-1] + \mu[1-(1-1/L)^{n_a}] \to \mu$. Hence $\Gamma \to \lambda\gamma(L) + (1-\lambda)\gamma(\mu) = \bar{\gamma}$.

Part (iv). $\bar{\gamma} = \lambda\gamma(L) + (1-\lambda)\gamma(\mu)$. Since $\gamma(\cdot)$ is strictly increasing, $\bar{\gamma}$ is strictly increasing in μ. Similarly, $\hat{\gamma} = \gamma(L)$ is strictly increasing in L (more features means more signals) and in ζ (each signal is more informative).

For the inequality $\hat{\gamma} > \bar{\gamma}$: $\hat{\gamma} - \bar{\gamma} = \gamma(L) - [\lambda\gamma(L) + (1-\lambda)\gamma(\mu)] = (1-\lambda)[\gamma(L) - \gamma(\mu)]$. Since $1-\lambda > 0$ and $\gamma(L) > \gamma(\mu)$ whenever $L > \mu$ (which holds by assumption $\mu < L$), we have $\hat{\gamma} > \bar{\gamma}$. □

Proof of Proposition 1. The proof uses only the three abstract conditions on γ.

Existence. At $\phi = \lambda$: $\gamma(\lambda) > 1/2$ by condition (ii), so $F(\lambda) = \psi(\gamma(\lambda)) > \lambda$ (since $\psi(\gamma) > \lambda$ iff $\gamma > 1/2$). At $\phi = 1$: $\gamma(1) < 1$ by condition (iii), so $F(1) = \psi(\gamma(1)) < 1$. By continuity of F, the intermediate value theorem gives $\phi^* \in (\lambda, 1)$ with $F(\phi^*) = \phi^*$.

Uniqueness. γ is strictly decreasing on $(\lambda, 1)$ by condition (i). Since ψ is strictly increasing, $F = \psi \circ \gamma$ is strictly decreasing. A strictly decreasing function crosses the identity at most once.

Stability. The ODE $\dot{\phi} = F(\phi) - \phi$ has ϕ^* as its unique equilibrium on $[\lambda, 1]$. Since F is strictly decreasing and crosses the identity once, $F(\phi) > \phi$ for $\phi < \phi^*$ and $F(\phi) < \phi$ for $\phi > \phi^*$. Hence ϕ^* is globally asymptotically stable for the ODE. The stochastic approximation argument (Robbins–Monro) extends to the discrete-time stochastic system; the full verification is given in the quadratic-specification proof below.

Perfect matching inconsistent. $F(1) = \psi(\gamma(1)) < 1$ since $\gamma(1) < 1$ by condition (iii). □

Proof of Proposition 1 (quadratic specification). At $\phi = \lambda$: $\gamma(\lambda) = \bar{\gamma} + (\hat{\gamma} - \bar{\gamma}) \cdot 4\lambda(1-\lambda)$. Since $\lambda \in (1/2, 1)$, $4\lambda(1-\lambda) \in (0, 1)$, so $\gamma(\lambda) > \bar{\gamma}$. We need $\bar{\gamma} > 1/2$. From the definition: $\bar{\gamma} = \lambda\gamma(L) + (1-\lambda)\gamma(\mu)$. For any $L \geq 1$ and $\zeta > 1/2$: $\gamma(L) > 1/2$. Also $\gamma(\mu) \geq \gamma(0) = 1/2$. Hence $\bar{\gamma} \geq \lambda \cdot (\text{something} > 1/2) + (1-\lambda) \cdot 1/2 > 1/2$. (The hypothesis $\hat{\gamma} > 1/2$ ensures this formally.) Since $\gamma(\lambda) > 1/2$, we have $\psi(\gamma(\lambda)) > \lambda$. To see this: $\psi(\gamma) = \lambda\gamma/[\lambda\gamma + (1-\lambda)(1-\gamma)]$. $\psi(\gamma) > \lambda$ iff $\lambda\gamma > \lambda[\lambda\gamma + (1-\lambda)(1-\gamma)]$ iff $\gamma(1-\lambda) > (1-\lambda)(1-\gamma)$ iff $\gamma > 1-\gamma$ iff $\gamma > 1/2$. So $F(\lambda) = \psi(\gamma(\lambda)) > \lambda$.

At $\phi = 1$: $4\phi(1-\phi) = 0$, so $\gamma(1) = \bar{\gamma}$. Then $F(1) = \psi(\bar{\gamma})$. We need $\psi(\bar{\gamma}) < 1$. Since $\psi(\gamma) < 1$ iff $(1-\lambda)(1-\gamma) > 0$ iff $\gamma < 1$, and $\bar{\gamma} < 1$ by hypothesis, $F(1) < 1$.

Since F is continuous on $[\lambda, 1]$, $F(\lambda) > \lambda$, and $F(1) < 1$, by the intermediate value theorem there exists $\phi^* \in (\lambda, 1)$ with $F(\phi^*) = \phi^*$.

