

Social Learning with Limited Attention: Negative Reviews Persist under Newest First

Jackie Baek*

Atanas Dinev†

Thodoris Lykouris‡

First version: June 2024

Current Version: May 2025§

Abstract

We study a model of *social learning from reviews* where customers are computationally limited and make purchases based on reading only the first few reviews displayed by the platform. Under this limited attention, we establish that the *review ordering* policy can have a significant impact. In particular, the popular *Newest First* ordering induces a negative review to persist as the most recent review longer than a positive review. This phenomenon, which we term the *Cost of Newest First*, can make the long-term revenue unboundedly lower than a counterpart where reviews are exogenously drawn for each customer.

We show that the impact of the *Cost of Newest First* can be mitigated under dynamic pricing, which allows the price to depend on the set of displayed reviews. Under the optimal dynamic pricing policy, the revenue loss is at most a factor of 2. On the way, we identify a structural property for this optimal dynamic pricing: the prices should ensure that the probability of a purchase is always the same, regardless of the state of reviews. We also consider a setting where product quality evolves over time according to a Markov chain; we find that Newest First better tracks current quality but still leads to lower revenue, highlighting a trade-off between customer belief accuracy and revenue. Finally, we support our theoretical findings with numerical simulations and an empirical analysis on reviews from Tripadvisor.

*Stern School of Business, New York University, baek@stern.nyu.edu

†Massachusetts Institute of Technology, adinev@mit.edu

‡Massachusetts Institute of Technology, lykouris@mit.edu

§An extended abstract appeared at the ACM Conference on Economics and Computation (EC) 2024.

1 Introduction

The use of product reviews to inform customer purchase decisions has become ubiquitous in a variety of online platforms, ranging from electronic commerce to accommodation and recommendation platforms. While the online nature of such platforms may hinder the ability of customers to confidently evaluate the product compared to an in-person experience, reviews written by previous customers can shed light on the product’s quality. It is well established that product reviews play a significant role on customer purchase decisions [CM06, ZZ10, Luc16].

The process in which reviews impact product purchases can be seen as a problem of *social learning*, which generically studies how agents update their beliefs for an unknown quantity of interest (e.g., product quality) based on observing actions of past agents (e.g., reading reviews written by past customers). The typical assumption in the literature of social learning with reviews is that, when deciding whether to purchase a product, customers consider either *all* reviews provided by previous customers [CIMS17, IMSZ19, GHKV23] or a summary statistic such as their average rating [BS18, CLT21, AMMO22]. The motivation for the latter assumption is that customers have limited time and computational power and thus rely on a summary statistic, often provided by the platform (see Section 1.3 for a further discussion on these lines of work).

However, in practice, a common scenario may be somewhere “in between” the above two assumptions: customers read a small number of reviews in detail. Existing works have found that the textual content of a review contains important information that goes beyond its numeric score and such information can heavily influence purchase decisions [GI10, AGI11, LDRF⁺13, LLS19, VL21, LYMZ22]. Therefore, customers look beyond the average review rating and read a small number of reviews in detail. In particular, [Kav21] finds that 76% of customers read between 1 and 9 reviews before making a purchase. This motivates the main questions of our work:

*When customers read a limited number of reviews, how does this impact social learning?
Are there operational decisions that should be reconsidered due to this limited attention?*

We study a model that builds on existing models of social learning, where we incorporate the behavior that customers read a limited number of reviews. This simple change has two key implications. First, we show that the *ordering* of those reviews plays a critical role in impacting social learning and revenue. This is in contrast to existing models, where customers read all reviews or rely solely on a summary statistic, in which the review ordering has no effect. Our work thus introduces a new operational lever and enables us to compare review ordering policies. Second, unlike models in the literature where customer beliefs eventually converge, beliefs in our model do not stabilize but rather *fluctuate indefinitely* because customers read only a small number of reviews. This provides a fundamentally different perspective on social learning when customers read a limited number of reviews.

Concretely, we study a model for a single product (formalized in Section 2), where a platform makes decisions regarding how reviews are *ordered* and how the product is *priced*. Customers arrive sequentially and each customer takes only the first c reviews into account to inform their purchase decision, where c is a small constant. We assume that the t ’th customer’s valuation can be decomposed as the sum of a) an idiosyncratic valuation Θ_t that is known to them and b) a product quality X_t that has a fixed mean μ ; the latter quantity is unknown to the customer and can only be inferred via the reviews. The randomness of X_t reflects variability in the service quality, product defects, or exogenous factors affecting the customer experience [CIMS17]. We assume that when a customer reads a review written by customer s , they observe X_s (see Section 1.3 for a discussion

of this assumption). Each customer uses c reviews to update their belief about μ , and makes a purchase if their estimate of their valuation is higher than the price. In the event of a purchase, they leave a review that future customers can read.

1.1 Our contribution

A popular review ordering policy is to display reviews in reverse chronological order (newest to oldest); we refer to this policy as σ^{NEWEST} . This is the default option in platforms such as Airbnb, Tripadvisor and Macy’s¹ as it allows customers to get access to the most up-to-date reviews. In the context of our model, under the σ^{NEWEST} ordering, a customer considers the c most recent reviews. The set of these c reviews evolves as a *stochastic process* over customer arrivals: when a new purchase happens and thus a new review is provided, this review replaces the c -th most recent review.

Cost of Newest First. By analyzing the steady state of the aforementioned stochastic process, we observe that the σ^{NEWEST} ordering policy induces an undesirable behavior where negative reviews are read more than positive reviews, leading to a significant loss in overall revenue (Section 3).

To illustrate this phenomenon, consider a simple setting where customers only read the first review ($c = 1$) and the probability of a purchase is higher when the review is positive. When the t ’th customer arrives, if the first review is positive, this customer is more likely to buy the product and subsequently leave a review; the new review from the t ’th customer then becomes the “first review” for the $(t + 1)$ ’th customer. On the other hand, if the first review is negative for the t ’th customer, they are less likely to buy the product, and hence the same negative review remains as the “first review” for the $(t + 1)$ ’th customer. Therefore, negative reviews persist under the newest first ordering: a review will stay longer as the first review if it is negative compared to if it is positive.

We show that this arises due to the *endogeneity* of the stochastic process that σ^{NEWEST} induces and results in a loss of long-term revenue. To formalize this notion, we compare σ^{NEWEST} to an *exogenous process* where each arriving customer sees an independently drawn random set of reviews; we refer to this ordering policy as σ^{RANDOM} . We establish that the long-term revenue under σ^{NEWEST} is strictly smaller than that of σ^{RANDOM} under any non-degenerate instance (Theorem 3.1) and that the revenue under σ^{NEWEST} can be arbitrarily smaller, in a multiplicative sense, compared to σ^{RANDOM} (Theorem 3.2). We refer to this phenomenon as the COST OF NEWEST FIRST (CoNF).

Dynamic pricing mitigates the impact of CoNF. Seeking to mitigate this phenomenon, we consider the impact of optimizing the product’s price (Section 4). We show that even under the *optimal* static price, the CoNF remains arbitrarily large (Theorem 4.1). However, if we allow for dynamic pricing, where the price can depend on the state of the reviews, we show that the CoNF is upper bounded by a factor of 2 when the idiosyncratic valuation distribution is non-negative and gracefully decays with its negative mass otherwise (Theorem 4.2).

This improvement stems from the fact that dynamic pricing allows us to *change the steady-state distribution* of the stochastic process. Recall that, under σ^{NEWEST} , the stochastic process spends more time on states with negative reviews than states with positive reviews. The optimal dynamic

¹This statement is based on access on May 14, 2025. Many other platforms such as Amazon and Yelp list NEWEST as the second default and have their own ordering mechanism as the default option.

pricing policy sets prices so that the purchase probability is *equal* across all states (Theorem 4.3) — this ensures that the steady-state distribution under σ^{NEWEST} is the same as that of σ^{RANDOM} .

A broader implication of this result is that, when purchase decisions depend on the state of the first reviews, platforms that offer state-dependent prices can be arbitrarily better off than platforms that are unaware of this phenomenon and statically optimize prices (Corollary 4.2).

Non-stationarity in product quality. One potential benefit of Newest First is that newer reviews are more representative of current product quality, when the quality changes over time. To formalize this, we define a model with non-stationarity where the product quality changes over time according to a Markov chain (Section 5). We define a *belief error* metric that represents the difference between the customer’s posterior belief and the current product quality. Indeed, the belief error is lower under Newest First compared to Random, where the gap is higher when the level of non-stationarity is low (i.e., quality switches are less frequent). However, we show that the CoNF still exists under non-stationarity; revenue is lower under Newest First compared to Random. This introduces a trade-off between revenue and belief error. Revenue is always higher under Random while the belief error is better (lower) under Newest First, but the benefit of Newest First for belief error diminishes when quality switches are more frequent. Therefore, when the level of non-stationarity is high, the benefit of using Newest First diminishes.

Numerical evaluations. Lastly, we conduct three sets of numerical evaluations to support our theoretical findings (Section 6). First, we simulate our model on a finite horizon to test the impact of the initial condition, as our theoretical results rely on an infinite horizon. Second, we numerically compare the CoNF for a set of instances under static and dynamic pricing, as our theoretical results on the CoNF provide worst-case bounds. Lastly, we leverage real review data from Tripadvisor, an online platform where the default review ordering policy is Newest First. We evaluate 109 hotel pages, and we find that for 79 of them, the average review rating of the first 10 reviews is lower than the average review rating of all reviews. These numerical evaluations provide robustness checks and support our main theoretical findings.

1.2 Key Behavioral Assumption and Managerial Implications

In this paper, we posit that customers have *limited attention*, i.e., they read a limited number of the top-ranked reviews (see Section 1.3 for a discussion on this assumption). Under this behavioral assumption, our paper sheds light to three novel managerial implications:

1. **Impact of review ordering.** Our paper identifies that if customers only read a limited number of the top-ranked reviews, then the *ordering* of reviews becomes an important operational lever. This lever was previously overlooked by the literature as past works did not model customers’ limited attention. Under this assumption, we identify and characterize the *Cost of Newest First* phenomenon, that presenting reviews in a Newest First order leads to the persistence of negative reviews, which negatively impacts customer beliefs and consequently reduces platform revenue.
2. **Incorporate reviews in dynamic pricing algorithms.** We find that dynamic pricing strategies mitigate the Cost of Newest First. Theoretically, the negative impact is bounded

by a factor of two, and our numerical experiments show that this effect is often negligible. Practically speaking, platforms already widely employ sophisticated and highly dynamic pricing algorithms [Pat17]. Our results imply that the *current set of reviews* is an important feature that platforms should incorporate in their pricing algorithms, alongside other important factors such as competitors’ prices, demand forecasts, seasonal trends, and others.

3. **Cost of Newest First under non-stationarity.** One potential benefit of Newest First is that newer reviews are more representative of current product quality. We show that indeed, when product quality changes over time, then the customer’s belief of product quality is on average more accurate (closer to the current product quality) under Newest First compared to Random. However, even in the non-stationary setting, revenue is lower under Newest First, hence there is a trade-off between revenue and accuracy of beliefs. Under Newest First, the accuracy of beliefs decreases as the level of non-stationarity increases (product quality changes more frequently). Therefore, the benefit of Newest First is limited when changes in the product quality are frequent.

1.3 Related Work and Comparison of Key Modeling Assumptions

Social learning and incentivized exploration. Classical models of social learning from [Ban92] and [BHW92] study a setting in which there is an unknown state of the world and each agent observes an independent, noisy signal about the state as well as the actions of past agents. The agent uses this information to update their beliefs and then takes an action. In this setting, undesirable “herding” behavior can arise: agents may converge to taking the wrong action. Conceptually closer to our work, [Say18] shows that dynamic pricing can mitigate the aforementioned herding behavior. Subsequent works study how social learning is affected by the agent’s signal distribution [SS00], prior for the state [CDO22], heterogeneous preferences [GPR06, LS16], as well as the structure of their observations [AMMO22]. From a different perspective, there is a stream of literature that aim to design mechanisms to *help* the learning process, either by modifying the information structure [KMP14, MSS20, BPS18] or by incentivizing exploration through payments [FKKK14, KKM⁺17].

Social learning with reviews. Closer to our work, several papers focus on the setting where customers learn about a product’s quality through reviews [HPZ17, CIMS17, BS18, SR18, IMSZ19, CLT21, AMMO22, GHKV23, Bon23, CSS24]. This literature induces several modeling differences compared to classical social learning. First, agents do not receive independent signals of the unknown state (the product quality). Second, agents not only observe the binary purchase decision of previous agents, but also the reviews of previous agents who purchased the product. We highlight the key modeling assumptions of our work and how they relate to existing works.

No self-selection bias. In prior works of social learning with reviews, the main difficulty stems from the *self-selection bias*, the idea that only customers who value the product highly will buy the product and hence these customers leave reviews with higher ratings. In the presence of self-selection bias, [CIMS17], [BS18] and [IMSZ19] study conditions in which customer beliefs eventually successfully learn the quality of a product, where customers update their beliefs based on the entire history of reviews. [BS18, SR18, CLT21, AMMO22] consider models in which customers only incorporate summary statistics of prior reviews (e.g., average rating) into their beliefs. [Bon23] analyzes how the magnitude of the self-selection bias depends on the product’s quality and polarization. [CSS24] consider a model where the platform’s pricing decision affect the review ratings

and characterizes the impact of the price on the average rating. This is also empirically supported by [BMZ12] which shows that Groupon discounts lead to lower ratings. [HPZ17] study a two-stage model which quantifies both self-selection bias and under-reporting bias (reviews are provided only by customers with extreme experiences); see references within for further related work.

In contrast, our work studies a model where self-selection bias does not arise. Specifically, we assume that customer t 's valuation can be decomposed as $V_t = \Theta_t + X_t$, where Θ_t is customer-specific and X_t has a fixed mean μ shared across customers. The quantity μ is the unknown quantity of interest for all customers. Our model assumes that a review reveals X_t . In contrast, prior works assume that a review reveals $\Theta_t + X_t$ and one cannot separate the contribution from each term. This means that, in our model, the customer-specific valuation and the pricing decision affect the purchase probability but do not affect the review itself conditioning on a purchase. Although our assumption makes it “easier” for customers to learn μ , we study a new phenomenon that arises due to the fact that customers only read a small number of reviews.

The practical motivation for our modeling assumption of no selection bias is the following. On most online platforms, a review is composed by both a numeric score (e.g., 4 out of 5) and a textual description that further explains the reviewer’s thoughts. Within our model, one interpretation is that the numeric score reveals $\Theta_t + X_t$, but one can use the textual content of the review to separate Θ_t from X_t . Therefore, we assume that reading the text of the review reveals X_t , but we also assume that each customer only reads a small number of reviews since reading the text takes time.

[GHKV23] also studies a setting with no self-selection bias (without limited attention) with dynamic pricing. Their model assumes that customers are partitioned into a finite number of types and only read reviews written by customers of the same type. This overcomes self-selection bias as customers of the same type can be thought of as having the same value of Θ_t in our model.

Individual reviews vs. average rating. Most online platforms show the average rating of all reviews for each product, as well as individual reviews. In our model, the average rating is captured through the customers’ *prior belief* about the product quality μ , which captures any information customers have before reading individual reviews (e.g., average rating, product features, etc.). The question we study is: conditioning on all other available information, how does the review ordering policy affect purchasing decisions? For example, if a customer starts with a high prior due to a 4.8 out of 5 average rating but still reads a few reviews, the order in which those reviews are presented can influence their final belief and purchase decision. Our paper therefore speaks directly to the incremental effect of review ordering, holding other information constant. Existing empirical studies support the assumption that customers read individual reviews, and that this can have a significant impact on purchasing behavior [AGI11, LLS19, VL21, LYMZ22]. As [AGI11] emphasize, product quality is multi-dimensional, and consumers have heterogeneous preferences over these dimensions—details that are often lost in average ratings. For example, one consumer might take seriously a negative review that criticizes a laptop’s screen being too large, while another might disregard the review entirely because they prefer a larger display. [LYMZ22] studies the relative impact of individual reviews and the average rating, and they find a swaying effect of individual reviews. [VL21] and [LLS19] claim that the *top* reviews have a substantial effect on customer behavior, which highlights the importance of review ordering.

Limited attention in operations. Our paper also connects to a growing literature in operations that models customer decision-making under limited attention in various operational contexts. For

example, [AHB19] and [BA21] model consumers who face costs in acquiring information about products, leading to non-standard choice behavior and strategic pricing by firms. [FMMP19] and [GL24] model limited consideration sets, where consumers only attend to a subset of available products when making purchasing decisions. [GMC⁺21] model impatient customers who sequentially view product recommendations and may stop searching once a product surpasses their threshold utility. Our model complements this literature by applying similar behavioral assumptions of limited attention, but in the context of customers reading reviews rather than evaluating products directly.

Belief convergence vs. stochastic process. In the existing literature, social learning is deemed “successful” if the customer’s estimate of product quality converges to the true quality. This convergence can either be that their belief distribution converges to a single point [IMSZ19, AMMO22], or that the customer’s scalar estimate of the product quality (e.g., average rating) converges to the true quality [CIMS17, BS18]. In our setting, customers update their beliefs based on the first c reviews and these c reviews evolve across customers as a stochastic process. Therefore, customer beliefs do not converge but rather oscillate based on the state of those c reviews, even as the number of customers goes to infinity.

Closer to our work, [PSX21] study a similar model (without limited attention) and show that the initial review can have an effect on the proclivity of customer purchases and the number of reviews. This bias introduced by initial reviews is also empirically observed by the work of [LKAK18]. Unlike our model, this effect diminishes over time as the product acquires more reviews and the initial review becomes less salient. Our results on the COST OF NEWEST FIRST can thus be viewed as a stronger version of the result in [PSX21]: we show that, in the presence of limited attention, the effect of negative reviews persists even in the steady-state of the system.

Fully Bayesian vs. non-Bayesian. Existing papers differ in whether customers incorporate information from reviews in a fully Bayesian or non-Bayesian manner. For example, [IMSZ19] and [AMMO22] study a fully Bayesian setting where all distributions (prior on μ , distribution of Θ_t) and purchasing behaviors are common knowledge and each customer forms a posterior belief on μ using the information given to them. In contrast, [CIMS17] and [CLT21] assume that customers use a simple non-Bayesian rule when making their purchase decision. Moreover, [BS18] study both Bayesian and non-Bayesian update rules and compare them.

In our paper, customers use a Bayesian *framework* but their update rule is not *fully* Bayesian. Specifically, customers start with a Beta prior on μ . This prior need not be correct as we assume that μ is a fixed number. We assume reviews are binary (0 or 1), and customers read c of them and update their beliefs *assuming* that these reviews are independent draws from $Bernoulli(\mu)$. However, this is not necessarily the “correct” Bayesian update rule for the customers, due to the *endogeneity* of the stochastic process of the first c reviews. That is, under the σ^{NEWEST} ordering rule, negative reviews are more likely to persist as the most recent review — hence even if $\mu = 1/2$, the most recent review is more likely to be negative than positive. Therefore, a fully Bayesian customer should take this phenomenon into account when updating their beliefs. We assume that customers do *not* account for this (and hence are not *fully* Bayesian), and we study the impact of how this endogeneity impacts the steady state of the process. Our model also implicitly assumes that customers do not use additional information about previous customers’ non-purchase decisions (which are typically non-observable) and the platform’s pricing policy (which is often opaque).

Finally, our model is flexible in that it allows customers to map their belief distribution to a scalar estimate of μ in an arbitrary manner, e.g., the mean of the belief distribution (which is considered in most prior work) as well as a pessimistic estimate thereof (as studied in [GHKV23]).

Other work on social learning with reviews. [HC24] studies the design of rating systems motivated by the idea that older reviews become less relevant. In a setting where the product’s quality changes, they show that a moving average rating system is optimal in reflecting the true quality. Social learning with reviews has also been studied for ranking [MSSV23], incorporating quality variability [DLT21], dealing with non-stationary environments [BPS22], and has been applied to green technology adoption [RHP23].

2 Model

We consider a platform that repeatedly offers a product to customers that arrive in consecutive rounds $t = 1, 2, \dots$. The customer makes a purchase decision based on a finite number of reviews and the price; if a purchase occurs, they leave a new review for the product. We consider the platform’s decisions regarding the *ordering of reviews* as well as the *price*.

Customer valuation. The customer at round t has a realized valuation $V_t = X_t + \Theta_t \in \mathbb{R}$ for the product, where X_t and Θ_t represent the contribution from the product’s unobservable and observable parts respectively. Specifically, when customer t makes a purchase, $X_t \in \{0, 1\}$ is drawn independently from $Bernoulli(\mu)$ where $\mu \in (0, 1)$ is the same across all customers and is unknown to them. Contrastingly, the quantity Θ_t is customer-specific and is known to customer t before they purchase. We assume that, at every round t , Θ_t is drawn independently from a distribution \mathcal{F} with bounded support. The platform knows the distribution \mathcal{F} but not Θ_t .

If the customer at round t knew their exact valuation V_t , then they would purchase the product if and only if $V_t \geq p_t$, where p_t is the price of the product at time t . However, X_t is unknown and hence so is V_t . We assume that the customers read reviews to learn about μ , and their purchase decision depends on their belief about μ after reading the reviews. Note that customers cannot aim to estimate X_t beyond estimating μ , since X_t is drawn independently for each customer.

Review generation. If the customer at round t purchases the product, they write a review that future customers may read. The review rating given by customer t is equal to X_t (see Section 1.3 for a discussion of this assumption). We often refer to a review with $X_t = 1$ as *positive* and to a review with $X_t = 0$ as *negative*. We refer to $X_t \in \{0, 1, \perp\}$ as the *rating of customer t ’s review* where $X_t = \perp$ if customer t did not purchase the product.

2.1 Customer Purchase Behavior

We describe the customer purchase behavior at one round, taking the price and the review ordering as fixed. Customers have a prior $\text{Beta}(a, b)$ for the value of μ , for some fixed $a, b > 0$. This prior need not be correct and could be based on information about the features of the product or summary statistics of all reviews which are subject to self-selection bias (see discussion in Section 1.3). Customers read the first c reviews that are shown to them to update their prior. Formally, letting $\mathbf{Z}_t = (Z_{t,1}, \dots, Z_{t,c}) \in \{0, 1\}^c$ denote the ratings of the first c reviews shown, the customer creates

the following posterior for the unobservable quality μ :

$$\Phi_t := \text{Beta} \left(a + \sum_{i=1}^c Z_{t,i}, b + c - \sum_{i=1}^c Z_{t,i} \right).$$

This corresponds to the natural posterior update for μ if each $Z_{t,i}$ is an independent draw from $\text{Bernoulli}(\mu)$.² Note that the customer places equal weight on the first c reviews. Based on this posterior, the customer creates an estimate \hat{V}_t for their valuation. We assume that there is a mapping $h(\Phi_t) \in (0, 1)$ from their posterior to a real number that represents an estimate of the fixed valuation μ . For example, the mapping $h(\Phi_t) = \mathbb{E}[\Phi_t]$ represents risk-neutral customers, while if $h(\Phi_t)$ corresponds to the ϕ -quantile of Φ_t for $\phi < 0.5$, this represents pessimistic customers [GHKV23]. The customer then forms their estimated valuation $\hat{V}_t := \Theta_t + h(\Phi_t)$ and buys the product at price p_t if and only if $\hat{V}_t \geq p_t$. Finally, the customer leaves a review $X_t \sim \text{Bernoulli}(\mu)$ if they bought the product, otherwise $X_t = \perp$.

To ease exposition, we often use n to denote the number of positive ratings (i.e., $n = \sum_{i=1}^c Z_{t,i}$), and we overload notation to denote by $h(n)$ to refer to $h(\text{Beta}(a + n, b + c - n))$. We make the natural assumption that higher number of positive ratings leads to a higher purchase probability.