Uniqueness. We show F is strictly decreasing on $[\lambda, 1]$. The diversity index $d(\phi) = 4\phi(1-\phi)$ has derivative $d'(\phi) = 4(1-2\phi) < 0$ for $\phi > 1/2$. Since $\lambda > 1/2$, $d(\phi)$ is strictly decreasing on $[\lambda, 1]$. Since $\hat{\gamma} > \bar{\gamma}$ (Proposition 2(iv)), $\gamma(\phi) = \bar{\gamma} + (\hat{\gamma} - \bar{\gamma})d(\phi)$ is strictly decreasing. Since $\psi(\gamma) = \lambda\gamma/[\lambda\gamma + (1-\lambda)(1-\gamma)]$ is strictly increasing in γ (the numerator increases and the denominator decreases), the composition $F = \psi \circ \gamma$ is strictly decreasing.

A strictly decreasing function and the strictly increasing identity function $\mathrm{id}(\phi) = \phi$ cross at most once. Since existence gives at least one crossing, there is exactly one.

Stability. The update rule (5) takes the form

$$\phi_{t+1} = \phi_t + a_t[h(\phi_t) + \xi_t],$$

where $a_t = 1/(n_t + 1)$, $h(\phi) = F(\phi) - \phi$, and $\xi_t = \hat{\phi}_t - \psi(\gamma_t)$ is the deviation of the realized flow composition $\hat{\phi}_t$ from its conditional mean $\psi(\gamma_t)$, so that $\mathbb{E}[\xi_t \mid \mathcal{F}_t] = 0$. We verify the conditions of a standard stochastic approximation theorem (see, e.g., Borkar, 2008, Theorem 2.1; the compact state space $[\lambda, 1]$ simplifies the verification):

(i) *Step sizes.* Since $\gamma(\phi) \leq \gamma(\lambda) < 1$ for all $\phi \in [\lambda, 1]$, each product receives at least a fraction $(1-\lambda)(1-\gamma(\lambda)) > 0$ of the unit mass each period (from type-b consumers who buy A). Hence $n_t \geq c \cdot t$ for $c = (1-\lambda)(1-\gamma(\lambda)) > 0$, ensuring $a_t = O(1/t)$, $a_t \to 0$, $\sum a_t = \infty$, and $\sum a_t^2 < \infty$.

(ii) *Mean field.* $h(\phi) = F(\phi) - \phi$ is Lipschitz continuous on $[\lambda, 1]$ (since F is a smooth composition of ψ and γ), with $h(\phi^*) = 0$.

(iii) *Noise.* ξ_t is a martingale difference with bounded second moment: $\mathbb{E}[\xi_t^2 \mid \mathcal{F}_t] \leq 1$ since $\hat{\phi}_t \in [0, 1]$.

(iv) *Global attractivity.* The ODE $\dot{\phi} = h(\phi)$ has ϕ^* as its unique equilibrium on $[\lambda, 1]$, and $h(\phi) > 0$ for $\phi < \phi^*$, $h(\phi) < 0$ for $\phi > \phi^*$ (since F is strictly decreasing and crosses the identity once). Hence ϕ^* is globally asymptotically stable for the ODE.

These conditions guarantee $\phi_t \to \phi^*$ almost surely.

Remaining claims. Since $\phi^* \in (\lambda, 1)$ by construction, $\phi^* < 1$ (partial matching). Since $\gamma(\phi)$ is strictly decreasing on $[\lambda, 1]$ and $\phi^* \in (\lambda, 1)$, we have $\gamma^* = \gamma(\phi^*) \in (\gamma(1), \gamma(\lambda)) = (\bar{\gamma}, \bar{\gamma} + (\hat{\gamma} - \bar{\gamma}) \cdot 4\lambda(1-\lambda)) \subset (\bar{\gamma}, \hat{\gamma})$. At the steady state, each consumer identifies the correct product with probability γ^* and buys it, so the fraction of all consumers correctly matched is γ^*. □

Proof of Propositions 4 and 5. *Part (i): ϕ^* and γ^* increasing in μ.* Increasing μ raises $\bar{\gamma} = \lambda\gamma(L) + (1-\lambda)\gamma(\mu)$ (since $\gamma(\mu)$ is increasing in μ). This shifts $\gamma(\phi) = \bar{\gamma} + (\hat{\gamma} - \bar{\gamma}) \cdot 4\phi(1-\phi)$ upward for every ϕ. Since ψ is increasing, $F(\phi) = \psi(\gamma(\phi))$ shifts upward for every ϕ.

At the old fixed point ϕ^*_{old}, $F(\phi^*_{\text{old}})$ now exceeds ϕ^*_{old} (because F has shifted up). Since F is strictly decreasing and the identity is strictly increasing, the new crossing point ϕ^*_{new} must be to the right: $\phi^*_{\text{new}} > \phi^*_{\text{old}}$.