Assumption 2.1. *The estimate $h(n)$ is strictly increasing in the number of positive reviews n .*

We also assume that the idiosyncratic valuation has positive mass on non-negative values.

Assumption 2.2. *The distribution \mathcal{F} has positive mass on non-negative values: $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq 0] > 0$.*

We denote the above problem instance as $\mathcal{E}(\mu, \mathcal{F}, a, b, c, h)$ for product quality μ , prior parameters a, b , idiosyncratic distribution \mathcal{F} , customers' attention budget c , and an estimate mapping h .

2.2 Platform Decisions

We consider two platform decisions, review ordering and pricing.

Review ordering. With respect to ordering, since customers only take the first c reviews into account, choosing an ordering is equivalent to selecting a set of c reviews to show. Let $I_t = \{\tau < t : X_\tau \neq \perp\}$ be the set of previous rounds in which a review was submitted and let $\mathcal{H}_t = \{I_t, \{X_\tau\}_{\tau < t}, \{p_\tau\}_{\tau < t}, \{\mathbf{Z}_\tau\}_{\tau < t}\}$ be the observed history before round t . At round t , the platform maps (possibly in a randomized way) its observed history \mathcal{H}_t to the set of review ratings $\sigma(\mathcal{H}_t)$ corresponding to the c reviews shown. We study the steady-state distribution of the system; to avoid initialization corner cases, we assume that at time $t = 1$, there is an infinite pool of reviews $\{X_\tau\}_{\tau=-\infty}^{-1}$ where $X_\tau \stackrel{\text{i.i.d.}}{\sim} \text{Bernoulli}(\mu)$. We consider the following review ordering policies:

- σ^{NEWEST} selects the c newest reviews. This is formally defined as $\sigma^{\text{NEWEST}}(\mathcal{H}_t) = (Z_{t,i})_{i=1}^c$ where $Z_{t,i}$ is the rating of the i -th most recent review.
- σ^{RANDOM} shows c random reviews. This corresponds to ratings being drawn independently from $\text{Bernoulli}(\mu)$ at each round t ; i.e., $\sigma^{\text{RANDOM}}(\mathcal{H}_t) = (Z_{t,i})_{i=1}^c$ where $Z_{t,i} \stackrel{\text{i.i.d.}}{\sim} \text{Bernoulli}(\mu)$. We use σ^{RANDOM} as a benchmark because it maximizes revenue across all policies that only consider the relative arrival sequence and are agnostic to the exact time and rating of reviews (see Appendix E).

²The reviews \mathbf{Z}_t are not necessarily independent draws from $\text{Bernoulli}(\mu)$, hence the customers are not completely Bayesian. See the second-to-last point in Section 1.3 for a detailed discussion.

Note that, under σ^{RANDOM} , the customers' and platform's actions at the current round do not influence the reviews shown in future rounds (i.e. the reviews are exogenous). In contrast, under σ^{NEWEST} , the customers' and platform's actions influence what reviews are shown in future rounds (i.e. the reviews are endogenous to the underlying stochastic process).

Pricing. The platform also decides on the pricing policy, where the price at each round can depend on the set of c displayed reviews. We denote a pricing policy by a function $\rho : \{0, 1\}^c \rightarrow \mathbb{R}$, which maps the set of displayed review ratings $\mathbf{z} \in \{0, 1\}^c$ to a price. We study two classes of pricing policies: static and dynamic. We let Π^{STATIC} be the set of pricing policies ρ that assign a fixed price p , i.e., $\rho(\mathbf{z}) = p$ for any \mathbf{z} . Similarly, Π^{DYNAMIC} includes the set of pricing policies ρ where $\rho(\mathbf{z})$ can depend on the review ratings \mathbf{z} .

2.3 Revenue and the Cost of Newest First

For an ordering policy σ and pricing policy ρ , we define the revenue as the steady-state revenue:³

$$\text{REV}(\sigma, \rho) := \liminf_{T \rightarrow \infty} \mathbb{E} \left[\frac{\sum_{t=1}^T \rho(\mathbf{Z}_t) \mathbf{1}_{\Theta_t + h(\Phi_t) \geq \rho(\mathbf{Z}_t)}}{T} \right]. \quad (1)$$

Our main focus lies in understanding the effect of the ordering policy on the revenue. Specifically, we compare the revenues of the ordering policies σ^{NEWEST} and σ^{RANDOM} . For a pricing policy ρ , we define the COST OF NEWEST FIRST (CONF) as (for all policies ρ we consider, $\text{REV}(\sigma^{\text{NEWEST}}, \rho) > 0$):

$$\chi(\rho) := \frac{\text{REV}(\sigma^{\text{RANDOM}}, \rho)}{\text{REV}(\sigma^{\text{NEWEST}}, \rho)}.$$

If $\rho \in \Pi^{\text{STATIC}}$, i.e., $\rho(\mathbf{z}) = p$ for any $\mathbf{z} \in \{0, 1\}^c$ we use $\text{REV}(\sigma, p)$ and $\chi(p)$ as shorthand for the corresponding steady-state revenue and CONF.

In Section 4, we study the CONF when the platform can optimize its pricing policy over a class of policies. For a class of pricing policies Π , we define similarly the optimal revenue within-class with respect to an ordering policy σ and the corresponding CONF as:

$$\text{REV}(\sigma, \Pi) := \sup_{\rho \in \Pi} \text{REV}(\sigma, \rho) \quad \text{and} \quad \chi(\Pi) := \frac{\text{REV}(\sigma^{\text{RANDOM}}, \Pi)}{\text{REV}(\sigma^{\text{NEWEST}}, \Pi)} \quad \text{respectively.}$$

To ease exposition, we make the mild assumption that $0 < \text{REV}(\sigma, \Pi) < \infty$.⁴

Finally, we note that, unlike classical revenue maximization works, our focus is *not* on identifying policies that maximize revenue but rather in comparing the performance of different ordering policies (σ^{NEWEST} vs σ^{RANDOM}) and different pricing policies (Π^{STATIC} vs Π^{DYNAMIC}).

Remark 2.1. *Our assumption of Bernoulli reviews and Beta prior is made for ease of exposition. Our results extend to a more general model where reviews come from an arbitrary distribution with finite support (not only $\{0, 1\}$) and the estimator h arbitrarily maps reviews to an estimate for the fixed valuation μ ; see Appendix A for details.*

³The policies we consider have a stationary distribution so, in our analysis, we replace the \liminf with a \lim .

⁴For the policies we consider, $\text{REV}(\sigma, \Pi) > 0$ by Assumption 2.2 and the fact that $h(0) > 0$; $\text{REV}(\sigma, \Pi) < \infty$ is satisfied if the maximum revenue from a customer's idiosyncratic valuation is finite, i.e., $\sup_p p \mathbb{P}_{\Theta \sim \mathcal{F}}(\Theta \geq p) < +\infty$.

3 Cost of Newest First (CoNF) with a Fixed Static Price

Throughout this section, we assume a static pricing policy where the price p is fixed and given. We establish the main phenomenon of the COST OF NEWEST FIRST by showing that $\chi(p) > 1$ under very mild conditions on the price p . We then show that $\chi(p)$ can be arbitrarily large.

Recall that $h(n)$ refers to $h(\text{Beta}(a+n, b+c-n))$, where $n \in \{0, \dots, c\}$ is the number of positive reviews. We first introduce an assumption on two natural conditions that the price satisfies.

Assumption 3.1. *For our results on a fixed price p , we assume that $p > 0$ and that it is:*

1. Non-absorbing: *if the purchase probability is positive for any displayed review ratings; i.e., for all $n \in \{0, 1, \dots, c\}$, i.e., $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(n) \geq p] > 0$.*
2. Non-degenerate: *if the purchase probability given all negative review ratings is strictly smaller than given all positive review ratings, i.e., $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(0) \geq p] < \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(c) \geq p]$.*

A non-absorbing price guarantees that σ^{NEWEST} does not get “stuck” in a zero-revenue state. A non-degenerate price implies that the review ratings *matter*, since there exist distinct review ratings where the purchase probability differs.

3.1 Existence of Cost of Newest First

Our main result is that the revenue under σ^{RANDOM} is strictly higher than that of σ^{NEWEST} ; that is, $\chi(p) > 1$. As a building block towards this result, we first provide simple and interpretable closed form expressions for $\text{REV}(\sigma^{\text{RANDOM}}, p)$ and $\text{REV}(\sigma^{\text{NEWEST}}, p)$ for a static price p , which are given in Propositions 3.1 and 3.2 respectively. Let $\text{Bern}(\mu)$ denote the Bernoulli distribution with success probability μ and $\text{Binom}(c, \mu)$ denote the Binomial distribution with c i.i.d. $\text{Bern}(\mu)$ trials.

Proposition 3.1 (Revenue of σ^{RANDOM}). *For any fixed price p ,*

$$\text{REV}(\sigma^{\text{RANDOM}}, p) = p \mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N) \geq p] \right].$$

Proof. By definition, σ^{RANDOM} displays c i.i.d. $\text{Bern}(\mu)$ reviews at every round t . As a result, the number of positive reviews is distributed as $\text{Binom}(c, \mu)$, yielding expected revenue, at every round t , equal to the right hand side of the theorem. Given that this quantity does not depend on t , recalling Eq.(1), it equals the steady-state revenue. \square

Unlike σ^{RANDOM} which displays c i.i.d. $\text{Bern}(\mu)$ reviews at every round, the reviews displayed by σ^{NEWEST} are an endogenous function of the history. The proof of the next result underlies the technical crux of this section and is presented in Section 3.2.

Proposition 3.2 (Revenue of σ^{NEWEST}). *For any fixed price p satisfying Assumption 3.1,*

$$\text{REV}(\sigma^{\text{NEWEST}}, p) = \frac{p}{\mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N) \geq p]} \right]}.$$

Intuitively, when the newest reviews are positive, the customer is more likely to buy the product and leave a new review, which then updates the set of newest reviews. On the other hand, when

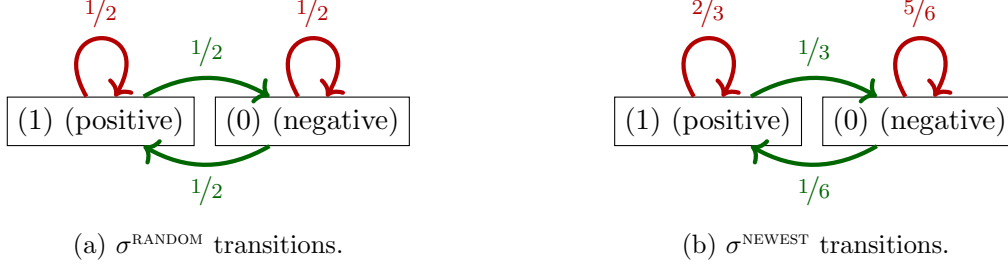


Figure 1: The state transitions in the instance of the Example 3.1 under σ^{RANDOM} and σ^{NEWEST} .

the newest reviews are negative, the customer is less likely to buy, and hence the set of newest reviews is less likely to be updated. This implies that σ^{NEWEST} spends more time in a state with negative reviews (which yield lower revenue) compared to σ^{RANDOM} . This phenomenon is the driver of our main result and we refer to it as the COST OF NEWEST FIRST (CoNF).

Theorem 3.1 (COST OF NEWEST FIRST). *For any fixed price p satisfying Assumption 3.1, the revenue of σ^{NEWEST} is strictly smaller than that of σ^{RANDOM} , i.e., $\text{REV}(\sigma^{\text{RANDOM}}, p) > \text{REV}(\sigma^{\text{NEWEST}}, p)$.*

Proof. We show that the expression of Proposition 3.1 is higher than the one of Proposition 3.2. By Jensen's inequality, $\mathbb{E}[X] \geq \frac{1}{\mathbb{E}[\frac{1}{X}]}$ for any non-negative random variable X and equality is achieved if and only if X is a constant. Letting $X(N) := \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N) \geq p]$ be the purchase probability in a state with N positive reviews, we apply the inequality for $X = X(N)$,

$$\mathbb{E}_{N \sim \text{Binom}(c, \mu)} [\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N) \geq p]] \geq \frac{1}{\mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N) \geq p]} \right]}.$$

Multiplying with $p > 0$ on both sides we obtain that $\text{REV}(\sigma^{\text{RANDOM}}, p) \geq \text{REV}(\sigma^{\text{NEWEST}}, p)$. By Assumption 3.1, it holds that $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(0) \geq p] < \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(c) \geq p]$ and thus $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N) \geq p]$ is not a constant random variable when $N \sim \text{Binom}(c, \mu)$. Therefore, the inequality is strict. \square

We describe a simple example that provides intuition on the CoNF established in Theorem 3.1, which is also illustrated in Figure 1.

Example 3.1. Suppose $\mu = \frac{1}{2}$, $\mathcal{F} = \mathcal{U}[0, 1]$, $c = a = b = 1$, $h(\Phi_t) = \mathbb{E}[\Phi_t] = \frac{N+1}{3}$, and $p = 1$. Under σ^{RANDOM} the probability that the review shown is positive is 0.5 (see Figure 1a). Under σ^{NEWEST} the purchase probability is $\frac{2}{3}$ when the review is positive and $\frac{1}{3}$ when it is negative. Thus, under σ^{NEWEST} , transitioning from a positive to a negative review is twice as likely as transitioning from a negative to a positive review (see Figure 1b). Hence, a negative rating is twice as likely as positive rating. This leads to lower revenue in steady state for σ^{NEWEST} compared to σ^{RANDOM} .

Lastly, we generalize our result beyond revenue loss by analyzing the *number of positive reviews* among the c reviews. We show that the expected number of positive reviews is smaller under σ^{NEWEST} compared to σ^{RANDOM} (proof in Appendix B.1). Let π_n^{NEWEST} (respectively π_n^{RANDOM}) denote the steady-state probability of observing n positive reviews under σ^{NEWEST} (respectively σ^{RANDOM}).

Proposition 3.3. *For any price p satisfying Assumption 3.1, the average number of positive reviews under σ^{NEWEST} is smaller than under σ^{RANDOM} . Formally, $\mathbb{E}_{N \sim \pi_n^{\text{NEWEST}}}[N] < \mathbb{E}_{N \sim \pi_n^{\text{RANDOM}}}[N]$.*

3.2 Characterization of Revenue under Newest (Proof of Proposition 3.2)

We provide the proof of Proposition 3.2, which contains the main technical crux of this section. We first introduce some notation. Recalling that $Z_{t,i}$ denotes the rating of the i -th most recent review at round t , we refer to the c most recent reviews by the vector $\mathbf{Z}_t = (Z_{t,1}, \dots, Z_{t,c})$. We note that \mathbf{Z}_t is a time-homogeneous Markov chain with a finite state space $\{0, 1\}^c$. Given that we assume an infinite pool of initial reviews, $\mathbf{Z}_1 = (X_{-1}, \dots, X_{-c})$ where $X_i \sim \text{Bern}(\mu)$ for $i \in \{-1, \dots, -c\}$.

With respect to its transition dynamics of this Markov chain, for the state $\mathbf{Z}_t = (z_1, \dots, z_c) \in \{0, 1\}^c$, the purchase probability is $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(\sum_{i=1}^c z_i) \geq p]$. If there is no purchase, no review is given and the state remains $\mathbf{Z}_{t+1} = (z_1, \dots, z_c)$. If there is a purchase, \mathbf{Z}_t transitions to the state $\mathbf{Z}_{t+1} = (1, z_1, \dots, z_{c-1})$ if the review is positive (with probability μ) and to the state $\mathbf{Z}_{t+1} = (0, z_1, \dots, z_{c-1})$ if the review is negative (with probability $1 - \mu$).

Because the price p is non-absorbing (Assumption 3.1), for every state of reviews $(z_1, \dots, z_c) \in \{0, 1\}^c$, the purchase probability is positive, and the probability that a new review is positive is strictly greater than zero (since $\mu \in (0, 1)$). Then, \mathbf{Z}_t can reach every state from every other state with positive probability (i.e. it is a single-recurrence-class Markov Chain with no transient states), and hence \mathbf{Z}_t has a unique stationary distribution, which we denote by π . Our next lemma exactly characterizes the form of this stationary distribution.

Lemma 3.1. *The stationary distribution π of \mathbf{Z}_t under any price p satisfying Assumption 3.1 is*

$$\pi_{(z_1, \dots, z_c)} = \kappa \cdot \frac{\mu^{\sum_{i=1}^c z_i} (1 - \mu)^{c - \sum_{i=1}^c z_i}}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(\sum_{i=1}^c z_i) \geq p]},$$

where $\kappa = 1/\mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N) \geq p]} \right]$ is a normalizing constant.

Proof sketch. If the reviews were drawn i.i.d. at each round, the probability of state (z_1, \dots, z_c) would be exactly $\mu^{\sum_{i=1}^c z_i} (1 - \mu)^{c - \sum_{i=1}^c z_i}$, which is the numerator. However, the set of newest reviews is only updated when there is a purchase, which occurs with probability $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(\sum_{i=1}^c z_i) \geq p]$. Hence, we multiply the numerator by $1/\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(\sum_{i=1}^c z_i) \geq p]$, which is the expected number of rounds until there is a new review under state (z_1, \dots, z_c) ; in fact, we show such a property holds for general Markov chains (Lemma B.2). A formal proof is provided in Appendix B.2. \square

Equipped with Lemma 3.1, we now prove Proposition 3.2.

Proof of Proposition 3.2. Using Eq. (1), the revenue of σ^{NEWEST} can be written as

$$\begin{aligned} \text{REV}(\sigma^{\text{NEWEST}}, p) &= p \liminf_{T \rightarrow \infty} \frac{\mathbb{E} \left[\sum_{t=1}^T \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(\sum_{i=1}^c Z_{t,i}) \geq p] \right]}{T} \\ &= p \sum_{(z_1, \dots, z_c) \in \{0, 1\}^c} \pi_{(z_1, \dots, z_c)} \mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + h\left(\sum_{i=1}^c z_i\right) \geq p \right], \end{aligned}$$

where the second step expresses the revenue of the stationary distribution via the Ergodic theorem. Expanding $\pi_{(z_1, \dots, z_c)}$ based on Lemma 3.1, the $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(\sum_{i=1}^c z_i) \geq p]$ term cancels out and:

$$\text{REV}(\sigma^{\text{NEWEST}}, p) = p \cdot \kappa \cdot \left(\sum_{(z_1, \dots, z_c) \in \{0,1\}^c} \mu^{\sum_{i=1}^c z_i} (1 - \mu)^{c - \sum_{i=1}^c z_i} \right).$$

Note that the term in the parenthesis equals 1, since it is a sum over all probabilities of $\text{Binom}(c, \mu)$. This yields $\text{REV}(\sigma^{\text{NEWEST}}, p) = p \cdot \kappa$, which gives the expression in the theorem. \square

3.3 Cost of Newest First can be arbitrarily bad

Theorem 3.1 implies that, for any price p satisfying Assumption 3.1, the CONF is strictly greater than 1, i.e., $\chi(p) > 1$.⁵ We now show that it can be arbitrarily large. We first provide a closed-form expression for $\chi(p)$ by dividing the expressions in Propositions 3.1 and 3.2 (see Appendix B.3 for proof details).

Lemma 3.2. *For any price p satisfying Assumption 3.1, the CoNF is given by:*

$$\chi(p) = \sum_{i,j \in \{0, \dots, c\}} \mu^{i+j} (1 - \mu)^{2c-i-j} \binom{c}{i} \binom{c}{j} \frac{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(i) \geq p]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(j) \geq p]}.$$

Theorem 3.2. *For any continuous value distribution \mathcal{F} with positive mass on a bounded support, and any $M > 0$, there exists a price p satisfying Assumption 3.1 such that $\chi(p) > M$.*

Proof sketch. One summand in the right hand side of Lemma 3.2 contains the ratio of the purchase probability of all reviews being positive compared to all reviews being negative:

$$\beta(p) := \frac{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(c) \geq p]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(0) \geq p]},$$

which quantifies how much the reviews affect the purchase probability. Since all other terms are non-negative, the CONF is lower bounded by this summand, i.e., $\chi(p) \geq \mu^c (1 - \mu)^c \beta(p)$. The full proof (provided in Appendix B.4) shows that, for any M the latter quantity can become equal to M by an appropriate selection of price that makes Newest to be mostly “stuck” (purchase probability when all reviews are negative is close to 0 while the purchase probability when all reviews are positive is at least a constant). \square

We complement Theorem 3.2 by showing that if \mathcal{F} has Monotone Hazard Rate (MHR), i.e., letting F and f be its cumulative and density functions, its hazard rate function $\frac{f(u)}{1-F(u)}$ is non-decreasing in u , then $\chi(p)$ is non-decreasing in p (proof in Appendix B.5).

Proposition 3.4. *Suppose that \mathcal{F} is a continuous distribution with support $[\underline{\theta}, \bar{\theta}]$ and has MHR. Then $\chi(p)$ is non-decreasing for $p \in (\underline{\theta} + h(c), \bar{\theta} + h(0))$.*

Remark 3.1. *In Appendix B.6, we show that $\chi(p) \leq \beta(p)$. This suggests that, when $\beta(p)$ is small, the COST OF NEWEST FIRST is also small. The former occurs when review ratings have small impact on purchases. For example, when $\mathcal{F} = \mathcal{U}[0, \bar{\theta}]$ and $\bar{\theta} \rightarrow \infty$, the idiosyncratic variability dominates the variability from estimating μ through reviews, yielding $\beta(p) \rightarrow 1$ and thus $\chi(p) \rightarrow 1$.*

⁵For any price p satisfying Assumption 3.1, the denominator $\text{REV}(\sigma^{\text{NEWEST}}, p)$ of $\chi(p)$ is strictly positive.

3.4 Limited Attention: Main Driver for the Persistence of Negative Reviews

We further investigate the key drivers of the persistence of negative reviews. We show that this phenomenon is not an artifact of the specific belief update rule of our model; rather the limited attention of customers is the main driver behind our result.

The phenomenon holds for generic belief updating rules. Consider any updating rule where at any round t , the purchase probability when the last review is positive is strictly greater than when it is negative. Specifically, let $q(r, t)$ be the purchase probability of the customer at round $t \in \mathbb{N}$ when the newest review is $r \in \{0, 1\}$, and suppose $q(0, t) < q(1, t)$ for every t . Suppose that any new review is equally likely to be positive or negative ($\mu = 1/2$). We show that at any round, the newest review is more likely to be negative than positive (Proposition B.2, Appendix B.7).

The phenomenon disappears without limited attention. To isolate the effect of limited attention, we contrast to the setting where customers do not have limited attention and they can take the full history into account ($c \rightarrow \infty$). We show that there exist instances where, under any limited attention \tilde{c} , the CoNF can be arbitrarily large, while it goes to 1 when $c \rightarrow \infty$ (Proposition B.3, Appendix B.8); the latter result holds more broadly under mild assumptions on the purchase probability (Proposition B.4, Appendix B.8).

4 Dynamic Pricing Mitigates the Impact of CoNF

In this section, we allow the platform to optimize the pricing policy ρ , while the review ordering policy is either σ^{NEWEST} or σ^{RANDOM} . We assume that the platform knows the true underlying quality μ .⁶ Recall that Π^{STATIC} and Π^{DYNAMIC} are the classes of static and dynamic pricing policies respectively, and that the revenue and CoNF for a class Π are defined respectively as

$$\text{REV}(\sigma, \Pi) := \sup_{\rho \in \Pi} \text{REV}(\sigma, \rho) \quad \text{and} \quad \chi(\Pi) := \frac{\text{REV}(\sigma^{\text{RANDOM}}, \Pi)}{\text{REV}(\sigma^{\text{NEWEST}}, \Pi)}.$$

The main results of this section (Section 4.1) establish that the CoNF can be arbitrarily large for the optimal static pricing policy (Theorem 4.1) but that it is bounded by a small constant for the optimal dynamic pricing policy (Theorem 4.2). The main technical challenge of this section is in proving Theorem 4.2. To do this, we first *characterize* the optimal dynamic pricing policies under both σ^{NEWEST} and σ^{RANDOM} and derive exact expressions for their long-term revenue (Section 4.2). In doing so, we derive a structural property of the optimal dynamic pricing policy under σ^{NEWEST} : the prices ensure that the purchase probability is always equal regardless of the state of reviews.

4.1 Cost of Newest First under Optimal Static and Dynamic Pricing

We first establish that when optimizing over static prices, the CoNF can be arbitrarily large for any number of reviews c . Note that this is not implied by Theorem 3.2, since here we assume the platform always chooses the *optimal* static price for a given instance.

Theorem 4.1. *For any instance where the support of \mathcal{F} is $[0, \bar{\theta}]$, it holds that $\chi(\Pi^{\text{STATIC}}) \geq \frac{\mu^c h(c)}{h(0) + \bar{\theta}}$.*

⁶We assume μ is fixed over time and the platform has access to enough reviews to estimate μ arbitrarily well.

This implies that $\chi(\Pi^{\text{STATIC}}) \rightarrow +\infty$ if $h(c)$ is held constant and $\bar{\theta} \rightarrow 0$ and $h(0) \rightarrow 0$. Intuitively, this means that when the variability in the customer's idiosyncratic valuation Θ_t is negligible compared to the variability in review-inferred quality estimates, σ^{NEWEST} spends a disproportionate time in the state with no positive reviews, which leads to unbounded CONF.

Proof of Theorem 4.1. The optimal revenue under σ^{NEWEST} is at most $h(0) + \bar{\theta}$. This is because any price $p > h(0) + \bar{\theta}$ induces a purchase probability of zero when all reviews are negative and thus σ^{NEWEST} gets “stuck” in a zero revenue state. Thus, $\max_{p \in \mathbb{R}} \text{REV}(\sigma^{\text{NEWEST}}, p) \leq h(0) + \bar{\theta}$.

Under σ^{RANDOM} , if all reviews are positive, the non-negativity of the value distribution implies that a price $p = h(c)$ induces a purchase with probability one. The probability of this event is μ^c , which implies that $\max_{p \in \mathbb{R}} \text{REV}(\sigma^{\text{RANDOM}}, p) \geq \mu^c h(c)$. Combining the two inequalities we obtain

$$\chi(\Pi^{\text{STATIC}}) = \frac{\max_{p \in \mathbb{R}} \text{REV}(\sigma^{\text{RANDOM}}, p)}{\max_{p \in \mathbb{R}} \text{REV}(\sigma^{\text{NEWEST}}, p)} \geq \frac{\mu^c h(c)}{h(0) + \bar{\theta}}.$$

□

Next, we show an upper bound under dynamic pricing, which is the main result of this section.

Theorem 4.2. *For any instance, it holds that $\chi(\Pi^{\text{DYNAMIC}}) \leq \frac{2}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq 0]}$.*

In contrast to static pricing where the CONF can be arbitrarily bad, Theorem 4.2 shows that its negative impact is uniformly bounded under dynamic pricing. If the idiosyncratic valuation Θ_t is always non-negative, the upper bound on $\chi(\Pi^{\text{DYNAMIC}})$ is 2. Theorem 4.2 applies even when Θ_t can be negative: for example if it is non-negative with probability 1/2, then $\chi(\Pi^{\text{DYNAMIC}}) \leq 4$.

The proof of Theorem 4.2 (Section 4.3) relies on characterizing the optimal dynamic pricing policies under both σ^{NEWEST} and σ^{RANDOM} and their corresponding revenues (Section 4.2). Recall that with static pricing, σ^{NEWEST} spends a disproportionate amount of time in a negative review state compared to σ^{RANDOM} . In contrast, the optimal dynamic pricing sets prices so that the purchase probability is equal across all review states, leading to σ^{NEWEST} and σ^{RANDOM} spending the same amount of time in each review state (Section 4.2). This allows us to bound the ratio of demands under σ^{NEWEST} and σ^{RANDOM} by a factor of $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq 0]$. Finally, we bound the ratio of the optimal prices under σ^{NEWEST} and σ^{RANDOM} , which we decompose in two terms stemming from the customer's belief about μ and customer specific valuation; each term is bounded by 1 (Section 4.3).

We complement this result by a lower bound (proof in Appendix C.1) which shows that the COST OF NEWEST FIRST still exists even under optimal dynamic pricing.

Proposition 4.1. *For any $\alpha < 4/3$, there exists an instance such that $\chi(\Pi^{\text{DYNAMIC}}) > \alpha$.*

Remark 4.1. *Even under optimal dynamic pricing, it is still the case that σ^{RANDOM} induces no smaller revenue than σ^{NEWEST} , i.e., $\chi(\Pi^{\text{DYNAMIC}}) \geq 1$ (see Appendix C.2).*

Remark 4.2. *If we have the additional knowledge of $h(n) \leq u$ for some $u \geq 0$, then we can improve the result of Theorem 4.2 to $\chi(\Pi^{\text{DYNAMIC}}) \leq \frac{2\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq -u]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq 0]}$ (see Appendix C.3).*

4.2 Characterization of Optimal Dynamic Pricing under Newest First

Our key technical contribution characterizes the optimal dynamic pricing policy under σ^{NEWEST} . We show that it satisfies a structural property: the purchase probability is equal regardless of the

review state. We define the policies that satisfy this property as *review-offsetting* policies. For a state of reviews $\mathbf{z} = (z_1, \dots, z_c) \in \{0, 1\}^c$, we denote by $N_{\mathbf{z}} = \sum_{i=1}^c z_i$ the number of positive review ratings.

Definition 4.1. A dynamic pricing policy ρ is *review-offsetting* if there exists an offset $a \in \mathbb{R}$ such that $\rho(\mathbf{z}) = h(N_{\mathbf{z}}) + a$ for all $\mathbf{z} \in \{0, 1\}^c$.

Note that for a review-offsetting policy ρ , the purchase probability at any state $\mathbf{z} \in \{0, 1\}^c$ is $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N_{\mathbf{z}}) \geq h(N_{\mathbf{z}}) + a] = \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq a]$, where the last term does not depend on \mathbf{z} . Hence, review-offsetting policies induce equal purchase probability regardless of the state of reviews.

The main result of this section establishes that under σ^{NEWEST} , there is a review-offsetting dynamic pricing policy that maximizes revenue and characterizes the corresponding offset. We note that this characterization is the only place where we require the platform to know the true quality μ .

Theorem 4.3. Let $p^* \in \arg \max_{p \in \mathbb{R}} p \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + \mathbb{E}_{N \sim \text{Binom}(c, \mu)}[h(N)] \geq p]$. Under σ^{NEWEST} , the review-offsetting policy with offset $a^* = p^* - \mathbb{E}_{N \sim \text{Binom}(c, \mu)}[h(N)]$ is an optimal dynamic pricing policy.

To prove Theorem 4.3, we need a characterization of the revenue for a dynamic pricing policy ρ (similar to Proposition 3.2). This is established in the following lemma (proof in Appendix C.4).

Lemma 4.1. For any dynamic pricing policy ρ with positive purchase probabilities in all states,⁷

$$\text{REV}(\sigma^{\text{NEWEST}}, \rho) = \frac{\mathbb{E}_{Y_1, \dots, Y_c \sim i.i.d. \text{Bern}(\mu)}[\rho(\mathbf{Y})]}{\mathbb{E}_{Y_1, \dots, Y_c \sim i.i.d. \text{Bern}(\mu)} \left[\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N_{\mathbf{Y}}) \geq \rho(\mathbf{Y})]} \right]}.$$

Proof of Theorem 4.3. The high-level idea of the proof is that any dynamic pricing policy ρ can be improved by a particular review-offsetting policy. Specifically, for any dynamic pricing policy ρ and any review state $\mathbf{z} \in \{0, 1\}^c$, we define a policy $\tilde{\rho}_{\mathbf{z}}$ to be the review-offsetting policy with offset $a_{\mathbf{z}} := \rho(\mathbf{z}) - h(N_{\mathbf{z}})$ and show that the revenue of one of $\{\tilde{\rho}_{\mathbf{z}}\}_{\mathbf{z} \in \{0, 1\}^c}$ is at least the revenue of ρ .

To establish this we express the revenue of any review-offsetting policy $\tilde{\rho}$ with offset a , which satisfies that $\tilde{\rho}(\mathbf{z}) = a + h(N_{\mathbf{z}})$ for all states \mathbf{z} . As a result, (a) its expected price is $\mathbb{E}[\tilde{\rho}(\mathbf{Y})] = a + \bar{h}$ where $\bar{h} = \mathbb{E}_{Y_1, \dots, Y_c \sim i.i.d. \text{Bern}(\mu)}[h(N_{\mathbf{Y}})]$ and (b) its purchase probability is always equal to $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N_{\mathbf{Y}}) \geq \tilde{\rho}(\mathbf{Y})] = \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + \bar{h} \geq a + \bar{h}]$. By Lemma 4.1, its revenue is equal to

$$\text{REV}(\sigma^{\text{NEWEST}}, \tilde{\rho}) = (a + \bar{h}) \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + \bar{h} \geq a + \bar{h}]. \quad (2)$$

Note that, across all review-offsetting policies, the offset that maximizes the above revenue is $a^* = p^* - \bar{h}$; we refer to the corresponding review-offsetting policy as ρ^* . We next show that the revenue of any dynamic pricing policy is upper bounded by the revenue of ρ^* , concluding the proof. Given that the policy $\tilde{\rho}_{\mathbf{z}}$ is review-offsetting with offset $a_{\mathbf{z}}$, the above implies

$$\text{REV}(\sigma^{\text{NEWEST}}, \tilde{\rho}_{\mathbf{z}}) = \underbrace{(\rho(\mathbf{z}) - h(N_{\mathbf{z}}) + \bar{h})}_{A_{\mathbf{z}}} \cdot \underbrace{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N_{\mathbf{z}}) \geq \rho(\mathbf{z})]}_{(B_{\mathbf{z}})^{-1}}.$$

⁷In this lemma and the following proofs we again use the notation $\mathbf{Y} = (Y_1, \dots, Y_c)$ and $N_{\mathbf{Y}} = \sum_{i=1}^c Y_i$.

Let $\alpha_z = \mu^{N_z}(1 - \mu)^{c - N_z}$ be the probability that c i.i.d. $\text{Bern}(\mu)$ trials result in z . Hence $\mathbb{E}_{Y_1, \dots, Y_c \sim i.i.d. \text{Bern}(\mu)}[\rho(\mathbf{Y})] = \mathbb{E}_{Y_1, \dots, Y_c \sim i.i.d. \text{Bern}(\mu)}[\rho(\mathbf{Y}) - h(N_{\mathbf{Y}}) + \bar{h}] = \sum_{z \in \{0,1\}^c} \alpha_z A_z$. Applying Lemma 4.1 we can thus express the revenue of ρ as

$$\text{REV}(\sigma^{\text{NEWEST}}, \rho) = \frac{\sum_{z \in \{0,1\}^c} \alpha_z A_z}{\sum_{z \in \{0,1\}^c} \alpha_z B_z} \leq \max_{z \in S} \frac{A_z}{B_z} = \max_{z \in \{0,1\}^c} \text{REV}(\sigma^{\text{NEWEST}}, \tilde{\rho}_z) \leq \text{REV}(\sigma^{\text{NEWEST}}, \rho^*)$$

where the first inequality follows from the following natural convexity property (the proof can be found in Appendix C.5): for any $\{\tilde{\alpha}_i, \tilde{A}_i, \tilde{B}_i\}_{i \in S}$ with $\tilde{\alpha}_i, \tilde{B}_i > 0$ for all i , $\frac{\sum_{i \in S} \tilde{\alpha}_i \tilde{A}_i}{\sum_{i \in S} \tilde{\alpha}_i \tilde{B}_i} \leq \max_{i \in S} \frac{\tilde{A}_i}{\tilde{B}_i}$. \square

Intuition behind the optimal dynamic pricing. Intuitively, the term p^* in Theorem 4.3 is the optimal price when facing a *single* customer with a random valuation $\Theta + \mathbb{E}_{N \sim \text{Binom}(c, \mu)}[h(N)]$ for $\Theta \sim \mathcal{F}$. The selected offset makes the purchase probability equal to the purchase probability under the “single customer” setting with the optimal price p^* . This intuition enables us to characterize the optimal revenue of dynamic policies. To do this, a useful quantity in our characterization results is the optimal revenue for a given valuation distribution represented by a random variable V , i.e., $r^*(V) := \max_{p \in \mathbb{R}} p \mathbb{P}_V[V \geq p]$. For the distributions we consider, V has bounded support, and hence $r^*(V) < \infty$. The proof of the following corollary can be found in Appendix C.6.

Corollary 4.1. *The revenue of the optimal dynamic pricing policy under σ^{NEWEST} equals the optimal revenue from selling to a single customer with valuation $\Theta + \mathbb{E}_{N \sim \text{Binom}(c, \mu)}[h(N)]$. That is,*

$$\max_{\rho \in \Pi^{\text{DYNAMIC}}} \text{REV}(\sigma^{\text{NEWEST}}, \rho) = r^*(\Theta + \mathbb{E}_{N \sim \text{Binom}(c, \mu)}[h(N)]).$$

Remark 4.3. In Appendix C.7, we also characterize the complete set of optimal policies for σ^{NEWEST} ; the stated policy is the unique optimal policy under mild regularity conditions (Appendix C.8).

Comparison between optimal dynamic pricing for Newest and Random. The following proposition (proof in Appendix C.9) derives the optimal dynamic pricing policy for σ^{RANDOM} . Specifically, for every state of reviews, the optimal price is the revenue-maximizing price for that state.

Proposition 4.2. *For every review state $z \in \{0,1\}^c$, any optimal dynamic pricing policy under σ^{RANDOM} sets $\rho^{\text{RANDOM}}(z) \in \arg \max_p p \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N_z) \geq p]$. This implies that*

$$\text{REV}(\sigma^{\text{RANDOM}}, \Pi^{\text{DYNAMIC}}) = \mathbb{E}_{N \sim \text{Binom}(c, \mu)}[r^*(\Theta + h(N))].$$

We next compare the optimal dynamic pricing policies under σ^{NEWEST} and σ^{RANDOM} . The following proposition (proof in Appendix C.10) shows that under mild regularity conditions, states z with $h(N_z) > \bar{h}$ result in higher price under σ^{NEWEST} , while states z which $h(N_z) < \bar{h}$ result in lower price under σ^{NEWEST} . Intuitively, ρ^{NEWEST} charges higher prices in review states z with “high” ratings and lower prices in review states z with “low” ratings compared to ρ^{RANDOM} in order to induce the same purchase probability in every review state.

Proposition 4.3. *Let $\bar{h} = \mathbb{E}_{N \sim \text{Binom}(c, \mu)}[h(N)]$. Under mild regularity conditions (see Appendix C.10), the unique dynamic pricing policies ρ^{NEWEST} and ρ^{RANDOM} satisfy:*

- $\rho^{\text{NEWEST}}(z) \geq \rho^{\text{RANDOM}}(z)$ for review states $z \in \{0,1\}^c$ with $h(N_z) > \bar{h}$,
- $\rho^{\text{NEWEST}}(z) \leq \rho^{\text{RANDOM}}(z)$ for review states $z \in \{0,1\}^c$ with $h(N_z) < \bar{h}$,
- $\rho^{\text{NEWEST}}(z) = \rho^{\text{RANDOM}}(z)$ review states $z \in \{0,1\}^c$ with $h(N_z) = \bar{h}$.

4.3 Cost of Newest First is Bounded under Dynamic Pricing (Theorem 4.2)

We now prove Theorem 4.2, leveraging the results of Section 4.2 that characterize the optimal dynamic pricing policies. For convenience, we denote $\mathbb{E}_{N \sim \text{Binom}(c, \mu)}$ by \mathbb{E}_N and $\bar{h} = \mathbb{E}_{N \sim \text{Binom}(c, \mu)}[h(N)]$. By Proposition 4.2 and Corollary 4.1, we can express the COST OF NEWEST FIRST as:

$$\chi(\Pi^{\text{DYNAMIC}}) = \frac{\mathbb{E}_{N \sim \text{Binom}(c, \mu)}[r^*(\Theta + h(N))]}{r^*(\Theta + \bar{h})} = \mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\frac{r^*(\Theta + h(N))}{r^*(\Theta + \bar{h})} \right], \quad (3)$$

where the denominator does not depend on N and can thus move inside the expectation. We now focus on the quantity inside the expectation for a particular realization of N . For any price $p_n > 0$,

$$\begin{aligned} \frac{r^*(\Theta + h(n))}{r^*(\Theta + \bar{h})} &= \frac{p^*(\Theta + h(n)) \cdot \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(n) \geq p^*(\Theta + h(n))]}{p^*(\Theta + \bar{h}) \cdot \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + \bar{h} \geq p^*(\Theta + \bar{h})]} \\ &\leq \underbrace{\frac{p^*(\Theta + h(n))}{p_n}}_{\text{PRICE RATIO}(p_n, n)} \cdot \underbrace{\frac{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(n) \geq p^*(\Theta + h(n))]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + \bar{h} \geq p_n]}}_{\text{DEMAND RATIO}(p_n, n)}. \end{aligned} \quad (4)$$

The inequality replaces the revenue-maximizing price $p^*(\Theta + \bar{h})$ by another price p_n , which can only increase the ratio. Given that we operate with dynamic prices, we are allowed to select a different price for any number of positive reviews N .

In particular, a price of $\bar{h} + p^*(\Theta + h(n)) - h(n)$ ensures that the demand ratio is one. However, if $p^*(\Theta + h(n)) < h(n)$, the denominator in the price ratio can be unboundedly small. To simultaneously bound the expected price and demand ratios, we select $\tilde{p}_n = \bar{h} + \max(p^*(\Theta + h(n)) - h(n), 0)$.

Lemma 4.2. *The expected price ratio is at most $\mathbb{E}[\text{PRICE RATIO}(\tilde{p}_N, N)] \leq 2$.*

Lemma 4.3. *For any $n \in \{0, 1, \dots, c\}$, the demand ratio is $\text{DEMAND RATIO}(\tilde{p}_n, n) \leq \frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq 0]}$.*

Proof of Theorem 4.2. The proof directly combines (3), (4), and Lemmas 4.2 and 4.3. \square

What is left is to prove the lemmas that bound the expected price ratio and the demand ratio.

Proof of Lemma 4.2. Given that customer valuations are additive with a belief and an idiosyncratic component, the optimal price (numerator of price ratio) can be similarly decomposed as:

$$p^*(\Theta + h(n)) = \underbrace{h(n)}_{\text{Belief about } \mu} + \underbrace{p^*(\Theta + h(n)) - h(n)}_{\text{Idiosyncratic valuation}}$$

The expected price ratio $\mathbb{E}[\text{PRICE RATIO}(\tilde{p}_N, N)]$ for $\tilde{p}_n = \bar{h} + \max(p^*(\Theta + h(n)) - h(n), 0)$ is:

$$\mathbb{E}_N \left[\frac{h(N)}{\bar{h} + \max(p^*(\Theta + h(N)) - h(N), 0)} \right] + \mathbb{E}_N \left[\frac{p^*(\Theta + h(N)) - h(N)}{\bar{h} + \max(p^*(\Theta + h(N)) - h(N), 0)} \right].$$

Given that the denominators in both terms are positive and $\bar{h} = \mathbb{E}_N[h(N)]$, each of those terms can be upper bounded by 1, concluding the proof. \square

Proof of Lemma 4.3. For every number of positive reviews n , we distinguish two cases based on where the maximum in \tilde{p}_n lies. If $p^*(\Theta + h(n)) \geq h(n)$ the demand ratio is equal to 1. Otherwise,

$$\text{DEMAND RATIO}(\tilde{p}_n, n) = \frac{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq p^*(\Theta + h(n)) - h(n)]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq 0]} \leq \frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq 0]}.$$

□

4.4 Broader Implication: Cost of Ignoring State-dependent Customer Behavior

Suppose a platform uses σ^{NEWEST} , and they are *not aware* of the phenomenon that customer's purchase decisions depend on the state of reviews — they instead assume the purchase behavior is constant (i.e., purchase probability does not depend on the state of the newest c reviews). Then, if the platform uses a standard data-driven approach to optimize prices (e.g., do price experimentation and estimate demand from data), the optimal revenue is $\text{REV}(\sigma^{\text{NEWEST}}, \Pi^{\text{STATIC}})$. In contrast, a platform can estimate separate demands for each state of reviews and employ a dynamic pricing policy to earn $\text{REV}(\sigma^{\text{NEWEST}}, \Pi^{\text{DYNAMIC}})$. By comparing $\text{REV}(\sigma^{\text{NEWEST}}, \Pi^{\text{DYNAMIC}})$ with $\text{REV}(\sigma^{\text{NEWEST}}, \Pi^{\text{STATIC}})$, we show that the revenue loss from not accounting for this state-dependent behavior can be arbitrarily large. The next result (proof in Appendix C.11) follows from Theorems 4.1 and 4.2.

Corollary 4.2. *For any $M > 0$, there exists an instance such that $\frac{\text{REV}(\sigma^{\text{NEWEST}}, \Pi^{\text{DYNAMIC}})}{\text{REV}(\sigma^{\text{NEWEST}}, \Pi^{\text{STATIC}})} > M$.*

5 Cost of Newest First in the Presence of Non-Stationarity

As discussed in the introduction, customers prefer to read more recent reviews; this explains the practical popularity of σ^{NEWEST} . One potential reason for the popularity of Newest First is that customers prefer to read newer reviews because of their belief that newer reviews are more representative of the current product quality compared to older reviews. This might lead one to believe that Newest First is better than Random when the product quality μ changes over time. To study this phenomenon, we focus on a simple model where just one review is displayed ($c = 1$) and the product quality $\mu^{(t)}$ evolves according to a Markov Chain with a high value μ_H and a low value μ_L . The transition dynamics of the product quality Markov Chain $\mu^{(t)}$ are such that at from any state with probability $\xi \in (0, 1]$ the Markov Chain $\mu^{(t)}$ transitions to a new state which is equally likely to be μ_H or μ_L and otherwise it remains at the current state. The customer's purchase behavior given reviews is the same as in Section 2.1. Let $q_0 = \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(0) \geq p]$ and $q_1 = \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(1) \geq p]$ be the purchase probabilities for a negative and a positive review respectively.

Our first result in this setting shows that, for any qualities μ_L and μ_H , the Cost of Newest First continues to arise, i.e., the revenue under the Newest First ordering policy is strictly worse than the revenue under a random ordering policy σ^{RANDOM} . To define the latter, notice that, given that the switching probability ξ is independent of the product quality, the steady-state distribution for the quality is μ_H with probability $1/2$ and μ_L with probability $1/2$. As a result, we define σ^{RANDOM} as the review from a random time period, which is thus drawn from the distribution $\text{Bernoulli}(\frac{\mu_H + \mu_L}{2})$.