For γ^*: the fixed-point condition is $\psi(\gamma^*) = \phi^*$. Since ψ is strictly increasing, ψ^{-1} exists and is also strictly increasing. Hence $\gamma^* = \psi^{-1}(\phi^*)$ is increasing in ϕ^*, so γ^* increases when μ increases.

Part (ii). The argument is identical: increasing L or ζ raises both $\hat{\gamma}$ and $\bar{\gamma}$ (Proposition 2(iv)), shifting $\gamma(\phi)$ upward and hence shifting the fixed point rightward.

Part (iii): Diminishing marginal effect of μ. $\partial\bar{\gamma}/\partial\mu = (1-\lambda)\gamma'(\mu)$. From the formula $\gamma(m) = \zeta^m/[\zeta^m + (1-\zeta)^m]$, write $\ell(m) = (\zeta/(1-\zeta))^m$, so $\gamma(m) = \ell/(\ell+1)$. Differentiating: $\gamma'(m) = \ell'/(\ell+1)^2 = \ell \ln(\zeta/(1-\zeta))/(\ell+1)^2 = \gamma(m)(1-\gamma(m)) \ln(\zeta/(1-\zeta))$.

The factor $\gamma(m)(1-\gamma(m))$ achieves its maximum of $1/4$ at $\gamma = 1/2$ (i.e., $m = 0$) and is strictly decreasing for $\gamma > 1/2$ (i.e., $m > 0$). Hence $\gamma'(\mu)$ is highest at $\mu = 0$, and $\partial\bar{\gamma}/\partial\mu$ is diminishing. Since γ^* is increasing in $\bar{\gamma}$ (Proposition 1 and the fixed-point argument in Part (i)), the diminishing marginal effect on $\bar{\gamma}$ translates into a diminishing marginal effect on γ^*, though the mapping is not one-to-one because the sensitivity $\partial\gamma^*/\partial\bar{\gamma}$ also depends on the equilibrium composition.

Part (iv): Baseline quality as the binding constraint. From (3): $\partial\gamma(\phi)/\partial\bar{\gamma} = 1 - 4\phi(1-\phi)$ and $\partial\gamma(\phi)/\partial\hat{\gamma} = 4\phi(1-\phi)$. We need $1 - 4\phi(1-\phi) > 4\phi(1-\phi)$, i.e., $8\phi^*(1-\phi^*) < 1$, which holds iff $\phi^* > \underline{\phi} \equiv (2+\sqrt{2})/4$. Since ϕ^* is increasing in $\bar{\gamma}$, L, and ζ (Parts (i)–(ii)), there exists a threshold in each primitive above which the condition is satisfied. Below $\underline{\phi}$, the diversity gap $\hat{\gamma} - \bar{\gamma}$ matters more than the floor itself. □

Proof of Corollary 5. When a fraction β of consumers follow the recommender, recommender-routed buyers of product A have type-a share $\phi_{\text{rec}} = \lambda(1+\rho)/[\lambda(1+\rho) + (1-\lambda)(1-\rho)] > \phi^*$ (for $\rho > 2\gamma^* - 1$). The review base composition is $\phi_{\text{total}} = \beta\phi_{\text{rec}} + (1-\beta)\psi(\gamma(\phi_{\text{total}}))$. Increasing β raises ϕ_{total} (more homogeneous reviews), which reduces $\gamma(\phi_{\text{total}})$ since γ is decreasing. Hence $\gamma^*_H(\beta)$ is decreasing in β. □

Proof of Proposition 8*. Part (i).* Consider the aggregate match rate $M(\beta) = \beta(1+\rho)/2 + (1-\beta)\gamma_H^*(\beta)$. At $\beta = 0$, $M = \gamma^*$. At $\beta = 1$, $M = (1+\rho)/2 > \gamma^*$ when $\rho > 2\gamma^* - 1$. We show $M'(\beta) < 0$ near $\beta = 1$. The derivative is:

$$M'(\beta) = \frac{1+\rho}{2} - \gamma_H^*(\beta) + (1-\beta)\frac{d\gamma_H^*}{d\beta}.$$

As $\beta \to 1$, the third term vanishes; the first two approach $(1+\rho)/2 - \gamma_H^*(1)$. But a misspecified recommender has match rate bounded by $(1+\rho_1)/2$ along its observed dimension, so the marginal review-reader contributes welfare that includes the unparameterized dimension, exceeding $(1+\rho)/2$. Hence removing the last review-readers is costly: $M'(\beta) < 0$ near $\beta = 1$. Since $M(1) > M(0)$ and M is decreasing near $\beta = 1$, M achieves its maximum at some $\beta' < 1$ with $M(\beta') > M(1) > M(0)$. Hence the optimum $\beta^* \in (0,1)$ and $M(\beta^*)$ exceeds both boundary values.