Proposition 5.1. *For any probability of change $\xi \in (0, 1]$, product qualities μ_L and μ_H with $\mu_L < \mu_H$, and any price p satisfying Assumption 3.1, the revenue of σ^{NEWEST} is strictly smaller than the revenue of σ^{RANDOM} , i.e., $\text{REV}(\sigma^{\text{RANDOM}}, p) > \text{REV}(\sigma^{\text{NEWEST}}, p)$.*

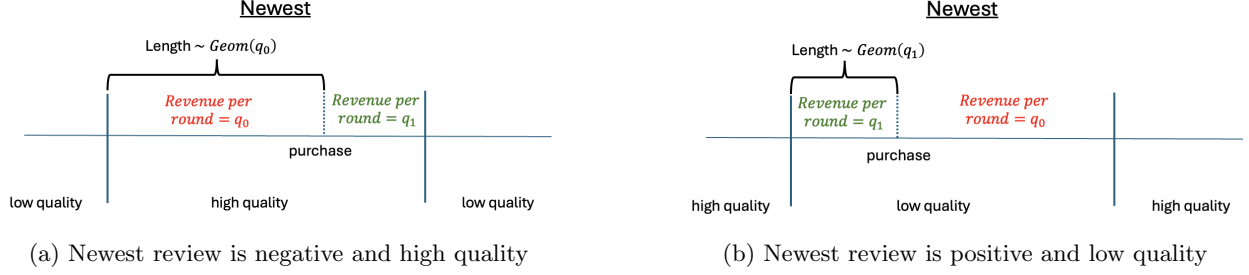


Figure 2: When the newest review is negative and the state transitions from low to high, it takes $\text{Geometric}(q_0)$ rounds to obtain a new review (Figure 2 (a)). On the other hand, when the newest review is positive and the state transitions from high to low it takes $\text{Geometric}(q_1)$ rounds to obtain a new review (Figure 2 (b)).

Proof sketch. For simplicity, we here assume that $\mu_L = 0$ and $\mu_H = 1$; the full proof in Appendix D.1 proves the theorem for general product qualities. The crux of the proof is that, when transitioning from a low to high (high to low) state, σ^{NEWEST} needs more (less) time than σ^{RANDOM} to obtain a purchase and thus update the newest review; see Figure 2 for an illustration. For periods in which the newest review has the same quality as the environment, σ^{NEWEST} and σ^{RANDOM} operate identically if one couples the qualities for the two processes. \square

We next formalize the idea that Newest First is better than Random in representing the product quality. To do this, for any review ordering policy σ , pricing policy ρ , and probability of change $\xi \in (0, 1]$, we define the belief error as a measure of difference between the customer's posterior and the product quality at the current round in steady state.

$$\text{BELIEFERROR}(\sigma, \rho; \xi) := \liminf_{T \rightarrow \infty} \mathbb{E} \left[\frac{\sum_{t=1}^T (\mathbb{E}[\Phi_t] - \mu_t)^2}{T} \right].$$

The following proposition establishes that, when the posterior after a positive (respectively, negative) review matches μ_H (respectively, μ_L), σ^{NEWEST} induces lower belief error than σ^{RANDOM} for any probability of change $\xi \in (0, 1)$. This explains the intuitive appeal of Newest First.

Proposition 5.2. *Let $\mu_L = \mathbb{E}[\text{Beta}(a, b + 1)]$ and $\mu_H = \mathbb{E}[\text{Beta}(a + 1, b)]$ and price p satisfying Assumption 3.1. For any $\xi \in (0, 1)$, $\text{BELIEFERROR}(\sigma^{\text{NEWEST}}, p; \xi) < \text{BELIEFERROR}(\sigma^{\text{RANDOM}}, p; \xi)$. Moreover, $\text{BELIEFERROR}(\sigma, p; \xi)$ increases with ξ for $\sigma = \sigma^{\text{NEWEST}}$ and is constant in ξ for $\sigma = \sigma^{\text{RANDOM}}$.*

Intuitively, more frequent quality switches (larger ξ) implies that the newest review comes from the previous product quality (similar to Figure 2) with higher probability in steady state. The formal proof is provided in Appendix D.2.

6 Numerical evaluation on synthetic and real data

We conduct two sets of simulations on synthetic data to validate our main results under variants of our theoretical model (Sections 6.1 and 6.2). Subsequently, we further corroborate our theoretical findings by analyzing real data of reviews from Tripadvisor (Section 6.3).

6.1 Cold start: Evaluating the impact of the initial condition

Recall that our model assumes that the process starts with an infinite pool of reviews (see Section 2.2). Then, the σ^{RANDOM} policy is equivalent to the c reviews being drawn from i.i.d. $\text{Bern}(\mu)$ variables at each time step. Numerically, we simulate the random policy from a *cold start*, where it has to choose reviews from a finite historical pool of reviews from previous time steps and this review pool is initialized with a single review. Specifically, we test the following policies:

- Random-finite: We maintain a review pool which is initialized with c reviews drawn from $\text{Bern}(\mu)$. At each time, the policy draws c reviews at random from this pool. If there is a purchase, a new review is added to the review pool.
- Random-infinite: At each time, c reviews are drawn i.i.d. from $\text{Bern}(\mu)$.
- Newest First: The most recent c reviews are chosen from the review pool.

We run the above three policies for two instances where $\mu \in \{0.1, 0.5\}$, $c = 1$, the prior belief is $\text{Beta}(0.5, 0.5)$, the distribution of Θ_t is $\mathcal{F} = \mathcal{U}[0, 1]$, and the price is $p = 1$. We plot the revenue over 5000 time steps, averaged over 10000 runs in Figure 3.

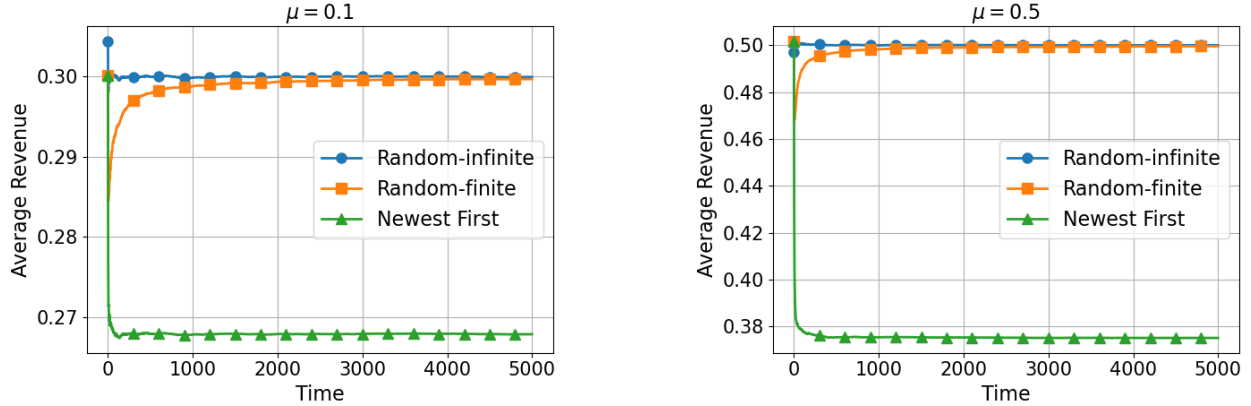


Figure 3: Average revenue for Random-finite, Random-infinite and Newest First as a function of time.

Figure 3 shows that Random-finite behaves similarly to Random-infinite, and there is a clear gap between these policies and Newest First, which represents the CoNF. As time increases, the finite pool of reviews increases under Random-finite, and hence the revenue converges to Random-infinite. We also observe that Random-finite yields a slightly lower revenue than Random-infinite in the initial rounds. This is because there is a similar phenomenon to CoNF that arises when comparing Random-finite to Random-infinite. Specifically, when the review pool is small, if those reviews are all negative, then any random sample from this pool will be negative, and hence it will take longer to get a new purchase, compared to when reviews are drawn from an infinite pool. As the time horizon increases and the review pool gets larger, the behavior of Random-finite approaches Random-infinite. Next, Random-finite takes longer to converge to Random-infinite when $\mu = 0.1$ compared to $\mu = 0.5$. This is because when the product is lower quality, there are fewer purchases, and hence it takes longer for the review pool to grow in size.

6.2 Empirical validation that dynamic pricing mitigates the impact of CoNF

Recall that our main results of Section 4 was that dynamic pricing mitigates the impact of CoNF. Under dynamic pricing, the worst case value for the CoNF, $\chi(\Pi^{\text{DYNAMIC}})$, is at most 2 (Theorem 4.2), while $\chi(\Pi^{\text{STATIC}})$ can be unbounded (Theorem 4.1). Since these results are worst-case bounds, in this section, we compare the values of $\chi(\Pi^{\text{STATIC}})$ and $\chi(\Pi^{\text{DYNAMIC}})$ on a class of instances.

We consider instances where the customer-specific valuation is distributed as $\mathcal{F} = \mathcal{U}[0, \epsilon]$, $c = 1$, $\mu = 0.5$, the prior belief is $\text{Beta}(a, a)$, and $h(\cdot)$ is the mean of the posterior. We vary ϵ , which corresponds to the variability in the customer-specific valuation, and we also vary the magnitude of a , which represents the strength of the prior belief. A low prior strength, corresponding to lower values for a , implies that a review has a large impact on a customer’s belief, whereas a higher prior strength implies that the customer’s belief is minimally impacted by reviews. We vary ϵ from 0 to 3, while we vary $a \in \{0.05, 0.5, 5\}$. We plot the value of $\chi(\Pi^{\text{STATIC}})$ and $\chi(\Pi^{\text{DYNAMIC}})$ in Figure 4.

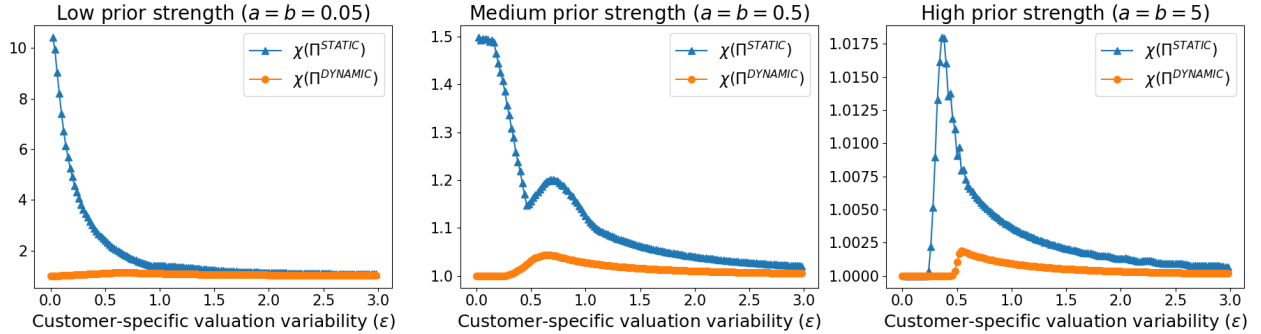


Figure 4: Comparison of $\chi(\Pi^{\text{STATIC}})$ and $\chi(\Pi^{\text{DYNAMIC}})$.

We observe that the CoNF is small (less than 1.1) under dynamic pricing under all instances. For static pricing, the CoNF can be large, especially when the prior strength is low and ϵ is small. This behavior is expected from the theory, as the instances that were constructed to show that the CoNF can be arbitrarily bad are instances where the prior strength is low (and thus the variability in review-inferred quality estimates is high) while the variability in customer-specific valuation is also low (Theorem 4.1). Importantly, we observe that $\chi(\Pi^{\text{DYNAMIC}})$ is strictly smaller than $\chi(\Pi^{\text{STATIC}})$ for all instances that were tested. These results corroborate the message that dynamic pricing mitigates the impact of the Cost of Newest First.

6.3 Evidence supporting CoNF from Tripadvisor

We validate our theoretical findings using a dataset of reviews from Tripadvisor, an online platform whose default review ordering policy is σ^{NEWEST} . We focused on 109 hotels in the region of Times Square, New York, where we retrieved the newest 1000 reviews for each of these hotels. We searched “Times Square” on Tripadvisor (March 20, 2024) and collected data from the first 21 hotels with more than 1000 reviews.⁸ Subsequently, we made the same query and collected 88 additional hotels (August 6, 2024). A hotel’s page on Tripadvisor displays the average numerical rating of all reviews, as well as details of the first 10 reviews. More reviews can be shown if the user clicks for the next

⁸We initially retrieved 3 more hotels that had less than 1000 reviews (respectively 394, 770, and 957) but later removed them to have a consistent rule. All of those three hotels exhibit CoNF and thus including them would only strengthen our empirical findings.

page of reviews. Each review has a rating (an integer between 1 and 5), the date the review was written, information on the user who wrote the review, and the text of the review.

Recall that Proposition 3.3 states that the average number of positive reviews from the first c reviews is smaller under σ^{NEWEST} compared to σ^{RANDOM} . We corroborate this result by computing the empirical distributions for each hotel. Since the hotel’s page on Tripadvisor shows 10 reviews by default, we use $c = 10$.

For each hotel, we compute two empirical distributions, which we call $\hat{\pi}^{\text{NEWEST}}$ and $\hat{\pi}^{\text{RANDOM}}$. Let d_1 and d_2 be the dates of when the earliest and the latest review was written (among the reviews retrieved for this hotel). For each date between d_1 and d_2 , we compute what the 10 newest reviews were on that date and compute their average rating; this becomes a sample in our empirical distribution $\hat{\pi}^{\text{NEWEST}}$. The number of samples is thus simply the number of days between d_1 and d_2 , which, on average, corresponds to 5.95 years. The empirical distribution $\hat{\pi}^{\text{RANDOM}}$ is computed by randomly sampling 10 reviews out of all reviews retrieved from the particular hotel and computing their average rating, where the number of samples is equal to that of $\hat{\pi}^{\text{NEWEST}}$.

Since Tripadvisor did not use the σ^{RANDOM} policy, $\hat{\pi}^{\text{RANDOM}}$ is not exactly what the distribution of reviews would have been under the counterfactual world where Tripadvisor always used σ^{RANDOM} . However, if the review rating of a customer does not depend on the historical reviews that they read, then the empirical distribution is equal to the distribution of review ratings that we would have observed if Tripadvisor used σ^{RANDOM} . We note that, if Tripadvisor used σ^{RANDOM} , there may have been more (or fewer) purchases, and hence there may have been a different number of observed reviews. Therefore, we are approximating the distribution of reviews under the σ^{RANDOM} policy by selecting a random set of reviews from the set of all reviews that were collected under σ^{NEWEST} .

Results. To validate our theoretical findings, we consider a null hypothesis positing that the probability of $\hat{\pi}^{\text{RANDOM}}$ having higher mean than $\hat{\pi}^{\text{NEWEST}}$ is exactly 50%. In our data, we observe that, for 79 out of 109, the mean of $\hat{\pi}^{\text{NEWEST}}$ is strictly smaller than that of $\hat{\pi}^{\text{RANDOM}}$ which implies a p-value of 5.4×10^{-7} (and thus rejects the null hypothesis). On average, the mean of $\hat{\pi}^{\text{NEWEST}}$ is smaller than that of $\hat{\pi}^{\text{RANDOM}}$ by 2.3%, or 0.092 points out of 5 in absolute terms. Figure 5 plots a histogram of the difference between these means for the 109 hotels.

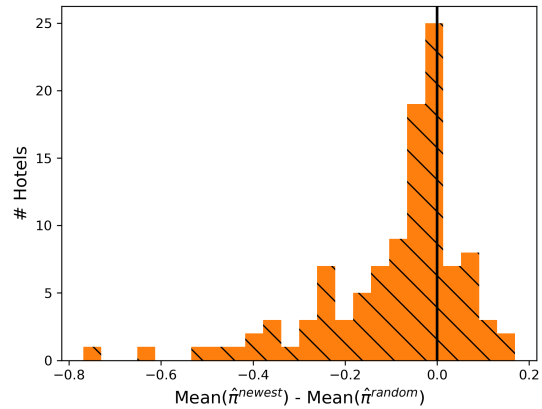


Figure 5: Histogram of the difference of means of $\hat{\pi}^{\text{NEWEST}}$ and $\hat{\pi}^{\text{RANDOM}}$ for the 109 hotels.

7 Conclusions

In this paper, we model the idea that customers read only a small number of reviews before making purchase decisions. This model gives rise to the COST OF NEWEST FIRST, the idea that, when reviews are ordered by Newest First, negative reviews will persist as the newest review longer than positive reviews. This phenomenon does not arise in models from the existing literature, since prior

works assume that customers incorporate either all reviews or a summary statistic of all reviews into their beliefs. We show that incorporating randomness into the review ordering or using dynamic pricing can alleviate the negative impact of the COST OF NEWEST FIRST.

Our work opens up a number of intriguing avenues for future research. First, existing literature on social learning studies the self-selection bias (which we do not consider in our model) – how does this self-selection bias interact with the COST OF NEWEST FIRST? Second, in terms of operational decisions, a platform contains multiple products — should it take the state of reviews into consideration when making display or ranking decisions? Third, given this limited attention behavior, are there alternative methods of disseminating relevant information from reviews? For example, one could succinctly summarize information from all reviews (via, e.g, generative AI) to be the most helpful for each customer. Lastly, on the theoretical side, our analysis fully characterizes the steady state of a stochastic process whose state remains unchanged with some state-dependent probability (Lemma B.2 which is the crux in the analysis of Lemma 3.1). It would be interesting to apply this result to other settings that exhibit a similar structure.

Acknowledgements.

We thank the anonymous reviewers from the 25th ACM Conference on Economics and Computation (EC 2024) for their thorough feedback that greatly improved the presentation of the paper. We are also grateful to the Simons Institute for the Theory of Computing as this work started during the Fall’22 semester-long program on *Data Driven Decision Processes*.

References

- [AGI11] Nikolay Archak, Anindya Ghose, and Panagiotis G Ipeirotis. Deriving the pricing power of product features by mining consumer reviews. *Management science*, 57(8):1485–1509, 2011. 1, 5
- [AHB19] Yücel Akçay, Felix Huettnner, and Tamer Boyacı. Consumer choice under limited attention when alternatives have different information costs. *Operations Research*, 67(3):671–699, 2019. 6
- [AMMO22] Daron Acemoglu, Ali Makhdoumi, Azarakhsh Malekian, and Asuman Ozdaglar. Learning from reviews: The selection effect and the speed of learning. *Econometrica*, 90(6):2857–2899, 2022. 1, 4, 6
- [BA21] Tamer Boyacı and Yücel Akçay. Pricing when customers have limited attention. *Management Science*, 67(1):14–31, 2021. 6
- [Ban92] Abhijit V Banerjee. A simple model of herd behavior. *The quarterly journal of economics*, 107(3):797–817, 1992. 4
- [BHW92] Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, 100(5):992–1026, 1992. 4
- [BMZ12] John W Byers, Michael Mitzenmacher, and Georgios Zervas. The groupon effect on yelp ratings: a root cause analysis. In *Proceedings of the 13th ACM conference on electronic commerce*, pages 248–265, 2012. 5

- [Bon23] Tommaso Bondi. Alone, together: A model of social (mis)learning from consumer reviews. In Kevin Leyton-Brown, Jason D. Hartline, and Larry Samuelson, editors, *Proceedings of the 24th ACM Conference on Economics and Computation, EC 2023, London, United Kingdom, July 9-12, 2023*, page 296. ACM, 2023. 4
- [BPS18] Kostas Bimpikis, Yiangos Papanastasiou, and Nicos Savva. Crowdsourcing exploration. *Management Science*, 64(4):1477–1973, 2018. 4
- [BPS22] Etienne Boursier, Vianney Perchet, and Marco Scarsini. Social learning in non-stationary environments. In *International Conference on Algorithmic Learning Theory*, pages 128–129. PMLR, 2022. 7
- [BS18] Omar Besbes and Marco Scarsini. On information distortions in online ratings. *Operations Research*, 66(3):597–610, 2018. 1, 4, 6, 37
- [CDO22] Ishita Chakraborty, Joyee Deb, and Aniko Oery. When do consumers talk? *Available at SSRN 4155523*, 2022. 4
- [CIMS17] Davide Crapis, Bar Ifrach, Costis Maglaras, and Marco Scarsini. Monopoly pricing in the presence of social learning. *Management Science*, 63(11):3586–3608, 2017. 1, 4, 6, 37
- [CLT21] Ningyuan Chen, Anran Li, and Kalyan Talluri. Reviews and self-selection bias with operational implications. *Management Science*, 67(12):7472–7492, 2021. 1, 4, 6
- [CM06] Judith A Chevalier and Dina Mayzlin. The effect of word of mouth on sales: Online book reviews. *Journal of marketing research*, 43(3):345–354, 2006. 1
- [CSS24] Christoph Carnehl, André Stenzel, and Peter Schmidt. Pricing for the stars: Dynamic pricing in the presence of rating systems. *Management Science*, 70(3):1755–1772, 2024. 4
- [DLT21] Gregory DeCroix, Xiaoyang Long, and Jordan Tong. How service quality variability hurts revenue when customers learn: Implications for dynamic personalized pricing. *Operations Research*, 69(3):683–708, 2021. 7
- [FKKK14] Peter Frazier, David Kempe, Jon Kleinberg, and Robert Kleinberg. Incentivizing exploration. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 5–22, 2014. 4
- [FMMP19] Jon Feldman, Vahab Mirrokni, S. Muthukrishnan, and Martin Pál. Assortment optimization under small consideration sets. *Operations Research*, 67(4):1010–1026, 2019. 6
- [GHKV23] Wenshuo Guo, Nika Haghtalab, Kirthevasan Kandasamy, and Ellen Vitercik. Leveraging reviews: Learning to price with buyer and seller uncertainty. In Kevin Leyton-Brown, Jason D. Hartline, and Larry Samuelson, editors, *Proceedings of the 24th ACM Conference on Economics and Computation, EC 2023, London, United Kingdom, July 9-12, 2023*, page 816. ACM, 2023. 1, 4, 5, 7, 8
- [GI10] Anindya Ghose and Panagiotis G Ipeirotis. Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics. *IEEE transactions on knowledge and data engineering*, 23(10):1498–1512, 2010. 1

- [GL24] Guillermo Gallego and Anran Li. A random consideration set model for demand estimation, assortment optimization, and pricing. *Operations Research*, 72(4):1192–1211, 2024. 6
- [GMC⁺21] Pin Gao, Yuhang Ma, Ningyuan Chen, Guillermo Gallego, Anran Li, Paat Rusmevichientong, and Huseyin Topaloglu. Assortment optimization and pricing under the multinomial logit model with impatient customers: Sequential recommendation and selection. *Operations Research*, 69(5):1509–1532, 2021. 6
- [GPR06] Jacob K Goeree, Thomas R Palfrey, and Brian W Rogers. Social learning with private and common values. *Economic theory*, 28:245–264, 2006. 4
- [HC24] Michael Hamilton and Titing Cui. Fresh rating systems: Structure, incentives, and fees. *Incentives, and Fees (June 24, 2024)*, 2024. 7
- [HPZ17] Nan Hu, Paul A Pavlou, and Jie Zhang. On self-selection biases in online product reviews. *MIS quarterly*, 41(2):449–475, 2017. 4, 5
- [IMSZ19] Bar Ifrach, Costis Maglaras, Marco Scarsini, and Anna Zseleva. Bayesian social learning from consumer reviews. *Operations Research*, 67(5):1209–1221, 2019. 1, 4, 6, 37
- [Kav21] M Kavanagh. The impact of customer reviews on purchase decisions. *Bizrate Insights*, available at: <https://bizrateinsights.com/resources/the-impact-of-customer-reviews-on-purchase-decisions/> (accessed 4 Feb 2024), 2021. 1
- [KKM⁺17] Sampath Kannan, Michael Kearns, Jamie Morgenstern, Mallesh Pai, Aaron Roth, Rakesh Vohra, and Zhiwei Steven Wu. Fairness incentives for myopic agents. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, pages 369–386, 2017. 4
- [KMP14] Ilan Kremer, Yishay Mansour, and Motty Perry. Implementing the “wisdom of the crowd”. *Journal of Political Economy*, 122(5):988–1012, 2014. 4
- [LDRF⁺13] Stephan Ludwig, Ko De Ruyter, Mike Friedman, Elisabeth C Brüggén, Martin Wetzels, and Gerard Pfann. More than words: The influence of affective content and linguistic style matches in online reviews on conversion rates. *Journal of marketing*, 77(1):87–103, 2013. 1
- [LKAK18] Gaël Le Mens, Balázs Kovács, Judith Avrahami, and Yaakov Kareev. How endogenous crowd formation undermines the wisdom of the crowd in online ratings. *Psychological science*, 29(9):1475–1490, 2018. 6
- [LLS19] Xiao Liu, Dokyun Lee, and Kannan Srinivasan. Large-scale cross-category analysis of consumer review content on sales conversion leveraging deep learning. *Journal of Marketing Research*, 56(6):918–943, 2019. 1, 5
- [LS16] Ilan Lobel and Evan Sadler. Preferences, homophily, and social learning. *Operations Research*, 64(3):564–584, 2016. 4
- [Luc16] Michael Luca. Reviews, reputation, and revenue: The case of yelp. com. *Com (March 15, 2016)*. *Harvard Business School NOM Unit Working Paper*, (12-016), 2016. 1

- [LYMZ22] Zhanfei Lei, Dezhi Yin, Sabyasachi Mitra, and Han Zhang. Swayed by the reviews: Disentangling the effects of average ratings and individual reviews in online word-of-mouth. *Production and Operations Management*, 31(6):2393–2411, 2022. 1, 5
- [MSS20] Yishay Mansour, Aleksandrs Slivkins, and Vasilis Syrgkanis. Bayesian incentive-compatible bandit exploration. *Operations Research*, 68(4):1132–1161, 2020. 4
- [MSSV23] Costis Maglaras, Marco Scarsini, Dongwook Shin, and Stefano Vaccari. Product ranking in the presence of social learning. *Operations Research*, 71(4):1136–1153, 2023. 7
- [Pat17] Scott Patterson. The high-speed trading behind your amazon purchase. *The Wall Street Journal*, March 2017. Accessed: 2025-05-14. 4
- [PSX21] Sungsik Park, Woochoel Shin, and Jinhong Xie. The fateful first consumer review. *Marketing Science*, 40(3):481–507, 2021. 6
- [RHP23] Hang Ren, Tingliang Huang, and Georgia Perakis. Impact of social learning on consumer subsidies and supplier capacity for green technology adoption. *Available at SSRN 4335284*, 2023. 7
- [Say18] Amin Sayedi. Pricing in a duopoly with observational learning. *Available at SSRN 3131561*, 2018. 4
- [SR18] Sven Schmit and Carlos Riquelme. Human interaction with recommendation systems. In *International Conference on Artificial Intelligence and Statistics*, pages 862–870. PMLR, 2018. 4
- [SS00] Lones Smith and Peter Sørensen. Pathological outcomes of observational learning. *Econometrica*, 68(2):371–398, 2000. 4
- [VL21] Prasad Vana and Anja Lambrecht. The effect of individual online reviews on purchase likelihood. *Marketing Science*, 40(4):708–730, 2021. 1, 5
- [ZZ10] Feng Zhu and Xiaoquan Zhang. Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of marketing*, 74(2):133–148, 2010. 1

A Generalizing beyond Beta-Bernoulli distributions (Remark 2.1)

We consider a generalization of the customer behavior model in Section 2. We only point to the modeling assumptions that change; everything else remains as in Section 2:

1. With respect to the customer valuation, the product's unobservable part μ_t is drawn from an arbitrary distribution \mathcal{D} with mean μ and finite support $R = \{r_1, \dots, r_s\} \in \mathbb{R}^s$ where $r_1 < \dots < r_s$ and $s \geq 2$. That is, \mathcal{D} is no longer restricted to be Bernoulli with $S = \{0, 1\}$ and μ in $(0, 1)$. This can capture a system where μ_t is the number of stars ($S = \{1, \dots, 5\}$) which is common in online platforms such as Tripadvisor, Airbnb, and Amazon.
2. With respect to the customer purchase behavior, when presented with a vector of c reviews $\mathbf{Z}_t = (Z_{t,1}, \dots, Z_{t,c}) \in S^c$, the customer maps them to an estimated valuation $\hat{V}_t = \Theta_t + \hat{h}(\mathbf{Z}_t)$, where $\hat{h} : S^c \rightarrow \mathbb{R}$ is an arbitrary fixed mapping. The model of Section 2.1 is a special case where the estimate \hat{h} is created via a two-stage process: the customer initially creates a posterior belief $\Phi_t = \text{Beta}(a + \sum_{i=1}^c Z_{t,i}, b + c - \sum_{i=1}^c Z_{t,i})$ and maps this posterior to an estimate $\hat{h}(\mathbf{Z}_t) = h(\Phi_t)$ via a mapping h . Unlike this special case (where the estimate can only depend on the number of positive reviews in the c displayed reviews), our generalization here allows an arbitrary mapping from \mathbf{Z}_t that can also take the order of reviews into consideration.
3. With respect to the customer review generation, the review is X_t in the event of a purchase and $X_t = \perp$ otherwise. The difference to Section 2 is that μ_t is not restricted to be Bernoulli.
4. We assume that the estimator $\hat{h}(\mathbf{z})$ is strictly increasing in each coordinate of the review ratings \mathbf{z} . This extends Assumption 2.1 in a way that can capture the order of the reviews.
5. We assume that the customer's smallest estimated valuation $\hat{V} = \Theta + \underline{h}$ where $\underline{h} = \hat{h}(r_1, \dots, r_1)$ has nonzero mass on non-negative numbers, i.e., $\mathbb{P}_{\Theta \sim \mathcal{F}}[\hat{V} \geq 0] > 0$. This extends Assumption 2.2 and ensures that there exist non-negative prices inducing positive purchase probability.

The main driver behind all results in Sections 3 and 4 is the characterization of the stationary distribution of the Markov chain $\mathbf{Z}_t = (Z_{t,1}, \dots, Z_{t,c})$ of the newest c reviews.

- Estimator \hat{h} generalizes the customer's purchase behavior. The purchase probability at a state with review ratings $\mathbf{z} \in R^c$ is $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(\mathbf{z}) \geq \rho(\mathbf{z})]$ where ρ is the pricing policy.
- The review generation process changes the transition dynamics of \mathbf{Z}_t upon a purchase, i.e., a new review takes one of the values $\{r_1, \dots, r_s\}$ (as opposed to $\{0, 1\}$ in the original model).
- Letting $g_{\mathcal{D}}$ be the probability mass function of \mathcal{D} , the stationary distribution of \mathbf{Z}_t is

$$\pi_{\mathbf{z}} = \kappa \cdot \frac{\prod_{i=1}^c g_{\mathcal{D}}(r_i)}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + \hat{h}(\mathbf{z}) \geq \rho(\mathbf{z})]} \quad \text{where } \kappa = \frac{1}{\mathbb{E}_{Z_1, \dots, Z_c \sim i.i.d. \mathcal{D}} \left[\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + \hat{h}(Z_1, \dots, Z_c) \geq \rho(\mathbf{z})]} \right]}.$$

- This is analogous to Lemma 3.1 and Lemma C.1 for static and dynamic pricing in the the Beta-Bernoulli model. In particular, Lemma C.1 replaces $g_{\mathcal{D}}(r_i)$ and $\hat{h}(\mathbf{z})$ by $\mu^{z_i}(1 - \mu)^{1-z_i}$ and $h(N_{\mathbf{z}})$ in the expression for π , and the expectation $\mathbb{E}_{Z_1, \dots, Z_c \sim i.i.d. \mathcal{D}}$ and $\hat{h}(Z_1, \dots, Z_c)$ with $\mathbb{E}_{Y_1, \dots, Y_c \sim \text{Bern}(\mu)}$ and $h(\sum_{i=1}^c Y_i)$ in the expression for κ . The proof is completely analogous.

Using this stationary distribution and similar steps as in the proofs of Proposition 3.2 and Lemma 4.1 the revenue of σ^{NEWEST} in the generalized model becomes:

$$\text{REV}(\sigma^{\text{NEWEST}}, \rho) = \frac{\mathbb{E}_{Z_1, \dots, Z_c \sim i.i.d. \mathcal{D}}[\rho(\mathbf{Z})]}{\mathbb{E}_{Z_1, \dots, Z_c \sim i.i.d. \mathcal{D}}\left[\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + \hat{h}(\mathbf{Z}) \geq \rho(\mathbf{Z})]}\right]}.$$

The revenue of σ^{RANDOM} (which shows c i.i.d. reviews from \mathcal{D}) directly follows by adapting the purchase probability in Proposition 3.1 and Proposition 4.2:

$$\text{REV}(\sigma^{\text{RANDOM}}, \rho) = \mathbb{E}_{Z_1, \dots, Z_c \sim i.i.d. \mathcal{D}}\left[\rho(\mathbf{Z}) \cdot \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + \hat{h}(Z_1, \dots, Z_c) \geq \rho(\mathbf{Z})]\right].$$

Theorem 3.1 and 4.2 then extend to our generalized model by analogously adapting their proofs. With respect to the negative results (Theorem 3.2 and Theorem 4.1), given that the model in the main body is a special case of the generalized model, they also directly extend.

Remark A.1. *The generalized model allows for customers who are fully Bayesian in estimating the fixed valuation given the ordering policy σ and use an estimator mapping $\hat{h} = \hat{h}_{\text{FULLYBAYESIAN}}$. Such customers could be reactive to σ and account for the effect of CONF; the extension of the CONF result may thus seem surprising. The reason why this occurs is that CONF evaluates the customers assuming that they follow the same behavioral model under σ^{RANDOM} and under σ^{NEWEST} (and does not consider the setting where customers are reactive to the ordering policy).*

B Supplementary material for Section 3

B.1 Average rating under σ^{newest} is smaller than under σ^{random} (Proposition 3.3)

Proposition 3.3 states that the average review rating of the c reviews displayed by σ^{NEWEST} is strictly smaller than the average rating of the c reviews displayed by σ^{RANDOM} . To prove the proposition we analyze the stationary distribution of the number of positive reviews under each of σ^{NEWEST} and σ^{RANDOM} . Lemma 3.1 implies that the stationary distribution of seeing n positive reviews under σ^{NEWEST} is

$$\pi_n^{\text{NEWEST}} = \kappa \binom{c}{n} \frac{\mu^n (1 - \mu)^{c-n}}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(n) \geq p]} \quad \text{where} \quad \kappa = 1 / \mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N) \geq p]} \right].$$

By Proposition 3.2, $\text{REV}(\sigma^{\text{NEWEST}}, p) = \kappa \cdot p$, so κ can be interpreted as the rate at which customers purchase under σ^{NEWEST} . In contrast, the corresponding stationary distribution under σ^{RANDOM} is

$$\pi_n^{\text{RANDOM}} = \binom{c}{n} \mu^n (1 - \mu)^{c-n}.$$

To prove the proposition, we first compare the behavior of σ^{NEWEST} and σ^{RANDOM} based on $n^* = \max\{n | \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(n) \geq p] \leq \kappa\}$ (i.e., the largest number of positive review ratings where the purchase probability is at most the average purchase rate $\kappa = 1 / \mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N) \geq p]} \right]$).

Lemma B.1. *For any price p satisfying Assumption 3.1, $\pi_n^{\text{NEWEST}} \leq \pi_n^{\text{RANDOM}}$ if $n > n^*$ and $\pi_n^{\text{NEWEST}} \geq \pi_n^{\text{RANDOM}}$ if $n \leq n^*$.*

Proof of Lemma B.1. By Lemma 3.1, it holds that $\pi_n^{\text{RANDOM}} = \pi_n^{\text{NEWEST}} \cdot \frac{1}{\kappa} \cdot \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(n) \geq p]$. By the definition of n^* and the monotonicity of $h(n)$, if $n \leq n^*$, $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(n) \geq p] \leq \kappa$ and thus $\pi_n^{\text{NEWEST}} \geq \pi_n^{\text{RANDOM}}$. The other case is analogous. \square

Proof of Proposition 3.3. To show that σ^{NEWEST} has smaller average rating compared to σ^{RANDOM} ,

$$\begin{aligned} \mathbb{E}_{N \sim \pi_n^{\text{RANDOM}}}[N] - \mathbb{E}_{N \sim \pi_n^{\text{NEWEST}}}[N] &= \sum_{m=0}^c m(\pi_m^{\text{RANDOM}} - \pi_m^{\text{NEWEST}}) \\ &= \underbrace{\sum_{m=n^*+1}^c m(\pi_m^{\text{RANDOM}} - \pi_m^{\text{NEWEST}})}_{(1)} - \underbrace{\sum_{m=0}^{n^*} m(\pi_m^{\text{NEWEST}} - \pi_m^{\text{RANDOM}})}_{(2)} \end{aligned}$$

Lemma B.1 yields $(\pi_m^{\text{RANDOM}} - \pi_m^{\text{NEWEST}}) \geq 0$ for $m \geq n^* + 1$ and thus

$$\begin{aligned} (1) &\geq (n^* + 1) \sum_{m=n^*+1}^c (\pi_m^{\text{RANDOM}} - \pi_m^{\text{NEWEST}}) = (n^* + 1)(\mathbb{P}_{N \sim \pi_n^{\text{RANDOM}}}[N > n^*] - \mathbb{P}_{N \sim \pi_n^{\text{NEWEST}}}[N > n^*]) \\ &= (n^* + 1)(\mathbb{P}_{N \sim \pi_n^{\text{NEWEST}}}[N \leq n^*] - \mathbb{P}_{N \sim \pi_n^{\text{RANDOM}}}[N \leq n^*]) \end{aligned}$$

where in the last equality we used that $\mathbb{P}[X > n^*] = 1 - \mathbb{P}[X \leq n^*]$ for any random variable X . Using Lemma B.1 again yields $(\pi_m^{\text{NEWEST}} - \pi_m^{\text{RANDOM}}) \geq 0$ for $m \leq n^*$ and thus

$$(2) \leq n^* \sum_{m=0}^{n^*} (\pi_m^{\text{NEWEST}} - \pi_m^{\text{RANDOM}}) = n^*(\mathbb{P}_{N \sim \pi_n^{\text{NEWEST}}}[N \leq n^*] - \mathbb{P}_{N \sim \pi_n^{\text{RANDOM}}}[N \leq n^*]).$$

Therefore,

$$(1) - (2) \geq \mathbb{P}_{N \sim \pi_n^{\text{NEWEST}}}[N \leq n^*] - \mathbb{P}_{N \sim \pi_n^{\text{RANDOM}}}[N \leq n^*].$$

To conclude the proof it remains to show that $\mathbb{P}_{N \sim \pi_n^{\text{NEWEST}}}[N \leq n^*] > \mathbb{P}_{N \sim \pi_n^{\text{RANDOM}}}[N \leq n^*]$. By Lemma B.1, $\pi_n^{\text{NEWEST}} \geq \pi_n^{\text{RANDOM}}$ for $n \leq n^*$. To show the claim it is enough to show that there is some $n \leq n^*$ such that $\pi_n^{\text{NEWEST}} > \pi_n^{\text{RANDOM}}$. We show that the purchase probability when all reviews are negative is strictly greater under σ^{NEWEST} than under σ^{RANDOM} i.e. $\pi_0^{\text{NEWEST}} > \pi_0^{\text{RANDOM}}$. By the monotonicity of h , $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(0) \geq p] \leq \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(n) \geq p]$ for all $n \in \{0, 1, \dots, c\}$. Since p satisfies Assumption 3.1 (and is thus non-degenerate), the inequality is strict when $n = c$. Taking the reciprocal and expectation of the last inequality yields the result. \square

B.2 Stationary Distribution under Newest First (Lemma 3.1)

We first state a general property characterizing how the stationary distribution of a Markov chain changes if we modify it so that in every state, the process remains there with some probability.

Lemma B.2. *Let \mathcal{M} be a Markov chain on a finite state space \mathcal{S} ($|\mathcal{S}| = m$) with a transition probability matrix $M \in \mathbb{R}^{m \times m}$ and a stationary distribution $\pi \in \mathbb{R}^m$. For any function $f : \mathcal{S} \rightarrow (0, 1]$, we define a new Markov chain \mathcal{M}_f on the same state space \mathcal{S} . At state $s \in \mathcal{S}$, \mathcal{M}_f transitions according to the matrix M with probability $f(s)$ and remains in s with probability $1 - f(s)$. Then $\pi_f(s) = \kappa \cdot \frac{\pi(s)}{f(s)}$ is a stationary distribution of \mathcal{M}_f where $\kappa = 1 / \sum_{s \in \mathcal{S}} \frac{\pi(s)}{f(s)}$ is a normalizing constant*

Proof of Lemma B.2. For any state $s \in \mathcal{S}$, the probability of a self-transition under \mathcal{M}_f is $M_f(s, s) = f(s)M(s, s) + 1 - f(s)$ since there are two ways to transition from s back to itself: (1) the $f(s)$ transition followed by a transition back to s via \mathcal{M} and (2) the $1 - f(s)$ transition that does not alter the current state. For all states $s \neq s'$, $M_f(s, s') = f(s)M(s, s')$ since the only way for \mathcal{M}_f to transition from s to s' is to take a $f(s)$ transition at s and follow the transitions of \mathcal{M} to get to s' .

The distribution $\left\{ \pi_f(s) = \kappa \cdot \frac{\pi(s)}{f(s)} \right\}_{s \in \mathcal{S}}$ is a probability distribution by the definition of the normalizing constant $\kappa = 1 / \sum_{s \in \mathcal{S}} \frac{\pi(s)}{f(s)}$. Using the transitions of \mathcal{M}_f , it holds that:

$$\begin{aligned}
\sum_{s' \in \mathcal{S}} \pi_f(s') M_f(s', s) &= \pi_f(s) M_f(s, s) + \sum_{s' \neq s} \pi_f(s') M_f(s', s) \\
&= \pi_f(s) \cdot (1 - f(s) + f(s)M(s, s)) + \sum_{s' \neq s} \pi_f(s') \cdot (f(s')M(s', s)) \\
&= \kappa \cdot \frac{\pi(s)}{f(s)} \cdot (1 - f(s) + f(s)M(s, s)) + \sum_{s' \neq s} \kappa \cdot \frac{\pi(s')}{f(s')} \cdot (f(s')M(s', s)) \\
&= \kappa \cdot \frac{\pi(s)}{f(s)} - \underbrace{\kappa \cdot \pi(s) + \sum_{s' \neq s} \kappa \cdot \pi(s') M(s', s)}_{=\kappa \cdot \pi(s) \text{ (as } \pi \text{ is stationary for } \mathcal{M})} = \pi_f(s).
\end{aligned}$$

Hence, π_f is a stationary distribution of M_f as $\pi_f(s) = \sum_{s' \in \mathcal{S}} \pi_f(s') M_f(s', s)$ for all states s . \square

Proof of Lemma 3.1. We will show that the stationary distribution of the newest c reviews \mathbf{Z}_t is

$$\pi_{(z_1, \dots, z_c)} = \kappa \cdot \frac{\mu^{\sum_{i=1}^c z_i} (1 - \mu)^{c - \sum_{i=1}^c z_i}}{\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + h(\sum_{i=1}^c z_i) \geq p \right]}$$

where $\kappa = 1 / \mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}} [\Theta + h(N) \geq p]} \right]$ is the normalizing constant. In the language of Lemma B.2, \mathbf{Z}_t corresponds to \mathcal{M}_f , the state space \mathcal{S} to $\{0, 1\}^c$, and f is a function that expresses the purchase probability at a given state, i.e., $f(z_1, \dots, z_c) = \mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + h(\sum_{i=1}^c z_i) \geq p \right]$. Note that with probability $1 - f(z_1, \dots, z_c)$, \mathbf{Z}_t remains at the same state (as there is no purchase).

To apply Lemma B.2, we need to show that whenever there is purchase, \mathbf{Z}_t transitions according to a Markov chain with stationary distribution $\mu^{\sum_{i=1}^c z_i} (1 - \mu)^{c - \sum_{i=1}^c z_i}$. Consider the Markov chain \mathcal{M} which always replaces the c -th last review with a new $\text{Bern}(\mu)$ review. This process has stationary distribution equal to the above numerator and \mathbf{Z}_t transitions according to \mathcal{M} upon a purchase, i.e., with probability $f(z_1, \dots, z_c)$. As a result, by Lemma B.2, π is a stationary distribution for \mathbf{Z}_t . As \mathbf{Z}_t is irreducible and aperiodic, this is the unique stationary distribution. \square

B.3 Closed-form expression for Cost of Newest First (Lemma 3.2)

Proof of Lemma 3.2. Dividing the expressions given in Propositions 3.1 and 3.2, the price term p cancels out and the CoNF can be expressed as:

$$\begin{aligned}\chi(p) &= \mathbb{E}_{N \sim \text{Binom}(c, \mu)} [\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N) \geq p]] \cdot \mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N) \geq p]} \right] \\ &= \left(\sum_{i=0}^c \mu^i (1 - \mu)^{c-i} \binom{c}{i} \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(i) \geq p] \right) \cdot \left(\sum_{j=0}^c \mu^j (1 - \mu)^{c-j} \binom{c}{j} \frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(j) \geq p]} \right) \\ &= \sum_{i, j \in \{0, \dots, c\}} \mu^{i+j} (1 - \mu)^{2c-i-j} \binom{c}{i} \binom{c}{j} \frac{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(i) \geq p]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(j) \geq p]}.\end{aligned}$$

□

B.4 CoNF can be arbitrarily large under static pricing (Theorem 3.2)

Proof of Theorem 3.2. One summand in the right hand side of Lemma 3.2 contains the ratio of the purchase probability of all reviews being positive compared to all reviews being negative:

$$\beta(p) := \frac{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(c) \geq p]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(0) \geq p]},$$

which quantifies how much the reviews affect the purchase probability. Since all other terms are non-negative, the CoNF is lower bounded by this summand, i.e., $\chi(p) \geq \mu^c (1 - \mu)^c \beta(p)$.

Since \mathcal{F} is bounded, suppose that its support is $[\underline{\theta}, \bar{\theta}]$. When all reviews are negative, selecting a price of $h(0) + \bar{\theta}$ results in a purchase probability of 0. Combined with the continuity of \mathcal{F} , this implies that, when $p \rightarrow h(0) + \bar{\theta}$, the purchase probability goes to 0. If, on the other hand, all reviews were positive, then using $h(c) > h(0)$ and that \mathcal{F} is continuous and has positive mass on its support, the purchase probability is positive; i.e., $\lim_{p \rightarrow h(0) + \bar{\theta}} \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(c) \geq p] > 0$. Therefore $\lim_{p \rightarrow h(0) + \bar{\theta}} \beta(p) = +\infty$, which implies that $\lim_{p \rightarrow h(0) + \bar{\theta}} \chi(p) = +\infty$ since $\chi(p) \geq \mu^c (1 - \mu)^c \beta(p)$.