Part (ii). The review system's equilibrium $\gamma_H^*(\beta)$ solves a modified fixed-point condition. Recommender-routed buyers of product A have type-a share $\phi_{\text{rec}} = \lambda(1+\rho)/[\lambda(1+\rho) + (1-\lambda)(1-\rho)]$, which exceeds ϕ^* for ρ sufficiently large. The total review base for A is a weighted mixture: $\phi_{\text{total}} = w\phi_{\text{rec}} + (1-w)\phi_{\text{rev}}$, where w is the share of A's reviews from recommender-routed buyers and ϕ_{rev} is the composition of review-guided buyers. Since $\phi_{\text{rec}} > \phi^*$, injecting recommender-routed reviews raises ϕ_{total} for any given ϕ_{rev}, reducing text accuracy $\gamma(\phi_{\text{total}})$. The review system's equilibrium shifts: $\gamma_H^* < \gamma^*$. But aggregate welfare $M(\beta^*) > M(0) = \gamma^*$ by the optimality of β^*, and $M(\beta^*) > M(1) = (1+\rho)/2$ because the interior maximum strictly exceeds the boundary value (from $M'(\beta) < 0$ near $\beta = 1$). □

Proof of Proposition 2*.* The inverse selection function is derived by solving $\phi = \psi(\gamma)$ for γ:

$$\psi^{-1}(\phi) = \frac{\phi(1-\lambda)}{\lambda(1-\phi) + \phi(1-\lambda)}.$$

The fixed-point condition $\psi^{-1}(\phi^*) = \gamma(\phi^*)$ becomes, in terms of $m = 1 - \phi$:

$$\frac{(1-m)(1-\lambda)}{(1-\lambda) + m(2\lambda - 1)} = \bar{\gamma} + 4\Delta m(1-m).$$

Define $f(m) = \psi^{-1}(1-m) - \gamma(1-m)$. We need $f(m^*) = 0$. Note $f(0) = 1 - \bar{\gamma} > 0$ and $f'(0) = -(R + 4\Delta) < 0$, so f is positive and decreasing near $m = 0$.

Lower bound. Expanding around $m = 0$: $\psi^{-1}(1-m) \approx 1 - Rm$ (since $\frac{d}{dm}\psi^{-1}(1-m)\big|_{m=0} = -R$, where $R = \lambda/(1-\lambda)$) and $\gamma(1-m) \approx \bar{\gamma} + 4\Delta m$. The linearized equation $1 - Rm = \bar{\gamma} + 4\Delta m$ yields $m = (1-\bar{\gamma})/(R + 4\Delta)$. To establish this as a lower bound, note that $\psi^{-1}(\phi)$ is strictly convex in ϕ on $(0,1)$: differentiating twice gives $(\psi^{-1})''(\phi) =$

$2\lambda(1-\lambda)(2\lambda-1)/[\lambda(1-\phi)+\phi(1-\lambda)]^3 > 0$ for $\lambda > 1/2$. Meanwhile, $\gamma(\phi) = \bar{\gamma}+(\hat{\gamma}-\bar{\gamma})\cdot 4\phi(1-\phi)$ is strictly concave: $\gamma''(\phi) = -8(\hat{\gamma}-\bar{\gamma}) < 0$. Hence $f(m) = \psi^{-1}(1-m) - \gamma(1-m)$ is convex in m (convex minus concave), and the linearized root undershoots: $m^* \geq (1-\bar{\gamma})/(R+4\Delta)$.

Upper bound. At any steady state, $\gamma^* = \bar{\gamma} + 4\Delta m^*(1-m^*) > \bar{\gamma}$. Since ψ is increasing, $\phi^* = \psi(\gamma^*) > \psi(\bar{\gamma})$, so $m^* = 1-\phi^* < 1-\psi(\bar{\gamma})$.

Match efficiency. The approximation $\gamma^* \approx 1 - Rm^*$ follows from the linear expansion: $\psi^{-1}(1-m) \approx 1 - Rm$ and the equilibrium condition $\psi^{-1}(\phi^*) = \gamma^*$. The approximation error is quadratic in m^*, arising from the higher-order terms in the Taylor expansion of ψ^{-1}. □

Proof of Corollary 1. The welfare gap closed is $(\gamma^* - \lambda)/(1-\lambda)$. From the comparative statics (Proposition 4), γ^* is increasing in μ and ζ. As $\bar{\gamma} \to 1$ (which requires both μ close to L and $\gamma(L)$ close to 1, the latter holding when L is large or ζ is close to 1), $\Delta \to 0$ and $m^* \to (1-\bar{\gamma})/(R+4\Delta) \to 0$, so $\gamma^* \to 1$ and the fraction approaches 1. □

Proof of Proposition 6. Let α denote the fraction of score-followers. The composition of A's buyer pool satisfies

$$\phi = \alpha\lambda + (1-\alpha)\psi(\gamma(\phi)), \tag{9}$$

which implicitly defines $\phi^*(\alpha)$ and hence $\gamma^*(\alpha) = \gamma(\phi^*(\alpha))$. Score-followers match at rate λ; text-readers match at rate $\gamma^*(\alpha)$. Hence $M(\alpha) = \alpha\lambda + (1-\alpha)\gamma^*(\alpha)$ with $M(0) = \gamma^*$ and $M(1) = \lambda$.