Lastly, for any price $p \in (\underline{\theta} + h(0), \bar{\theta} + h(0))$ it holds $0 < \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(0) \geq p] < \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(c) \geq p]$ and thus it is non-degenerate and non-absorbing. Combing this with the fact that $\bar{\theta} > 0$ (by Assumption 2.2), for every $M > 0$ there exists $\epsilon(M) > 0$ such that $p = \bar{\theta} + h(0) - \epsilon(M)$ satisfies Assumption 3.1 and has CoNF $\chi(p) > M$. □

B.5 Monotonicity of CoNF under monotone hazard rate (Proof of Proposition 3.4)

Proof of Proposition 3.4. By Lemma 3.2, the CoNF is given by

$$\begin{aligned}\chi(p) &= \sum_{i, j \in \{0, \dots, c\}} \mu^{i+j} (1 - \mu)^{2c-i-j} \binom{c}{i} \binom{c}{j} \frac{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(i) \geq p]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(j) \geq p]} \\ &= \sum_{i=0}^c \mu^{2i} (1 - \mu)^{2(c-i)} \binom{c}{i}^2 \\ &\quad + \sum_{i < j} \mu^{i+j} (1 - \mu)^{2c-i-j} \binom{c}{i} \binom{c}{j} \left(\frac{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(i) \geq p]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(j) \geq p]} + \frac{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(j) \geq p]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(i) \geq p]} \right).\end{aligned}$$

Letting F be the cumulative density function of \mathcal{F} and $\bar{F}(u) := 1 - F(u)$ be the survival function:

$$\chi(p) = \sum_{i=0}^c \mu^{2i} (1 - \mu)^{2(c-i)} \binom{c}{i}^2 + \sum_{i < j} \mu^{i+j} (1 - \mu)^{2c-i-j} \binom{c}{i} \binom{c}{j} \left(\frac{\bar{F}(p-h(i))}{\bar{F}(p-h(j))} + \frac{\bar{F}(p-h(j))}{\bar{F}(p-h(i))} \right).$$

We denote the ratio of the purchase probability with j and i positive reviews by $u_{i,j}(p) := \frac{\bar{F}(p-h(j))}{\bar{F}(p-h(i))}$. For $i < j$, this ratio is $u_{i,j}(p) \geq 1$ because $p - h(i) > p - h(j)$ due to the monotonicity of h (Assumption 2.1). Given that \mathcal{F} is continuous, $\bar{F}(x)$ is differentiable for any x such that $\bar{F}(x) \in (0, 1)$. Furthermore for $p \in (\underline{\theta} + h(c), \bar{\theta} + h(0))$, $\bar{F}(p-h(i)) \in (0, 1)$ for all $i \in \{0, 1, \dots, c\}$, and thus $u_{i,j}(p)$ is differentiable at p .

We now show that $u_{i,j}(p)$ is non-decreasing in p . Letting f be the probability density function of \mathcal{F} and taking the derivative of $u_{i,j}(p)$ with respect to p :

$$\begin{aligned} \frac{d}{dp} u_{i,j}(p) &= \frac{f(p-h(i))\bar{F}(p-h(j)) - f(p-h(j))\bar{F}(p-h(i))}{\bar{F}(p-h(j))^2} \\ &= \frac{\left(\frac{f(p-h(i))}{\bar{F}(p-h(i))} - \frac{f(p-h(j))}{\bar{F}(p-h(j))} \right) \bar{F}(p-h(j))\bar{F}(p-h(i))}{\bar{F}(p-h(j))^2}. \end{aligned}$$

Observe that $p-h(i) > p-h(j)$ since $i < j$ and the strict monotonicity of h . By the MHR property of \mathcal{F} : $\frac{f(p-h(i))}{\bar{F}(p-h(i))} \geq \frac{f(p-h(j))}{\bar{F}(p-h(j))}$, implying that $\frac{d}{dp} u_{i,j}(p) \geq 0$, and thus $u_{i,j}(p)$ is non-decreasing in p .

To finish the proof of the theorem, we rewrite $\chi(p)$ as a function of $\{u_{i,j}(p)\}_{i < j}$ as

$$\chi(p) = \sum_{i=0}^c \mu^{2i} (1 - \mu)^{2(c-i)} \binom{c}{i}^2 + \sum_{i < j} \mu^{i+j} (1 - \mu)^{2c-i-j} \binom{c}{i} \binom{c}{j} \left(u_{i,j}(p) + \frac{1}{u_{i,j}(p)} \right)$$

Note that the function $u + \frac{1}{u}$ is non-decreasing for $u \geq 1$. Since $u_{i,j}(p) \geq 1$ is non-decreasing in p , then $u_{i,j}(p) + \frac{1}{u_{i,j}(p)}$ is monotonically increasing in p for every $i < j$. Thus, $\chi(p)$ is non-decreasing in p . \square

Remark B.1. The same proof also extends to cases when $\underline{\theta} = -\infty$ and/or $\bar{\theta} = +\infty$.

B.6 Upper bounding CoNF by the effect of reviews on purchase (Remark 3.1)

Recall that $\beta(p)$ quantifies how much the review ratings affect the purchase probability. It is expected that when the review ratings have a small effect on the purchase probability then the CoNF is also small (i.e. small $\beta(p)$ would imply a small $\chi(p)$). We formalize this below.

Proposition B.1. For all prices p satisfying Assumption 3.1, the CoNF at most $\chi(p) \leq \beta(p)$.

Proof. The monotonicity of h , implies that for all $n \in \{0, 1, \dots, c\}$:

$$\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(c) \geq p] \geq \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(n) \geq p] \geq \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(0) \geq p].$$

Taking expectation over the number of positive reviews $N \sim \text{Binom}(c, \mu)$, the first inequality implies

$$\mathbb{E}_{N \sim \text{Binom}(c, \mu)} [\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N) \geq p]] \leq \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(c) \geq p]. \quad (5)$$

Similarly, taking expectation of the the reciprocal of the second inequality (which is well-defined as p satisfies Assumption 3.1 and is thus non-absorbing) implies that

$$\mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N) \geq p]} \right] \leq \frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(0) \geq p]}. \quad (6)$$

Expressing $\chi(\pi)$ as the ratio of the expressions in Propositions 3.1 and 3.2 and using (5) and (6):

$$\begin{aligned} \chi(p) &= \mathbb{E}_{N \sim \text{Binom}(c, \mu)} [\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N) \geq p]] \cdot \mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N) \geq p]} \right] \\ &\leq \frac{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(c) \geq p]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(0) \geq p]} = \beta(p) \end{aligned}$$

which concludes the proof. \square

B.7 Phenomenon holds under a generic update rule (Section 3.4)

Consider an item with product quality $\mu = 1/2$ and customers who read a single review ($c = 1$). Let $q(r, t)$ be the probability that the customer at round t purchases the item when the newest review is $r \in \{0, 1\}$. We consider any customer purchase model that satisfies the intuitive condition that the purchase probability given a negative review is strictly smaller than the purchase probability given a positive review. We also require that the purchase probability is always strictly smaller than 1.

Assumption B.1. *For every round $t \geq 1$, $q(0, t) < q(1, t) < 1$.*

We assume that the review at time $t = 0$ is randomly initialized and is positive with probability $1/2$.

Proposition B.2. *For any round $t \geq 2$, the newest review is negative with probability strictly greater than $1/2$.*

Proof. Fix any round t . Suppose the newest review r_t came from the previous round $t - 1$. Then, the review is negative with probability exactly equal to $1/2$. Next, suppose the newest review came from round $t - k$, where $k \geq 2$. Then, the probability that the review is negative is

$$\begin{aligned} \mathbb{P}(r_t = 0 \mid \text{review } r_t \text{ came from time } t - k) &= \frac{\mathbb{P}(r_t = 0, \text{review } r_t \text{ came from time } t - k)}{\mathbb{P}(\text{review } r_t \text{ came from time } t - k)} \\ &= \frac{\frac{1}{2} \prod_{\ell=1}^{k-1} (1 - q(0, t - \ell))}{\frac{1}{2} \prod_{\ell=1}^{k-1} (1 - q(0, t - \ell)) + \frac{1}{2} \prod_{\ell=1}^{k-1} (1 - q(1, t - \ell))} > \frac{1}{2}. \end{aligned}$$

The last inequality holds since $(1 - q(1, t - \ell)) < (1 - q(0, t - \ell))$ for all $\ell \in [1, k - 1]$ by Assumption B.1. Therefore, for any $k \geq 2$, if the review came from time $t - k$, then the review is negative with probability strictly greater than $1/2$. Since the purchase probability at any round is strictly less than 1 (by Assumption B.1), there is a non-zero probability that the review came from round $t - 2$ or before, which implies that the newest review is negative with probability strictly less than $1/2$. \square

B.8 CoNF exists due to limited attention (Section 3.4)

Suppose that the customer's product quality estimator h is the mean, $h[\Phi] = \mathbb{E}[\Phi]$ for all distributions Φ . We refer to an instance as tuple $(\mathcal{F}, \mu, a, b, p)$ where \mathcal{F} is the customer's idiosyncratic value distribution, μ is the product quality, a and b are the prior parameters, and p is the price the platform sets. For any limited attention window c , we define the CoNF as $\chi_c(p) = \frac{\text{REV}_c(\sigma^{\text{RANDOM}}, p)}{\text{REV}_c(\sigma^{\text{NEWEST}}, p)}$ where $\text{REV}_c(\sigma, \rho)$ is the revenue under review ordering policy σ and pricing policy ρ when customers read c reviews. Proposition B.3 establishes that the existence of CoNF is indeed due to the limited attention of the customers. In particular, for any limited attention window \tilde{c} and any $M > 0$ there exists an instance so that the $\chi_{\tilde{c}}(p) \geq M$, but $\lim_{c \rightarrow \infty} \chi_c(p) = 1$. Proposition B.4 then establishes that the latter result holds broadly under mild assumptions. As a result, CoNF disappears when customers observe full history.

Proposition B.3. *For any $\tilde{c} \geq 1$ and any $M > 0$ there exists an instance $\mathcal{E}_{\tilde{c}, M} = (\mathcal{F}, a, b, p)$ such that for the instance $\mathcal{E}_{\tilde{c}, M}$, it holds that a) $\chi_{\tilde{c}}(p) \geq M$ and b) $\lim_{c \rightarrow \infty} \chi_c(p) = 1$.*

Proof. For any $\tilde{c} \geq 1$ and $M > 0$ we define the instance $\mathcal{E}_{\tilde{c}, M}$ to have a customer value distribution of $\mathcal{F} = \text{Exponential}(\lambda_{\tilde{c}, M})$ where $\lambda_{\tilde{c}, M} = 3 \ln(2^{2\tilde{c}} M)$, a product quality $\mu = \frac{1}{2}$, prior parameters $a = b = 1$, and a price $p = 1 - \frac{1}{2+\tilde{c}}$.

To prove (a), note that, given $\mu = \frac{1}{2}$ we can express $\chi_{\tilde{c}}(p)$ using Lemma 3.2 as

$$\chi_{\tilde{c}}(p) = \sum_{i, j \in \{0, 1, \dots, \tilde{c}\}} \frac{1}{2^{2\tilde{c}}} \binom{\tilde{c}}{i} \binom{\tilde{c}}{j} \frac{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(i) \geq p]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(j) \geq p]} \geq \frac{1}{2^{2\tilde{c}}} \frac{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(\tilde{c}) \geq p]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(0) \geq p]}. \quad (7)$$

The inequality holds as one of the terms is $\frac{1}{2^{2\tilde{c}}} \frac{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(\tilde{c}) \geq p]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(0) \geq p]}$ and all other terms are non-negative.

Given that h is the mean and $a = b = 1$ then $h(\tilde{c}) = h(\text{Beta}(1 + \tilde{c}, 1)) = \frac{1+\tilde{c}}{2+\tilde{c}}$ and $h(0) = h(\text{Beta}(1, 1 + \tilde{c})) = \frac{1}{2+\tilde{c}}$. Combining this with the fact that $p = 1 - \frac{1}{2+\tilde{c}}$, the purchase probability with all positive reviews equals

$$\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{\tilde{c} + 1}{\tilde{c} + 2} \geq 1 - \frac{1}{\tilde{c} + 2} \right] = \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq 0] = 1 \quad (8)$$

as the exponential distribution is always non-negative. The purchase probability for all negative reviews equals

$$\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{1}{\tilde{c} + 1} \geq 1 - \frac{1}{\tilde{c} + 2} \right] = \mathbb{P}_{\Theta \sim \text{Bern}(\frac{1}{2^{2\tilde{c}}M})} \left[\Theta \geq 1 - \frac{2}{2 + \tilde{c}} \right] \leq \mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta \geq \frac{1}{3} \right] = \exp\left(-\frac{\lambda_{\tilde{c}, M}}{3}\right) \quad (9)$$

where in the first equality we used that $\tilde{c} \geq 1$ and in the last equality we used that $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq x] = \exp(-\lambda_{\tilde{c}, M} x)$. Using that $\lambda_{\tilde{c}, M} = 3 \ln(2^{2\tilde{c}} M)$, (9) yields $\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{1}{\tilde{c} + 1} \geq 1 - \frac{1}{\tilde{c} + 2} \right] \leq \frac{1}{2^{2\tilde{c}} M}$. Combining (8) and (9) with (7) yields $\chi_{\tilde{c}}(p) \geq \frac{1}{2^{2\tilde{c}}} 2^{2\tilde{c}} M = M$ as desired.

To prove (b), the formulas for the revenue of σ^{RANDOM} and σ^{NEWEST} given by Propositions 3.1 and 3.2 and the fact that $\mu = \frac{1}{2}$ and that h is the mean imply that $\chi_c(p)$ can be expressed as

$$\frac{\text{REV}_c(\sigma^{\text{RANDOM}}, p)}{\text{REV}_c(\sigma^{\text{NEWEST}}, p)} = \mathbb{E}_{N \sim \text{Binom}(c, \frac{1}{2})} \left[\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{1 + N}{2 + c} \geq p \right] \right] \mathbb{E}_{N \sim \text{Binom}(c, \frac{1}{2})} \left[\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{1 + N}{2 + c} \geq p \right]} \right].$$

For any number of positive reviews $N \in \{0, 1, \dots, c\}$, the purchase probability equals

$$\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{1+N}{2+c} \geq p \right] = \mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta \geq 1 - \frac{1}{2+\tilde{c}} - \frac{1+N}{2+c} \right] = \exp(-\lambda_{\tilde{c},M}(1 - \frac{1}{2+\tilde{c}} - \frac{1+N}{2+c})) \quad (10)$$

Using (10) in the expression for $\chi_c(p)$, the term $\exp(-\lambda_{\tilde{c},M}(1 - \frac{1}{2+\tilde{c}}))$ cancels out yielding

$$\begin{aligned} \chi_c(p) &= \mathbb{E}_{N \sim \text{Binom}(c, \frac{1}{2})} \left[\exp(-\lambda_{\tilde{c},M}(1 - \frac{1}{2+\tilde{c}} - \frac{1+N}{2+c})) \right] \mathbb{E}_{N \sim \text{Binom}(c, \frac{1}{2})} \left[\exp(\lambda_{\tilde{c},M}(1 - \frac{1}{2+\tilde{c}} - \frac{1+N}{2+c})) \right] \\ &= \mathbb{E}_{N \sim \text{Binom}(c, \frac{1}{2})} \left[\exp(\lambda_{\tilde{c},M} \frac{1+N}{2+c}) \right] \mathbb{E}_{N \sim \text{Binom}(c, \frac{1}{2})} \left[\exp(-\lambda_{\tilde{c},M} \frac{1+N}{2+c}) \right]. \end{aligned} \quad (11)$$

Let $\mathcal{E}^{\text{CONC}}$ be the event that $\{N \in [\frac{1}{2}c - c^{\frac{2}{3}}, \frac{1}{2}c + c^{\frac{2}{3}}]\}$. The Chernoff bound implies that $\mathbb{P}[\mathcal{E}^{\text{CONC}}] \geq 1 - 2e^{-\frac{c^{\frac{1}{3}}}{3}}$. The law of total expectation applied to the first expectation in (11) yields

$$\begin{aligned} \mathbb{E}_{N \sim \text{Binom}(c, \frac{1}{2})} \left[\exp(\lambda_{\tilde{c},M} \frac{1+N}{2+c}) \right] &= \underbrace{\mathbb{E}_{N \sim \text{Binom}(c, \frac{1}{2})} \left[\exp(\lambda_{\tilde{c},M} \frac{1+N}{2+c}) | \mathcal{E}^{\text{CONC}} \right] \mathbb{P}[\mathcal{E}^{\text{CONC}}]}_{=(1)} \\ &\quad + \underbrace{\mathbb{E}_{N \sim \text{Binom}(c, \frac{1}{2})} \left[\exp(\lambda_{\tilde{c},M} \frac{1+N}{2+c}) | (\mathcal{E}^{\text{CONC}})^{\perp} \right] \mathbb{P}[(\mathcal{E}^{\text{CONC}})^{\perp}]}_{=(2)} \end{aligned}$$

Notice that $\exp(\lambda_{\tilde{c},M}(\frac{1+N}{2+c})) \in [1, \exp(\lambda_{\tilde{c},M})]$ for all $N \in \{0, \dots, c\}$. Combining this with $\lim_{c \rightarrow \infty} \mathbb{P}[(\mathcal{E}^{\text{CONC}})^{\perp}] = 0$ yields that the term (2) goes to 0 as $c \rightarrow \infty$. Using the definition of $\mathcal{E}^{\text{CONC}}$, we can upper and lower bound the term (1) as

$$\exp\left(\lambda_{\tilde{c},M} \frac{1 + \frac{1}{2}c - c^{\frac{2}{3}}}{2+c}\right) \mathbb{P}[\mathcal{E}^{\text{CONC}}] \leq (1) \leq \exp\left(\lambda_{\tilde{c},M} \frac{1 + \frac{1}{2}c + c^{\frac{2}{3}}}{2+c}\right) \mathbb{P}[\mathcal{E}^{\text{CONC}}].$$

As the $\lim_{c \rightarrow \infty} \mathbb{P}[\mathcal{E}^{\text{CONC}}] = 1$, both right-hand side and left-hand side converge to $\exp\left(\frac{\lambda_{\tilde{c},M}}{2}\right)$.

Thus, $\lim_{c \rightarrow \infty} \mathbb{E}_{N \sim \text{Binom}(c, \frac{1}{2})} \left[\exp(\lambda_{\tilde{c},M} \frac{1+N}{2+c}) \right] = \exp\left(\frac{\lambda_{\tilde{c},M}}{2}\right)$. Similarly, the law of total expectation applied to the second expectation in (11) yields

$$\begin{aligned} \mathbb{E}_{N \sim \text{Binom}(c, \frac{1}{2})} \left[\exp(-\lambda_{\tilde{c},M} \frac{1+N}{2+c}) \right] &= \underbrace{\mathbb{E}_{N \sim \text{Binom}(c, \frac{1}{2})} \left[\exp(-\lambda_{\tilde{c},M} \frac{1+N}{2+c}) | \mathcal{E}^{\text{CONC}} \right] \mathbb{P}[\mathcal{E}^{\text{CONC}}]}_{=(3)} \\ &\quad + \underbrace{\mathbb{E}_{N \sim \text{Binom}(c, \frac{1}{2})} \left[\exp(-\lambda_{\tilde{c},M} \frac{1+N}{2+c}) | (\mathcal{E}^{\text{CONC}})^{\perp} \right] \mathbb{P}[(\mathcal{E}^{\text{CONC}})^{\perp}]}_{=(4)} \end{aligned}$$

Notice that $\exp(-\lambda_{\tilde{c},M}(\frac{1+N}{2+c})) \in [\exp(-\lambda_{\tilde{c},M}), 1]$ for all $N \in \{0, \dots, c\}$. Combining this with $\lim_{c \rightarrow \infty} \mathbb{P}[(\mathcal{E}^{\text{CONC}})^{\perp}] = 0$ yields that the term (4) goes to 0 as $c \rightarrow \infty$. Using the definition of $\mathcal{E}^{\text{CONC}}$, we can upper and lower bound the term (3) as

$$\exp\left(-\lambda_{\tilde{c},M} \frac{1 + \frac{1}{2}c + c^{\frac{2}{3}}}{2+c}\right) \mathbb{P}[\mathcal{E}^{\text{CONC}}] \leq (3) \leq \exp\left(-\lambda_{\tilde{c},M} \frac{1 + \frac{1}{2}c - c^{\frac{2}{3}}}{2+c}\right) \mathbb{P}[\mathcal{E}^{\text{CONC}}].$$

As the $\lim_{c \rightarrow \infty} \mathbb{P}[\mathcal{E}^{\text{CONC}}] = 1$, both right-hand side and left-hand side converge to $\exp\left(-\frac{\lambda_{\tilde{c},M}}{2}\right)$.

Thus, $\lim_{c \rightarrow \infty} \mathbb{E}_{N \sim \text{Binom}(c, \frac{1}{2})} \left[\exp(-\lambda_{\tilde{c},M} \frac{1+N}{2+c}) \right] = \exp\left(-\frac{\lambda_{\tilde{c},M}}{2}\right)$. Thus, $\lim_{c \rightarrow \infty} \chi_c(p) = 1$. \square

Proposition B.4. Fix an instance $(\mathcal{F}, \mu, a, b, p)$ such that \mathcal{F} is continuous, and suppose that c varies. If the purchase probability is lower bounded by $\delta > 0$ for all review states and for all c , then $\lim_{c \rightarrow \infty} \chi_c(p) = 1$.

The instance $\mathcal{E}_{\tilde{c}, M}$ of Proposition B.3 has a lower bound $\delta = \exp(-3 \ln(2^{2\tilde{c}} M)) = \frac{1}{2^{6\tilde{c}} M^3}$. This implies that the existence of a lower bound $\delta > 0$ does not imply a bounded CoNF with limited attention. We note that assuming a lower bound on the purchase probability is common in the literature on social learning with reviews [CIMS17, BS18, IMSZ19].

Proof of Proposition B.4. Let $\mathcal{E} = (\mathcal{F}, \mu, a, b, p)$ be an instance such that the purchase probability is lower bounded by $\delta > 0$ for any limited attention window $c \geq 1$ and number of positive reviews $N \in \{0, 1, \dots, c\}$, i.e.,

$$\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{N+a}{a+b+c} \geq p \right] \geq \delta \text{ for all } c \geq 1, N \in \{0, \dots, c\}. \quad (12)$$

Using the formulas for the revenue of σ^{RANDOM} and σ^{NEWEST} given by Propositions 3.1 and 3.2, the CoNF can be expressed as

$$\frac{\text{REV}_c(\sigma^{\text{RANDOM}}, p)}{\text{REV}_c(\sigma^{\text{NEWEST}}, p)} = \mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{N+a}{a+b+c} \geq p \right] \right] \mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{N+a}{a+b+c} \geq p \right]} \right].$$

Let $\mathcal{E}^{\text{CONC}}$ be the event that $\{N \in [\mu c - c^{\frac{2}{3}}, \mu c + c^{\frac{2}{3}}]\}$. The Chernoff bound implies that $\mathbb{P}[\mathcal{E}^{\text{CONC}}] \geq 1 - 2e^{-\frac{1}{3}}$. The law of total expectation applied to the first expectation yields

$$\begin{aligned} \mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{N+a}{a+b+c} \geq p \right] \right] &= \underbrace{\mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{N+a}{a+b+c} \geq p \right] | \mathcal{E}^{\text{CONC}} \right] \mathbb{P}[\mathcal{E}^{\text{CONC}}]}_{=(1)} \\ &\quad + \underbrace{\mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{N+a}{a+b+c} \geq p \right] | (\mathcal{E}^{\text{CONC}})^\perp \right] \mathbb{P}[(\mathcal{E}^{\text{CONC}})^\perp]}_{=(2)} \end{aligned}$$

Notice that $\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{N+a}{a+b+c} \geq p \right] \in [0, 1]$ for any $N \in \{0, \dots, c\}$. Combining this with $\mathbb{P}[(\mathcal{E}^{\text{CONC}})^\perp] \rightarrow_{c \rightarrow \infty} 0$, yields that (2) converges to zero as $c \rightarrow \infty$. Using the definition of $\mathcal{E}^{\text{CONC}}$ we can lower and upper bound the term (1) as

$$\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{\mu c + c^{\frac{2}{3}} + a}{a+b+c} \geq p \right] \mathbb{P}[\mathcal{E}^{\text{CONC}}] \geq (1) \geq \mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{\mu c - c^{\frac{2}{3}} + a}{a+b+c} \geq p \right] \mathbb{P}[\mathcal{E}^{\text{CONC}}].$$

Given that $\frac{\mu c + c^{\frac{2}{3}} + a}{a+b+c}, \frac{\mu c - c^{\frac{2}{3}} + a}{a+b+c} \rightarrow_{c \rightarrow \infty} \mu$ and \mathcal{F} is continuous $\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{\mu c + c^{\frac{2}{3}} + a}{a+b+c} \geq p \right], \mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{\mu c - c^{\frac{2}{3}} + a}{a+b+c} \geq p \right] \rightarrow_{c \rightarrow \infty} \mathbb{P}_{\Theta \sim \mathcal{F}} [\Theta + \mu \geq p]$. Combining this with $\lim_{c \rightarrow \infty} \mathbb{P}[\mathcal{E}^{\text{CONC}}] = 1$ yields that (1) converges to $\mathbb{P}[\Theta + \mu \geq p]$ as $c \rightarrow \infty$. Thus, $\lim_{c \rightarrow \infty} \mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{N+a}{a+b+c} \geq p \right] \right] =$

$\mathbb{P}[\Theta + \mu \geq p]$. The law of total expectation applied to the second expectation yields

$$\begin{aligned} \mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{N+a}{a+b+c} \geq p \right]} \right] &= \underbrace{\mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{N+a}{a+b+c} \geq p \right]} | \mathcal{E}^{\text{CONC}} \right]}_{=(3)} \mathbb{P}[\mathcal{E}^{\text{CONC}}] \\ &+ \underbrace{\mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{N+a}{a+b+c} \geq p \right]} | (\mathcal{E}^{\text{CONC}})^\perp \right]}_{=(4)} \mathbb{P}[(\mathcal{E}^{\text{CONC}})^\perp] \end{aligned}$$

Notice that $\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{N+a}{a+b+c} \geq p \right]} \in [0, \frac{1}{\delta}]$ for any $N \in \{0, \dots, c\}$ due to (12). Combining this with $\mathbb{P}[(\mathcal{E}^{\text{CONC}})^\perp] \rightarrow_{c \rightarrow \infty} 0$, yields that (4) converges to zero as $c \rightarrow \infty$. Using the definition of $\mathcal{E}^{\text{CONC}}$ we can lower and upper bound the term (3) as

$$\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{\mu c - c \frac{2}{3} + a}{a+b+c} \geq p \right]} \mathbb{P}[\mathcal{E}^{\text{CONC}}] \geq (3) \geq \frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{\mu c + c \frac{2}{3} + a}{a+b+c} \geq p \right]} \mathbb{P}[\mathcal{E}^{\text{CONC}}].$$

Given that $\frac{\mu c + c \frac{2}{3} + a}{a+b+c}, \frac{\mu c - c \frac{2}{3} + a}{a+b+c} \rightarrow_{c \rightarrow \infty} \mu$ and \mathcal{F} is continuous $\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{\mu c + c \frac{2}{3} + a}{a+b+c} \geq p \right], \mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{\mu c - c \frac{2}{3} + a}{a+b+c} \geq p \right] \rightarrow_{c \rightarrow \infty} \mathbb{P}_{\Theta \sim \mathcal{F}} [\Theta + \mu \geq p]$. Combining this with $\lim_{c \rightarrow \infty} \mathbb{P}[\mathcal{E}^{\text{CONC}}] = 1$ yields that (3) converges to $\frac{1}{\mathbb{P}[\Theta + \mu \geq p]}$ as $c \rightarrow \infty$. Thus, $\lim_{c \rightarrow \infty} \mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + \frac{N+a}{a+b+c} \geq p \right]} \right] = \frac{1}{\mathbb{P}[\Theta + \mu \geq p]}$, concluding the proof. \square

C Supplementary material for Section 4

C.1 CoNF still exists under optimal dynamic pricing (Proposition 4.1)

Proof of Proposition 4.1. For any $\alpha < 4/3$, we provide an instance with $\chi(\Pi^{\text{DYNAMIC}}) = \alpha$. The instance consists of a single review ($c = 1$), true quality $\mu < \frac{1}{3}$, estimate mappings $h(0) = \mu^2$, $h(1) = 1 - \mu^2$, and customer-specific valuation $\mathcal{F} = \mathcal{U}[0, 2\bar{h}]$ where $\bar{h} := \mathbb{E}_N[h(N)]$.

First, for any $A > 0$ and any price p the revenue of selling to a customer with valuation $V = Y + x$ where $Y \sim \mathcal{U}[0, A]$ is given by $p \cdot \mathbb{P}[V \geq p] = \frac{p(A+x-p)}{A}$ if $p \in [x, A+x]$. Thus, the revenue maximizing price is $p^* = \max(\frac{A+x}{2}, x) = \frac{\max(x, A) + x}{2}$ yielding an optimal revenue of

$$\max_p \{p \cdot \mathbb{P}[V \geq p]\} = \frac{\max(x, A) + x}{2} \cdot \frac{2A + x - \max(A, x)}{2A} = \frac{(\max(A, x) + x)^2}{4 \max(A, x)}. \quad (13)$$

Theorems 4.2 and 4.1 imply that $\chi(\Pi^{\text{DYNAMIC}}) = \frac{\mathbb{E}_N[r^*(\Theta + h(N))]}{r^*(\Theta + \bar{h})}$. Using (13) with $A = 2\bar{h}$, we obtain

$$r^*(\Theta + x) = \frac{(\max(2\bar{h}, x) + x)^2}{4 \max(2\bar{h}, x)} = \begin{cases} \frac{9}{8}\bar{h} & \text{if } x = \bar{h} \\ h(1) & \text{if } x = h(1) \\ \frac{(h(0) + 2\bar{h})^2}{8\bar{h}} & \text{if } x = h(0) \end{cases} \quad (14)$$

where $\max(2\bar{h}, h(1)) = h(1)$ because $2\mu(1 - \mu)(1 + 2\mu) = 2\mu(1 + \mu - 2\mu^2) < 2\mu(1 + \mu) < \frac{8}{9} < 1 - \mu^2$ for $\mu < \frac{1}{3}$ and thus

$$2\bar{h} = 2(\mu h(1) + (1 - \mu)h(0)) = 2(\mu(1 - \mu^2) + (1 - \mu)\mu^2) = 2\mu(1 - \mu)(1 + 2\mu) < 1 - \mu^2 = h(1).$$

Using (14), the CoNF can be expressed as:

$$\begin{aligned}
\chi(\Pi^{\text{DYNAMIC}}) &= \frac{\mathbb{E}_N[r^*(\Theta + h(N))]}{r^*(\Theta + \bar{h})} = \frac{\mu r^*(\Theta + h(1)) + (1 - \mu)r^*(\Theta + h(0))}{r^*(\Theta + \bar{h})} \\
&\stackrel{(1)}{=} \frac{\mu h(1) + (1 - \mu) \frac{(h(0) + 2\bar{h})^2}{8\bar{h}}}{\frac{9}{8}\bar{h}} = \frac{8}{9} \cdot \frac{\mu h(1)}{\bar{h}} + \frac{1 - \mu}{9} \left[\left(\frac{h(0)}{\bar{h}} \right)^2 + 4 \frac{h(0)}{\bar{h}} + 4 \right] \\
&\stackrel{(2)}{=} \frac{8}{9} - \frac{8}{9} \cdot \frac{(1 - \mu)h(0)}{\bar{h}} + \frac{1 - \mu}{9} \cdot \left(\frac{h(0)}{\bar{h}} \right)^2 + \frac{1 - \mu}{9} \cdot 4 \frac{h(0)}{\bar{h}} + \frac{1 - \mu}{9} \cdot 4 \\
&\stackrel{(3)}{=} \frac{8}{9} + \frac{1 - \mu}{9} \left[\left(\frac{h(0)}{\bar{h}} \right)^2 - 4 \frac{h(0)}{\bar{h}} + 4 \right] = \frac{8}{9} + \frac{1 - \mu}{9} \left(\frac{h(0)}{\bar{h}} - 2 \right)^2 \\
&\stackrel{(4)}{=} \frac{8}{9} + \frac{1 - \mu}{9} \left(\frac{\mu}{1 + \mu - 2\mu^2} - 2 \right)^2
\end{aligned}$$

where (1) follows by the expression for the optimal revenue (14), (2) follows by $\frac{\mu h(1)}{\bar{h}} = 1 - \frac{(1 - \mu)h(0)}{\bar{h}}$ which follows by the definition of \bar{h} , (3) follows by $-\frac{8}{9} \cdot \frac{(1 - \mu)h(0)}{\bar{h}} + \frac{1 - \mu}{9} \cdot 4 \frac{h(0)}{\bar{h}} = -\frac{1 - \mu}{9} \cdot 4 \frac{h(0)}{\bar{h}}$, and (4) follows by $\frac{h(0)}{\bar{h}} = \frac{\mu^2}{\mu(1 - \mu)(1 + 2\mu)} = \frac{\mu}{1 + \mu - 2\mu^2}$ because $h(0) = \mu^2$ and $\bar{h} = \mu(1 - \mu)(1 + 2\mu)$.

Thus, $\chi(\Pi^{\text{DYNAMIC}}) = \frac{8}{9} + \frac{1 - \mu}{9} \left(\frac{\mu}{1 + \mu - 2\mu^2} - 2 \right)^2$. The second term is $\frac{1 - \mu}{9} \left(\frac{\mu}{1 + \mu - 2\mu^2} - 2 \right)^2 \rightarrow \frac{4}{9}$ as $\mu \rightarrow 0$, and thus $\chi(\Pi^{\text{DYNAMIC}}) \rightarrow \frac{8}{9} + \frac{4}{9} = \frac{4}{3}$ as $\mu \rightarrow 0$. Therefore, for any $\alpha < 4/3$ there is some sufficiently small $\mu > 0$ such that $\chi(\Pi^{\text{DYNAMIC}}) > \alpha$. \square

C.2 σ^{random} is no worse than σ^{newest} under optimal dynamic pricing (Remark 4.1)

Theorem 3.1 established that $\chi(p) > 1$ for any price p satisfying Assumption 3.1. Here we show that if the platform optimizes over dynamic prices we have $\chi(\Pi^{\text{DYNAMIC}}) \geq 1$ i.e. σ^{RANDOM} has no smaller revenue than σ^{NEWEST} under optimal dynamic prices. Recall that $\chi(\Pi^{\text{DYNAMIC}}) = \frac{\max_{\rho \in \Pi^{\text{DYNAMIC}}} \text{REV}(\sigma^{\text{RANDOM}}, \rho)}{\max_{\rho \in \Pi^{\text{DYNAMIC}}} \text{REV}(\sigma^{\text{NEWEST}}, \rho)}$. We show the following result.

Proposition C.1. *For any problem instance, $\chi(\Pi^{\text{DYNAMIC}}) \geq 1$.*

Proof. By definition of $\chi(\Pi^{\text{DYNAMIC}})$, it is sufficient to show that

$$\max_{\rho \in \Pi^{\text{DYNAMIC}}} \text{REV}(\sigma^{\text{NEWEST}}, \rho) \leq \max_{\rho \in \Pi^{\text{DYNAMIC}}} \text{REV}(\sigma^{\text{RANDOM}}, \rho). \quad (15)$$

Letting $\bar{h} = \mathbb{E}_{N \sim \text{Binom}(c, \mu)}[h(N)]$ and $p^* \in \arg \max_{p \in \mathbb{R}} p \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + \bar{h} \geq p]$, Theorem 4.3 yields that the pricing policy given by $\rho^{\text{NEWEST}}(\mathbf{z}) = h(N_{\mathbf{z}}) + p^* - \bar{h}$ is optimal under σ^{NEWEST} and is thus the maximizer of the left-hand side of (15). To prove (15) it is sufficient to show that offering prices $\rho^{\text{NEWEST}}(\mathbf{z})$ under σ^{RANDOM} yields the same revenue as offering prices $\rho^{\text{NEWEST}}(\mathbf{z})$ under σ^{NEWEST} i.e. $\text{REV}(\sigma^{\text{RANDOM}}, \rho^{\text{NEWEST}}) = \text{REV}(\sigma^{\text{NEWEST}}, \rho^{\text{NEWEST}})$.

As σ^{RANDOM} shows c i.i.d. $\text{Bern}(\mu)$ reviews, the revenue of any pricing policy ρ is given by

$$\text{REV}(\sigma^{\text{RANDOM}}, \rho) = \mathbb{E}_{Z_1, \dots, Z_c \sim \text{i.i.d. Bern}(\mu)} \left[\rho(Z_1, \dots, Z_c) \mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + h\left(\sum_{i=1}^c Z_i\right) \geq \rho(Z_1, \dots, Z_c) \right] \right]. \quad (16)$$

For any reviews (Z_1, \dots, Z_c) the purchase probability under $\rho^{\text{NEWEST}}(Z_1, \dots, Z_c)$ equals

$$\mathbb{P}_{\Theta \sim \mathcal{F}} \left[\Theta + h\left(\sum_{i=1}^c Z_i\right) \geq h\left(\sum_{i=1}^c Z_i\right) + p^* - \bar{h} \right] = \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + \bar{h} \geq p^*] \quad (17)$$

and is independent of (Z_1, \dots, Z_c) . As $Z_1, \dots, Z_c \sim_{i.i.d.} \text{Bern}(\mu)$, the number of positive review ratings $\sum_{i=1}^c Z_i$ is distributed as $N \sim \text{Binom}(c, \mu)$; the expected price of $\rho^{\text{NEWEST}}(Z_1, \dots, Z_c)$ is thus

$$\mathbb{E}_{Z_1, \dots, Z_c \sim_{i.i.d.} \text{Bern}(\mu)} [\rho^{\text{NEWEST}}(Z_1, \dots, Z_c)] = \mathbb{E}_{N \sim \text{Binom}(c, \mu)} [h(N) + p^* - \bar{h}] = p^* \quad (18)$$

where the last equality uses the definition of $\bar{h} = \mathbb{E}_{N \sim \text{Binom}(c, \mu)} [h(N)]$. Combining (16), (17), and (18), we obtain $\text{REV}(\sigma^{\text{RANDOM}}, \rho) = p^* \mathbb{P}_{\Theta \sim \mathcal{F}} [\Theta + \bar{h} \geq p^*] = r^*(\Theta + \bar{h})$. The last expression is exactly equal to $\max_{\rho \in \Pi^{\text{DYNAMIC}}} \text{REV}(\sigma^{\text{NEWEST}}, \rho)$ by Corollary 4.1 which concludes the proof. \square

C.3 Stronger CoNF upper bound under further assumptions on h (Remark 4.2)

Proposition C.2. *For any instance with $h(n) \leq u, \forall n \in \{0, 1, \dots, c\}$: $\chi(\Pi^{\text{DYNAMIC}}) \leq \frac{2\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq -u]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq 0]}$.*

Proof. Similar to the proof of Theorem 4.2, we define $\tilde{p}_n = \bar{h} + \max(p^*(\Theta + h(n)) - h(n), 0)$ and let $p(n) = p^*(\Theta + h(n))$ for convenience. We refine the analysis of Lemma 4.3, to show that for any $n \in \{0, 1, \dots, c\}$, $\text{DEMAND RATIO}(\tilde{p}_n, n) \leq \frac{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq -u]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq 0]}$. We consider two cases:

- If $p(n) \geq h(n)$, then $\max(p(n) - h(n), 0) = p(n) - h(n)$ and thus

$$\text{DEMAND RATIO}(\tilde{p}_n, n) = \frac{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq p(n) - h(n)]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq \max(p(n) - h(n), 0)]} = 1 \leq \frac{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq -u]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq 0]}.$$

since $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq -u] \geq \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq 0]$ as $u > 0$. If $p(n) < h(n)$, then

- If $p(n) < h(n)$, then $\max(p(n) - h(n), 0) = 0$

$$\begin{aligned} \text{DEMAND RATIO}(\tilde{p}_n, n) &= \frac{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq p(n) - h(n)]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq \max(p(n) - h(n), 0)]} \\ &= \frac{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq p(n) - h(n)]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq 0]} \leq \frac{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq -h(n)]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq 0]} \leq \frac{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq -a]}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq 0]}. \end{aligned}$$

The first inequality uses $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq p(n) - h(n)] \leq \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq -h(n)]$ as the optimal price $p(n) \geq 0$. The second inequality uses $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq -h(n)] \leq \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq -u]$ as $h(n) \leq u$.

By Lemma 4.2, the expected price ratio is upper bounded as $\mathbb{E}[\text{PRICE RATIO}(\tilde{p}_N, N)] \leq 2$. Combining this with the aforementioned bound on the demand ratio the proof follows. \square

C.4 Dynamic pricing revenue of Newest First (Lemma 4.1)

As an analogue of Lemma 3.1, let $\mathbf{Z}_t = (Z_{t,1}, \dots, Z_{t,c}) \in \{0, 1\}^c$ denote the process of the newest c reviews. We note that \mathbf{Z}_t is a time-homogenous Markov chain on a finite state space $\{0, 1\}^c$.

If \mathbf{Z}_t is at state $\mathbf{z} = (z_1, \dots, z_c)$ it stays at that state if there is no purchase (with probability $1 - \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N_{\mathbf{z}}) \geq \rho(\mathbf{z})]$). If there is a purchase with a positive review, it transitions to $(1, z_1, \dots, z_{c-1})$ (with probability $\mu \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N_{\mathbf{z}}) \geq \rho(\mathbf{z})]$). If there is a purchase with a negative review, it transitions to $(0, z_1, \dots, z_{c-1})$ (with probability $(1 - \mu) \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N_{\mathbf{z}}) \geq \rho(\mathbf{z})]$). If $\mu \in (0, 1)$ and ρ induces positive purchase probabilities in all states, \mathbf{Z}_t is a single-recurrence-class Markov chain with no transient states. Thus, \mathbf{Z}_t admits a unique stationary distribution characterized in the following lemma. Recall that for a state of c reviews $\mathbf{z} \in \{0, 1\}^c$, $N_{\mathbf{z}} = \sum_{i=1}^c z_i$ denotes the number of positive review ratings.

Lemma C.1. *The stationary distribution of $\mathbf{Z}_t \in \{0, 1\}^c$ under any dynamic pricing policy ρ with positive purchase probabilities in all states is*

$$\pi_{\mathbf{z}} = \kappa \cdot \frac{\mu^{N_{\mathbf{z}}}(1 - \mu)^{c - N_{\mathbf{z}}}}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N_{\mathbf{z}}) \geq \rho(\mathbf{z})]} \text{ for } \mathbf{z} \in \{0, 1\}^c,$$

where $\kappa = 1/\mathbb{E}_{Y_1, \dots, Y_c \sim i.i.d. \text{Bern}(\mu)} \left[\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(\sum_{i=1}^c Y_i) \geq \rho(Y_1, \dots, Y_c)]} \right]$ is a normalizing constant.

Proof of Lemma C.1. Similar to the proof of Lemma 3.1, we invoke Lemma B.2. In the language of Lemma B.2, \mathbf{Z}_t corresponds to \mathcal{M}_f , the state space \mathcal{S} to $\{0, 1\}^c$, and f is a function that expresses the purchase probability at a given state, i.e., $f(z_1, \dots, z_c) = \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N_{\mathbf{z}}) \geq \rho(\mathbf{z})]$. Note that with probability $1 - f(z_1, \dots, z_c)$, \mathbf{Z}_t remains at the same state (as there is no purchase).

To apply Lemma B.2, we need to show that whenever there is purchase, \mathbf{Z}_t transitions according to a Markov chain with stationary distribution $\mu^{N_{\mathbf{z}}}(1 - \mu)^{c - N_{\mathbf{z}}}$. Consider the Markov chain \mathcal{M} which always replaces the c -th last review with a new $\text{Bern}(\mu)$ review. This process has stationary distribution equal to the above numerator and \mathbf{Z}_t transitions according to \mathcal{M} upon a purchase, i.e., with probability $f(z_1, \dots, z_c)$. As a result, by Lemma B.2, π is a stationary distribution for \mathbf{Z}_t . As \mathbf{Z}_t is irreducible and aperiodic, this is the unique stationary distribution. \square

Proof of Lemma 4.1. By Eq. (1) and the Ergodic theorem, we can express the revenue as

$$\begin{aligned} \text{REV}(\sigma^{\text{NEWEST}}, \rho) &= \liminf_{T \rightarrow \infty} \frac{\mathbb{E} \left[\sum_{t=1}^T \rho(Z_{t,1}, \dots, Z_{t,c}) \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(\sum_{i=1}^c Z_{t,i}) \geq \rho(Z_{t,1}, \dots, Z_{t,c})] \right]}{T} \\ &= \sum_{(z_1, \dots, z_c) \in \{0,1\}^c} \pi_{(z_1, \dots, z_c)} \rho(z_1, \dots, z_c) \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N_{\mathbf{z}}) \geq \rho(z_1, \dots, z_c)] \\ &= \kappa \cdot \sum_{(z_1, \dots, z_c) \in \{0,1\}^c} \mu^{\sum_{i=1}^c z_i} (1 - \mu)^{c - \sum_{i=1}^c z_i} \rho(z_1, \dots, z_c) \\ &= \kappa \cdot \mathbb{E}_{Y_1, \dots, Y_c \sim i.i.d. \text{Bern}(\mu)} [\rho(Y_1, \dots, Y_c)] \\ &= \frac{\mathbb{E}_{Y_1, \dots, Y_c \sim i.i.d. \text{Bern}(\mu)} [\rho(Y_1, \dots, Y_c)]}{\mathbb{E}_{Y_1, \dots, Y_c \sim i.i.d. \text{Bern}(\mu)} \left[\frac{1}{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(\sum_{i=1}^c Y_i) \geq \rho(Y_1, \dots, Y_c)]} \right]}. \end{aligned}$$

The third equality applies Lemma C.1 and cancels the term $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N_{\mathbf{z}}) \geq \rho(z_1, \dots, z_c)]$. \square

C.5 Ratio of averages is bounded by maximum ratio (Proof of Theorem 4.3)

Lemma C.2. *Let $\{\tilde{\alpha}_i, \tilde{A}_i, \tilde{B}_i\}_{i \in \mathcal{S}}$ be such that $\tilde{\alpha}_i, \tilde{B}_i > 0$ for all i . Then $\frac{\sum_{i \in \mathcal{S}} \tilde{\alpha}_i \tilde{A}_i}{\sum_{i \in \mathcal{S}} \tilde{\alpha}_i \tilde{B}_i} \leq \max_{i \in \mathcal{S}} \frac{\tilde{A}_i}{\tilde{B}_i}$. Equality is achieved if and only if $\frac{\tilde{A}_i}{\tilde{B}_i} = \frac{\tilde{A}_j}{\tilde{B}_j}$ for all $i, j \in \mathcal{S}$.*

Proof of Lemma C.2. We show the inequality by contradiction and assume that $\frac{\sum_{i \in \mathcal{S}} \tilde{\alpha}_i \tilde{A}_i}{\sum_{i \in \mathcal{S}} \tilde{\alpha}_i \tilde{B}_i} > \frac{\tilde{A}_j}{\tilde{B}_j}$ for all $j \in \mathcal{S}$. Given that the denominators are positive, this implies that for any $j \in \mathcal{S}$

$$\left(\sum_{i \in \mathcal{S}} \tilde{\alpha}_i \tilde{A}_i \right) (\tilde{\alpha}_j \tilde{B}_j) > (\tilde{\alpha}_j \tilde{A}_j) \left(\sum_{i \in \mathcal{S}} \tilde{\alpha}_i \tilde{B}_i \right).$$

Summing over $j \in \mathcal{S}$ we obtain the following which is a contradiction:

$$\left(\sum_{i \in \mathcal{S}} \tilde{\alpha}_i \tilde{A}_i\right) \left(\sum_{j \in \mathcal{S}} \tilde{\alpha}_j \tilde{B}_j\right) > \left(\sum_{j \in \mathcal{S}} \tilde{\alpha}_j \tilde{A}_j\right) \left(\sum_{i \in \mathcal{S}} \tilde{\alpha}_i \tilde{B}_i\right)$$