Implicit differentiation. Write $G(\phi,\alpha) = \alpha\lambda + (1-\alpha)\psi(\gamma(\phi)) - \phi$. At a fixed point, $G = 0$. Since $G_\phi = (1-\alpha)\psi'(\gamma)\gamma'(\phi) - 1 < -1$ (because $\psi' > 0$ and $\gamma' < 0$), the implicit function theorem gives

$$\frac{d\phi^*}{d\alpha} = -\frac{G_\alpha}{G_\phi} = \frac{\lambda - \psi(\gamma^*)}{1-(1-\alpha)\psi'(\gamma^*)\gamma'(\phi^*)} = \frac{\lambda - \phi^*}{1+(1-\alpha)|\psi'\gamma'|} < 0,$$

where $\lambda - \phi^* < 0$ and the denominator exceeds 1. Score-followers reduce ϕ^* (diversify the pool). Differentiating M:

$$\left.\frac{dM}{d\alpha}\right|_{\alpha=0} = (\lambda - \gamma^*) + \gamma'(\phi^*)\cdot\frac{\lambda-\phi^*}{1-\psi'(\gamma^*)\gamma'(\phi^*)}. \tag{10}$$

The first term is the direct loss from replacing informed consumers with uninformed ones. The second is the indirect diversity gain from a less homogeneous pool.

Existence and uniqueness of $\bar{\lambda}$. Define $\Phi(\lambda) \equiv dM/d\alpha\big|_{\alpha=0}$; the threshold $\bar{\lambda}$ is defined as the largest zero of Φ. We show $\Phi < 0$ near $\lambda = 1/2$ and $\Phi > 0$ for λ sufficiently close to 1.

Write the indirect gain as $B(\lambda) = |\gamma'(\phi^*)|\cdot(\phi^*-\lambda)/(1+|\psi'\gamma'|)$ and the direct loss as $D(\lambda) = \gamma^* - \lambda$, so $\Phi = B - D$.

(a) As $\lambda \to 1$: $\phi^* \to 1$, so $\gamma^* = \gamma(\phi^*) \to \gamma(1) = \bar{\gamma}$. Since $\bar{\gamma} < 1$ for all finite L and $\zeta < 1$, we have $D = \gamma^* - \lambda \to \bar{\gamma} - 1 < 0$, i.e., D becomes negative (score-followers actually match *better* than text-readers). Since $B \geq 0$ and $-D > 0$, $\Phi = B - D > 0$.

(b) As $\lambda \to 1/2^+$: at $\lambda = 1/2$ exactly, ψ is the identity ($\psi(\gamma) = \gamma$), so $\phi^* = \gamma^*$ and $\psi'(\gamma^*) = 1$. Hence $D = \gamma^* - 1/2 > 0$, $B = |\gamma'|(\phi^* - 1/2)/(1 + |\gamma'|)$, and $B/D = |\gamma'|/(1 + |\gamma'|) < 1$. So $\Phi = B - D < 0$ at $\lambda = 1/2$. By continuity, $\Phi < 0$ in a neighborhood of $\lambda = 1/2$.

(c) Existence and sign pattern. Since $\Phi = B - D$ with $B \geq 0$ and $D = \gamma^* - \lambda$, we show D is strictly decreasing in λ, which determines the sign pattern.

Total differentiation of the fixed-point system $\phi^* = \psi(\gamma^*, \lambda)$ and $\gamma^* = \gamma(\phi^*, \lambda)$ gives

$$\frac{d\gamma^*}{d\lambda} = \frac{\gamma_\lambda - |\gamma_\phi| \cdot \psi_\lambda}{1 + \psi_\gamma \cdot |\gamma_\phi|},$$

where $\gamma_\lambda = [\gamma(L) - \gamma(\mu)](1 - 4\phi^*(1 - \phi^*)) > 0$, $\psi_\lambda > 0$, $\psi_\gamma = \psi'(\gamma^*) > 0$, and $|\gamma_\phi| = |\gamma'(\phi^*)| > 0$. For $d\gamma^*/d\lambda < 1$, we need $\gamma_\lambda - |\gamma_\phi| \cdot \psi_\lambda < 1 + \psi_\gamma \cdot |\gamma_\phi|$. Since $\gamma_\lambda \leq \gamma(L) - \gamma(\mu) < 1$ (because $\gamma(L) < 1$ and $\gamma(\mu) \geq 1/2$) and $1 + \psi_\gamma \cdot |\gamma_\phi| > 1$: the inequality holds a fortiori. Hence $dD/d\lambda = d\gamma^*/d\lambda - 1 < 0$: D is strictly decreasing.

Since $D(\lambda)$ is strictly decreasing, continuous, positive near $\lambda = 1/2$ (from (b)), and negative near $\lambda = 1$ (from (a)), it has a unique zero λ_0. For $\lambda > \lambda_0$: $D < 0$ and $B \geq 0$, so $\Phi = B - D > 0$ (score-following improves welfare). For λ near $1/2$: $\Phi < 0$ (from (b)). Define $\bar{\lambda}$ as the largest value at which $\Phi(\bar{\lambda}) = 0$. Since $\Phi > 0$ for all $\lambda > \lambda_0$, we have $\bar{\lambda} \leq \lambda_0$. For $\lambda > \bar{\lambda}$, either $\lambda > \lambda_0$ (in which case $\Phi > 0$ directly) or $\lambda \in (\bar{\lambda}, \lambda_0]$ (in which case $\Phi > 0$ by the definition of $\bar{\lambda}$ as the largest zero).[36]

Comparative statics of $\bar{\lambda}$. We show $\bar{\lambda}$ is increasing in μ and ζ. Increasing μ raises $\bar{\gamma}$ and hence γ^* (Proposition 4(i)). At any fixed λ, the direct loss $D = \gamma^* - \lambda$ increases, shifting $\Phi = B - D$ downward. The region $\{\lambda : \Phi(\lambda) < 0\}$ expands, so $\bar{\lambda}$ shifts right.[37] The argument for ζ is analogous: increasing ζ raises both $\hat{\gamma}$ and $\bar{\gamma}$, strengthening the text signal at every composition. This increases D at each λ, shifting Φ downward and raising $\bar{\lambda}$.[38] □

Proof of Corollary 2. *Part (i).* At the steady state, product A's type-b reviewer stock is $n_b^A = (1 - \lambda)(1 - \gamma^*)n$, while product B's type-a reviewer stock is $n_a^B = \lambda(1 - \gamma^*)n$. Since

[36]Numerical computation confirms that Φ has a unique zero and is strictly increasing across the full parameter space, so $\bar{\lambda}$ is also the *smallest* zero. The analytical argument establishes the weaker claim that suffices for the proposition.

[37]More precisely: if $\Phi(\lambda; \mu) < 0$ for all $\lambda \leq \bar{\lambda}(\mu)$, then increasing μ makes Φ more negative at $\bar{\lambda}(\mu)$, so the largest zero must lie strictly to the right.

[38]For ζ, the indirect gain B also shifts; numerical verification confirms the net effect is monotone across the full parameter space.

$n_b^A/n_a^B = (1-\lambda)/\lambda = 1/R < 1$, the minority's reviewer stock on its "off" product is smaller. Coverage exhibits diminishing returns (coupon-collector): the marginal coverage of a type-b review on A for exclusive features is $\frac{L-\mu}{L}(1-\frac{1}{L})^{n_b^A}$, which is larger when n_b^A is smaller. Hence each minority reviewer's marginal contribution to A's text accuracy exceeds each majority reviewer's marginal contribution to B's, by a factor that is increasing in R.

Part (ii). $R = \lambda/(1-\lambda)$ is strictly increasing in λ. □

Proof of Proposition 3. At the steady state, product A's text accuracy for type-b readers depends on coverage of $\mathcal{F}_b$. The exclusive features $\mathcal{F}_b \setminus \mathcal{F}_a$ (of which there are $L-\mu$) are covered only by type-b reviewers, each of whom mentions a given feature with probability $1/L$. The expected number of exclusive features covered is $(L-\mu)[1-(1-1/L)^{n_b^A}]$, giving total type-$b$ coverage of $m_b^* = \mu + (L-\mu)[1-(1-1/L)^{n_b^A}]$ (shared features are fully covered in the large-n regime). The marginal coverage from one additional type-b review is

$$\frac{\partial C_b^{\text{excl}}}{\partial n_b^A} = \frac{L-\mu}{L}\left(1-\frac{1}{L}\right)^{n_b^A},$$

which is the probability that the new review covers a previously uncovered exclusive feature. If it does, the type-b reader's coverage increases from m_b^* to m_b^*+1, improving her accuracy by $\gamma(m_b^*+1)-\gamma(m_b^*) \approx \gamma'(m_b^*)$. The marginal welfare gain is the minority population weight $(1-\lambda)$ times the signal value $\gamma'(m_b^*)$ times the marginal coverage probability, giving (8). Part (i): increasing μ raises m_b^* (more shared features covered), and since γ' is decreasing for $m > 0$, the signal value falls; additionally, $(L-\mu)/L$ falls. Part (ii): increasing n_b^A raises m_b^* (reducing $\gamma'(m_b^*)$) and reduces the coupon-collector term. Part (iii): for fixed L, increasing $L-\mu$ (i.e., decreasing μ) raises $(L-\mu)/L$ directly and lowers m_b^* (fewer shared features means lower baseline coverage), increasing both the coupon-collector marginal and the signal value. □

Proof of Proposition 7 and Corollary 4. *Part (i).* Each type-a reviewer mentions a feature from $\mathcal{F}_a$ uniformly at random; each type-b reviewer mentions a feature from $\mathcal{F}_b$ uniformly at random. At the steady state, ϕ^* fraction of reviewers are type a and $1-\phi^*$ are type b. A feature f's mention probability per review is:

$$p(f) = \begin{cases} \phi^*/L + (1-\phi^*)/L = 1/L & \text{if } f \in \mathcal{F}_a \cap \mathcal{F}_b, \\ \phi^*/L & \text{if } f \in \mathcal{F}_a \setminus \mathcal{F}_b, \\ (1-\phi^*)/L & \text{if } f \in \mathcal{F}_b \setminus \mathcal{F}_a. \end{cases}$$

Since $\phi^* > 1/2$: $1/L > \phi^*/L > (1-\phi^*)/L$. By the law of large numbers, empirical mention frequencies converge to these values as $n \to \infty$, so the ranking by frequency places the μ shared features first, followed by the $L - \mu$ exclusive-a features, followed by the $L - \mu$ exclusive-b features, with probability approaching 1. For $k > \mu$, the top-k list includes all μ shared features and $k - \mu$ exclusive-a features. No exclusive-b features appear. A type-b reader attends to $\mathcal{F}_b = (\mathcal{F}_b \setminus \mathcal{F}_a) \cup (\mathcal{F}_a \cap \mathcal{F}_b)$; of the features in the summary, only the μ shared features are in her attention set. Her coverage is $\min(k, \mu) = \mu$, yielding accuracy $\gamma(\mu)$.

Part (ii). A type-a reader attends to $\mathcal{F}_a$; of the top-k features, both shared and exclusive-a features are in her set, giving coverage $\min(k, L)$. A type-b reader has coverage $\min(k, \mu)$. Hence $\Gamma_{\text{sum}}(k) = \lambda\gamma(\min(k, L)) + (1 - \lambda)\gamma(\min(k, \mu))$. For $\mu \le k \le L$: the type-b reader's coverage is capped at μ (no $\mathcal{F}_b$-exclusive features appear in the top-k list). At $k = L$: $\Gamma_{\text{sum}} = \lambda\gamma(L) + (1 - \lambda)\gamma(\mu) = \bar{\gamma}$. For $\mu \le k < L$: $\gamma(k) < \gamma(L)$ while $\gamma(\min(k, \mu)) = \gamma(\mu)$, so $\Gamma_{\text{sum}}(k) < \bar{\gamma}$. (For $k > L$, the summary begins to include $\mathcal{F}_b$-exclusive features, and Γ_{sum} increases above $\bar{\gamma}$; the policy-relevant regime is $k \le L$.)

Part (iii). At the steady state with full text, $\gamma^* > \bar{\gamma}$ (Proposition 1). For $k \le L$: $\Gamma_{\text{sum}}(k) \le \bar{\gamma} < \gamma^*$. We show $dM/d\alpha\big|_{\alpha=0} < 0$ and $M(1) < \gamma^*$, establishing that introducing summary-readers strictly reduces welfare at the margin and in the limit.

Endpoint. $M(1) = \Gamma_{\text{sum}}(k) \le \bar{\gamma} < \gamma^*$.

Marginal effect. Summary-readers diversify the buyer pool: their composition $\psi(\Gamma_{\text{sum}})$ is strictly below $\phi^* = \psi(\gamma^*)$ (since $\Gamma_{\text{sum}} < \gamma^*$ and ψ is increasing). The equilibrium with summary fraction α satisfies $\phi = \alpha\psi(\Gamma_{\text{sum}}) + (1 - \alpha)\psi(\gamma(\phi))$. Implicit differentiation (as in the proof of Proposition 6) gives

$$\frac{dM}{d\alpha}\bigg|_{\alpha=0} = \big(\Gamma_{\text{sum}} - \gamma^*\big) + |\gamma'(\phi^*)| \cdot \frac{\phi^* - \psi(\Gamma_{\text{sum}})}{1 + |\psi'(\gamma^*)\gamma'(\phi^*)|}.$$

The first term (direct loss) equals $-(\gamma^* - \Gamma_{\text{sum}})$. The second (indirect diversity gain) equals $|\gamma'| \cdot (\phi^* - \psi_S)/(1 + |F'(\phi^*)|)$, where $\psi_S = \psi(\Gamma_{\text{sum}})$ and $|F'(\phi^*)| = \psi'(\gamma^*) \cdot |\gamma'(\phi^*)|$. By the mean value theorem, $\phi^* - \psi_S = \psi(\gamma^*) - \psi(\Gamma_{\text{sum}}) = \psi'(\xi)(\gamma^* - \Gamma_{\text{sum}})$ for some $\xi \in (\Gamma_{\text{sum}}, \gamma^*)$. Substituting:

$$\frac{dM}{d\alpha}\bigg|_{\alpha=0} = (\gamma^* - \Gamma_{\text{sum}})\left[-1 + \frac{|\gamma'| \cdot \psi'(\xi)}{1 + |F'(\phi^*)|}\right].$$

Since $\gamma^* - \Gamma_{\text{sum}} > 0$, the sign equals the sign of the bracketed term. We need $|\gamma'| \cdot \psi'(\xi) < 1 + |F'|$, i.e., $|\gamma'| \cdot [\psi'(\xi) - \psi'(\gamma^*)] < 1$. Since ψ is strictly concave for $\lambda > 1/2$ ($\psi''(\gamma) = -2\lambda(1 - \lambda)(2\lambda - 1)/[\lambda\gamma + (1 - \lambda)(1 - \gamma)]^3 < 0$) and $\xi < \gamma^*$: $\psi'(\xi) > \psi'(\gamma^*)$, so the difference is positive. We bound it via a second application of the mean value theorem: $\psi'(\xi) - \psi'(\gamma^*) = |\psi''(\eta)|(\gamma^* - \xi)$ for some $\eta \in (\xi, \gamma^*)$.

We bound each factor. First, $|\gamma'(\phi^*)| = 4\Delta(2\phi^* - 1) < 4\Delta$. Second, since $\eta > \xi > \Gamma_{\text{sum}} \geq 1/2$ and $|\psi''|$ is decreasing in γ (the denominator $[\lambda\gamma + (1-\lambda)(1-\gamma)]^3$ is increasing): $|\psi''(\eta)| \leq |\psi''(1/2)| = 16\lambda(1-\lambda)(2\lambda-1)$. Third, $\gamma^* - \xi \leq \gamma^* - \frac{1}{2} < \frac{1}{2}$ (since $\gamma^* < 1$ by Proposition 1). Hence

$$|\gamma'| \cdot |\psi''(\eta)| \cdot (\gamma^* - \xi) < 4\Delta \cdot 16\lambda(1-\lambda)(2\lambda-1) \cdot \tfrac{1}{2} = 32\Delta\lambda(1-\lambda)(2\lambda-1).$$

Since $\Delta = (1-\lambda)[\gamma(L) - \gamma(\mu)] < (1-\lambda)/2$ (because $\gamma(L) < 1$ and $\gamma(\mu) \geq 1/2$):

$$32\Delta\lambda(1-\lambda)(2\lambda-1) < 16\lambda(1-\lambda)^2(2\lambda-1).$$

The right-hand side is the product of four positive terms λ, $(1-\lambda)$, $(1-\lambda)$, and $(2\lambda-1)$ times 16. Their sum is $1+\lambda \leq 2$, so by the AM–GM inequality their product is at most $((1+\lambda)/4)^4 \leq (1/2)^4 = 1/16$, and the inequality is strict because equality in AM–GM requires all factors to be equal, which would entail $\lambda = 1-\lambda$ and $\lambda = 2\lambda - 1$ simultaneously – a contradiction. Hence $16\lambda(1-\lambda)^2(2\lambda-1) < 16 \cdot (1/16) = 1$, and $dM/d\alpha\big|_{\alpha=0} < 0$.[39] □

Proof of Corollary 3. An entrant at $\phi_0 = \lambda$ has average score $\bar{s}_0 = \lambda$ (since $\bar{s}_A = \phi_A$). An incumbent at ϕ^* has score $\phi^* > \lambda$. The gap is $\phi^* - \lambda = (1-m^*) - \lambda = (1-\lambda) - m^*$. Substituting $m^* \approx (1-\bar{\gamma})/(R + 4\Delta)$ from Proposition 2 gives the stated expression.

The gap is increasing in μ (through $\bar{\gamma}$): by Proposition 4(i), increasing μ raises ϕ^*, which directly increases $\phi^* - \lambda$. In low-overlap markets (μ small, $\bar{\gamma}$ low), even the incumbent has a relatively diverse buyer pool, so the score advantage is smaller; the barrier is strongest in high-overlap markets where the incumbent's fan base is highly homogeneous. □

[39] For $k > L$, the summary includes $\mathcal{F}_b$-exclusive features, and for $k = 2L - \mu$, $\Gamma_{\text{sum}} = \hat{\gamma} > \gamma^*$: a "summary" that reproduces all features strictly improves on the decentralized equilibrium. But such a summary is the full review text, and the interesting regime is $k \ll L$.