As a result $\frac{\sum_{i \in \mathcal{S}} \tilde{\alpha}_i \tilde{A}_i}{\sum_{i \in \mathcal{S}} \tilde{\alpha}_i \tilde{B}_i} \leq \max_{i \in \mathcal{S}} \frac{\tilde{A}_i}{\tilde{B}_i}$. With respect to equality, let $j \in \arg \max_k \frac{\tilde{A}_k}{\tilde{B}_k}$. Thus,

$$\frac{\sum_{i \in \mathcal{S}} \tilde{\alpha}_i \tilde{A}_i}{\sum_{i \in \mathcal{S}} \tilde{\alpha}_i \tilde{B}_i} = \frac{\tilde{A}_j}{\tilde{B}_j}. \quad (19)$$

Multiplying by the denominators and rearranging the above can be rewritten as

$$\sum_{i \in \mathcal{S}} \tilde{\alpha}_i \tilde{B}_j \tilde{B}_i \left(\frac{\tilde{A}_i}{\tilde{B}_i} - \frac{\tilde{A}_j}{\tilde{B}_j} \right) = 0. \quad (20)$$

Since $\tilde{\alpha}_i > 0$, $\tilde{B}_i > 0$ for all $i \in \mathcal{S}$, and $\frac{\tilde{A}_i}{\tilde{B}_i} \leq \frac{\tilde{A}_j}{\tilde{B}_j}$ for all $i \in \mathcal{S}$, (20) holds only if $\frac{\tilde{A}_i}{\tilde{B}_i} = \frac{\tilde{A}_j}{\tilde{B}_j}$ for all $i \in \mathcal{S}$. For the “if” direction, suppose that $\frac{\tilde{A}_i}{\tilde{B}_i} = \frac{\tilde{A}_j}{\tilde{B}_j}$ for all i, j . In particular this holds, when $j \in \arg \max_k \frac{\tilde{A}_k}{\tilde{B}_k}$ is a maximizing index and $i \in \mathcal{S}$ is an arbitrary index. This implies that (20) holds for any $j \in \arg \max_k \frac{\tilde{A}_k}{\tilde{B}_k}$ and therefore so does (19), which concludes the proof. \square

C.6 Characterization of optimal revenue under Newest First (Corollary 4.1)

Proof of Corollary 4.1. Letting $\bar{h} = \mathbb{E}_{N \sim \text{Binom}(c, \mu)}[h(N)]$, Theorem 4.3 establishes that the optimal dynamic pricing policy is review-offsetting with offset $p^* - \bar{h}$. Using the expression for the revenue of review-offsetting policies with offset $a^* = p^* - \bar{h}$ given by (2):

$$\max_{\rho \in \Pi^{\text{DYNAMIC}}} \text{REV}(\sigma^{\text{NEWEST}}, \rho) = (a^* + \bar{h}) \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + \bar{h} \geq a^* + \bar{h}] = p^* \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + \bar{h} \geq p^*] = r^*(\Theta + \bar{h}).$$

\square

C.7 All optimal dynamic policies under Newest First (Remark 4.3)

Proposition C.3. *Suppose that the platform uses σ^{NEWEST} as the review ordering policy. Let $p_{\mathbf{z}} \in \arg \max_p p \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + \mathbb{E}_{N \sim \text{Binom}(c, \mu)}[h(N)] \geq p]$ be any revenue-maximizing price for $\mathbf{z} \in \{0, 1\}^c$. A dynamic pricing policy ρ^{NEWEST} is optimal if and only if it has the form*

$$\rho^{\text{NEWEST}}(\mathbf{z}) = h(N_{\mathbf{z}}) + p_{\mathbf{z}} - \mathbb{E}_{N \sim \text{Binom}(c, \mu)}[h(N)].$$

For an arbitrary dynamic pricing policy ρ , letting $\bar{h} = \mathbb{E}_{N \sim \text{Binom}(c, \mu)}[h(N)]$, the optimality condition of the theorem can be rewritten as⁹

$$\rho(\mathbf{z}) - h(N_{\mathbf{z}}) + \bar{h} \in \arg \max_p p \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + \bar{h} \geq p] \quad \text{for all } \mathbf{z} \in \{0, 1\}^c. \quad (21)$$

For any policy ρ and any state $\mathbf{z} \in \{0, 1\}^c$, we define a policy $\tilde{\rho}_{\mathbf{z}}$ to be the review-offsetting policy with offset $a_{\mathbf{z}} = \rho(\mathbf{z}) - h(N_{\mathbf{z}})$. The next lemma establishes that the revenue of ρ is upper-bounded by the revenue of one of the policies $\tilde{\rho}_{\mathbf{z}}$ and characterizes that equality is achieved if and only if each of $\tilde{\rho}_{\mathbf{z}}$ have the same revenue.

⁹If the customer-specific distribution \mathcal{F} is strictly regular, there is a unique optimal dynamic pricing policy.

Lemma C.3. *The revenue of ρ is at most the highest revenue of $\tilde{\rho}_{\mathbf{z}}$ over $\mathbf{z} \in \{0, 1\}^c$, i.e.,*

$$\text{REV}(\sigma^{\text{NEWEST}}, \rho) \leq \max_{\mathbf{z} \in \{0, 1\}^c} \text{REV}(\sigma^{\text{NEWEST}}, \tilde{\rho}_{\mathbf{z}}).$$

Equality holds if and only if $\text{REV}(\sigma^{\text{NEWEST}}, \tilde{\rho}_{\mathbf{z}}) = \text{REV}(\sigma^{\text{NEWEST}}, \tilde{\rho}_{\mathbf{z}'})$ for all $\mathbf{z}, \mathbf{z}' \in \{0, 1\}^c$.

To characterize all optimal dynamic pricing policies, the next lemma shows that $\tilde{\rho}_{\mathbf{z}}$ is optimal if and only if $\rho(\mathbf{z}) - h(N_{\mathbf{z}}) + \bar{h}$ is a revenue maximizing price when selling to a customer with valuation $\Theta + \bar{h}$.

Lemma C.4. *The policy $\tilde{\rho}_{\mathbf{z}}$ is optimal if and only if $\rho(\mathbf{z}) - h(N_{\mathbf{z}}) + \bar{h}$ is a revenue maximizing price when selling to a customer with valuation $\Theta + \bar{h}$ i.e. $\rho(\mathbf{z}) - h(N_{\mathbf{z}}) + \bar{h} \in \arg \max_p p \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + \bar{h} \geq p]$.*

Proof of Proposition C.3. For an optimal policy ρ , Lemma C.3 implies that

$$\text{REV}(\sigma^{\text{NEWEST}}, \rho) \leq \max_{\mathbf{z} \in \{0, 1\}^c} \text{REV}(\sigma^{\text{NEWEST}}, \tilde{\rho}_{\mathbf{z}})$$

where equality holds if and only if $\text{REV}(\sigma^{\text{NEWEST}}, \tilde{\rho}_{\mathbf{z}}) = \text{REV}(\sigma^{\text{NEWEST}}, \tilde{\rho}_{\mathbf{z}'})$ for all $\mathbf{z}, \mathbf{z}' \in \{0, 1\}^c$. The optimality of ρ thus implies that $\tilde{\rho}_{\mathbf{z}}$ must be optimal for all $\mathbf{z} \in \{0, 1\}^c$. By Lemma C.4, $\rho(\mathbf{z}) - h(N_{\mathbf{z}}) + \bar{h} \in \arg \max_p p \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + \bar{h} \geq p]$ for all $\mathbf{z} \in \{0, 1\}^c$ i.e. (21) holds.

Conversely, suppose that ρ satisfies (21). Lemma C.4 implies that each of $\tilde{\rho}_{\mathbf{z}}$ is optimal and thus yields the same revenue, i.e., $\text{REV}(\sigma^{\text{NEWEST}}, \tilde{\rho}_{\mathbf{z}}) = r^*(\Theta + \bar{h})$ for all \mathbf{z} . By the equality condition of Lemma C.3, $\text{REV}(\sigma^{\text{NEWEST}}, \rho) = \max_{\mathbf{z} \in \{0, 1\}^c} \text{REV}(\sigma^{\text{NEWEST}}, \tilde{\rho}_{\mathbf{z}}) = r^*(\Theta + \bar{h})$ and thus ρ is optimal. \square

To prove Lemma C.3 and Lemma C.4, we show a simplified expression for the revenue of any review offsetting policy.

Lemma C.5. *Let $\bar{h} = \mathbb{E}_{N \sim \text{Binom}(c, \mu)}[h(N)]$. For a review-offsetting policy $\tilde{\rho}$ with offset $a \in \mathbb{R}$:*

$$\text{REV}(\sigma^{\text{NEWEST}}, \tilde{\rho}) = (a + \bar{h}) \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + \bar{h} \geq a + \bar{h}].$$

Proof of Lemma C.3. Given that policy $\tilde{\rho}_{\mathbf{z}}$ is review-offsetting with offset $a_{\mathbf{z}}$, it holds that

$$\text{REV}(\sigma^{\text{NEWEST}}, \tilde{\rho}_{\mathbf{z}}) = \underbrace{(\rho(\mathbf{z}) - h(N_{\mathbf{z}}) + \bar{h})}_{A_{\mathbf{z}}} \cdot \underbrace{\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N_{\mathbf{z}}) \geq \rho(\mathbf{z})]}_{(B_{\mathbf{z}})^{-1}}.$$

Let $\alpha_{\mathbf{z}} = \mu^{N_{\mathbf{z}}}(1 - \mu)^{c - N_{\mathbf{z}}}$ be the probability that c i.i.d. $\text{Bern}(\mu)$ trials result in \mathbf{z} . Hence $\mathbb{E}_{Y_1, \dots, Y_c \sim i.i.d. \text{Bern}(\mu)}[\rho(\mathbf{Y})] = \mathbb{E}_{Y_1, \dots, Y_c \sim i.i.d. \text{Bern}(\mu)}[\rho(\mathbf{Y}) - h(N_{\mathbf{Y}}) + \bar{h}] = \sum_{\mathbf{z} \in \{0, 1\}^c} \alpha_{\mathbf{z}} A_{\mathbf{z}}$. Applying Lemma 4.1 we can thus express the revenue of ρ as

$$\text{REV}(\sigma^{\text{NEWEST}}, \rho) = \frac{\sum_{\mathbf{z} \in \{0, 1\}^c} \alpha_{\mathbf{z}} A_{\mathbf{z}}}{\sum_{\mathbf{z} \in \{0, 1\}^c} \alpha_{\mathbf{z}} B_{\mathbf{z}}} \leq \max_{\mathbf{z} \in \mathcal{S}} \frac{A_{\mathbf{z}}}{B_{\mathbf{z}}} = \max_{\mathbf{z} \in \{0, 1\}^c} \text{REV}(\sigma^{\text{NEWEST}}, \tilde{\rho}_{\mathbf{z}}) \leq \text{REV}(\sigma^{\text{NEWEST}}, \rho^*)$$

where the first inequality follows from the following natural convexity property (Lemma C.2, Appendix C.5): for any $\{\tilde{\alpha}_i, \tilde{A}_i, \tilde{B}_i\}_{i \in \mathcal{S}}$ with $\tilde{\alpha}_i, \tilde{B}_i > 0$ for all i , $\frac{\sum_{i \in \mathcal{S}} \tilde{\alpha}_i \tilde{A}_i}{\sum_{i \in \mathcal{S}} \tilde{\alpha}_i \tilde{B}_i} \leq \max_{i \in \mathcal{S}} \frac{\tilde{A}_i}{\tilde{B}_i}$. By Lemma C.2, equality holds if and only if $\frac{A_{\mathbf{z}}}{B_{\mathbf{z}}} = \text{REV}(\sigma^{\text{NEWEST}}, \tilde{\rho}_{\mathbf{z}})$ is the same for all $\mathbf{z} \in \{0, 1\}^c$. \square

Proof of Lemma C.4. Given that $\tilde{\rho}_z$ is a review-offsetting policy with offset $\rho(z) - h(N_z)$, Lemma C.5 implies

$$\text{REV}(\sigma^{\text{NEWEST}}, \tilde{\rho}_z) = \underbrace{(\bar{h} + \rho(z) - h(N_z))}_{p_z} \cdot \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N_z) \geq \rho(z)].$$

By adding and subtracting \bar{h} from each side the purchase probability can be rewritten as

$$\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N_z) \geq \rho(z)] = \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + \bar{h} \geq \bar{h} + \rho(z) - h(N_z)] = \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + \bar{h} \geq p_z].$$

Thus, the revenue of $\tilde{\rho}_z$ is equal to the revenue of offering a price of p_z to a customer with valuation $\Theta + \bar{h}$, which is maximized if and only $p_z = \rho(z) - h(N_z) + \bar{h} \in \arg \max_p p \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + \bar{h} \geq p]$. \square

C.8 Uniqueness of optimal dynamic pricing for Newest and Random (Remark 4.3)

We show that the optimal dynamic pricing policies under σ^{NEWEST} and σ^{RANDOM} are unique assuming mild regularity conditions on the customer-specific value distribution \mathcal{F} . For a continuous random variable V with bounded support, let $\bar{F}_V(p) = \mathbb{P}[V \geq p]$ be its survival function. As V has a continuous distribution, the inverse survival function $G_V(q) = (\bar{F}_V)^{-1}(q)$ is well-defined for $q \in (0, 1)$, i.e., for any quantile $q \in (0, 1)$, $G_V(q)$ is the unique price which induces a purchase probability q . Let $r(q) := qG_V(q)$ be the revenue as a function of the quantile q . We require a notion of strict regularity.

Definition C.1 (Strict Regularity). *A random variable V has a strictly regular distribution if the revenue function $r(q) = qG_V(q)$ is strictly concave in the quantile $q \in (0, 1)$.*

In order to show uniqueness of the optimal dynamic pricing policies, we extend the strict regularity assumption by imposing a few mild further conditions.

Definition C.2 (Well-behavedness). *A random variable V is well-behaved if: (1) V is continuous with bounded support; (2) V is strictly regular; (3) $\mathbb{P}[V > 0] > 0$.*

Condition (3) implies that we can achieve a strictly positive revenue by selling to a customer with a well-behaved valuation V .

Lemma C.6. *For any $x \geq 0$ and well-behaved random variable V , $V_x := V + x$ is also well-behaved.*

Proof of Lemma C.6. To prove that V_x is well-behaved we establish all three conditions of Definition C.2. First, the continuity and boundedness of V imply that $V_x = V + x$ also has these properties. Second, $qG_V(q)$ is strictly concave and qx is linear in q ; thus $qG_{V+x}(q) = qG_V(q) + qx$ is strictly concave. Third, as $\mathbb{P}[V > 0]$, adding a non-negative scalar x yields $\mathbb{P}[V_x > 0]$. \square

Lemma C.7. *For a well-behaved random variable \tilde{V} , the revenue-maximizing price p^* is unique.*

Proof of Lemma C.7. Letting the support of \tilde{V} be $[\underline{v}, \bar{v}]$ implies that $G_{\tilde{V}}(q) \in (\underline{v}, \bar{v})$ for all $q \in (0, 1)$, $\lim_{q \rightarrow 0} G_{\tilde{V}}(q) = \bar{v}$, and $\lim_{q \rightarrow 1} G_{\tilde{V}}(q) = \underline{v}$. Thus,

$$\lim_{q \rightarrow 0} r(q) = \lim_{q \rightarrow 0} qG_{\tilde{V}}(q) = \lim_{q \rightarrow 0} G_{\tilde{V}}(q) \cdot \lim_{q \rightarrow 0} q = 0.$$

By the third condition of well-behavedness, $\mathbb{P}_{\tilde{V}}[\tilde{V} > 0]$ which implies that $\bar{v} > 0$ and thus there exists some quantile \tilde{q} such that $r(\tilde{q}) > 0$. In particular, for $\tilde{q} = \frac{\bar{v} - \max(\underline{v}, 0)}{2(\bar{v} - \underline{v})}$, $r(\tilde{q}) = \tilde{q}G_{\tilde{V}}(\tilde{q}) > 0$ as $G_{\tilde{V}}(\tilde{q}) = \frac{\bar{v} + \max(0, \underline{v})}{2} > 0$. Since $\lim_{q \rightarrow 0} r(q) = 0$, $r(\tilde{q}) > 0$, and $r(q)$ is strictly concave on $(0, 1)$, it holds that either (a) $r(q)$ has a unique maximizer at $q = q^* \in (0, 1)$ or (b) $r(q)$ is strictly increasing for $q \in (0, 1)$. We consider these two cases separately below.

For case (a), the optimal price p^* has a quantile $q^* = \mathbb{P}_{\tilde{V}}[\tilde{V} \geq p^*] \in (0, 1)$ such that $r(q^*) > r(q)$ for all $q \in (0, 1) \setminus \{q^*\}$. The strict concavity of r implies

$$r(q^*) > \max(\lim_{q \rightarrow 1} r(q), \lim_{q \rightarrow 0} r(q)) = \max(\underline{v}, 0). \quad (22)$$

We show that $p^* = G_{\tilde{V}}(q^*)$ is the unique revenue-maximizing price. It suffices to show that this price provides strictly higher revenue than any other price $p \neq p^*$, i.e., $r(q^*) = p^* \mathbb{P}_{\tilde{V}}[\tilde{V} \geq p^*] > p \mathbb{P}_{\tilde{V}}[\tilde{V} \geq p]$. Let $q = \mathbb{P}_{\tilde{V}}[\tilde{V} \geq p]$ be the quantile associated with the price p .

- If $p \in (\underline{v}, \bar{v})$ the continuity of \tilde{V} implies that $q = \mathbb{P}_{\tilde{V}}[\tilde{V} \geq p] \in (0, 1)$ and thus the revenue of p^* is strictly larger than the revenue of p by the assumption of case (a) that $r(q^*) > r(q)$ for all $q \in (0, 1) \setminus \{q^*\}$.
- If $p \geq \bar{v}$, then $p \mathbb{P}_{\tilde{V}}[\tilde{V} \geq p] = 0$ which combined with $r(q^*) > 0$ yields that the revenue of p^* is strictly larger than the revenue of p .
- Lastly, if $p \leq \underline{v}$, then $p \mathbb{P}_{\tilde{V}}[\tilde{V} \geq p] \leq \underline{v} < r(q^*)$ by (22).

For case (b), $r(q)$ is strictly increasing for $q \in (0, 1)$. Let $p^* = \underline{v}$. Since $r(1) = \bar{v}$, $r(0) = 0$, and $r(1) > r(0)$ we obtain that $\underline{v} > 0$. To show that $p^* = \underline{v}$ is the unique revenue-maximizing price, it suffices to prove that $\bar{v} = p^* \mathbb{P}_{\tilde{V}}[\tilde{V} \geq p^*] > p \mathbb{P}_{\tilde{V}}[\tilde{V} \geq p]$ for any $p \neq p^*$.

- If $p \in (\underline{v}, \bar{v})$ the continuity of \tilde{V} implies that $q = \mathbb{P}_{\tilde{V}}[\tilde{V} \geq p] \in (0, 1)$ and thus the revenue of p^* is strictly larger than the revenue of p by the assumption of case (b) that $r(q)$ is increasing on $(0, 1)$.
- If $p \geq \bar{v}$ then $\mathbb{P}_{\tilde{V}}[\tilde{V} \geq p] = 0$ and thus $p \mathbb{P}_{\tilde{V}}[\tilde{V} \geq p] = 0$ which is strictly smaller than $p^* = \underline{v}$.
- Lastly if $p < \underline{v}$, then $p \mathbb{P}_{\tilde{V}}[\tilde{V} \geq p] \leq p < \underline{v} = p^* = p^* \mathbb{P}_{\tilde{V}}[\tilde{V} \geq p^*]$.

As a result, in both cases, we established that there exists a unique revenue-maximizing p^* . \square

Proposition C.4. *For a well-behaved customer-specific valuation Θ , the optimal dynamic pricing policies under σ^{NEWEST} and σ^{RANDOM} are unique.*

Proof. The valuation of a customer at a state with review ratings \mathbf{z} is $\Theta + h(N_{\mathbf{z}})$. By Proposition 4.2, any optimal dynamic pricing policy under σ^{RANDOM} outputs, at each state of review ratings \mathbf{z} , a revenue-maximizing price for a customer with valuation $\Theta + h(N_{\mathbf{z}})$. Combining Lemma C.6 with $x = h(N_{\mathbf{z}})$ and Lemma C.7, this price is unique for every state of review ratings \mathbf{z} .

By Proposition C.3, any optimal dynamic pricing policy under σ^{NEWEST} outputs, at each state of review ratings \mathbf{z} , a price of $\rho^{\text{NEWEST}}(\mathbf{z}) = h(N_{\mathbf{z}}) + p_{\mathbf{z}} - \bar{h}$ where $p_{\mathbf{z}}$ is a revenue-maximizing price for a customer with valuation $\Theta + \bar{h}$. Combining Lemma C.6 with $x = \bar{h}$ and Lemma C.7, the price $\rho^{\text{NEWEST}}(\mathbf{z})$ is unique for every state of review ratings \mathbf{z} . \square

C.9 Characterization of optimal dynamic pricing policy under σ^{random} (Proposition 4.2)

Proof of Proposition 4.2. By definition, σ^{RANDOM} displays c i.i.d. $\text{Bern}(\mu)$ reviews at every round t . If the displayed reviews are $\mathbf{z} = (z_1, \dots, z_c) \in \{0, 1\}^c$, the revenue obtained by offering price $\rho(\mathbf{z})$ is equal to

$$\rho(\mathbf{z}) \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N_{\mathbf{z}}) \geq \rho(\mathbf{z})].$$

Thus the optimal price is any revenue-maximizing price $\rho^*(\mathbf{z}) \in \arg \max_{p \in \mathbb{R}} p \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(N_{\mathbf{z}}) \geq p]$. Since the number of positive reviews $N_{\mathbf{z}}$ is distributed as $\text{Binom}(c, \mu)$, the optimal revenue is equal to $\mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[\max_{p \in \mathbb{R}} p \mathbb{P}[\Theta + h(N) \geq p] \right] = \mathbb{E}_{N \sim \text{Binom}(c, \mu)} \left[r^*(\Theta + h(N)) \right]$. \square

C.10 Comparing optimal dynamic pricing for Newest and Random (Proposition 4.3)

We compare the optimal dynamic pricing policies under σ^{NEWEST} and σ^{RANDOM} assuming that the customer-specific valuation Θ is well-behaved (Definition C.2, Appendix C.8). This is a sufficient condition for the optimal dynamic pricing policies under σ^{NEWEST} and σ^{RANDOM} to be unique (Proposition C.4, Appendix C.8). We restate Proposition 4.3 (with the exact *mild regularity conditions*).

Proposition C.5. *Let $\bar{h} = \mathbb{E}_{N \sim \text{Binom}(c, \mu)}[h(N)]$. For any well-behaved customer-specific valuation Θ , the unique dynamic pricing policies ρ^{NEWEST} and ρ^{RANDOM} satisfy:*

- $\rho^{\text{NEWEST}}(\mathbf{z}) \geq \rho^{\text{RANDOM}}(\mathbf{z})$ for review states $\mathbf{z} \in \{0, 1\}^c$ with $h(N_{\mathbf{z}}) > \bar{h}$
- $\rho^{\text{NEWEST}}(\mathbf{z}) \leq \rho^{\text{RANDOM}}(\mathbf{z})$ for review states $\mathbf{z} \in \{0, 1\}^c$ with $h(N_{\mathbf{z}}) < \bar{h}$
- $\rho^{\text{NEWEST}}(\mathbf{z}) = \rho^{\text{RANDOM}}(\mathbf{z})$ review states $\mathbf{z} \in \{0, 1\}^c$ with $h(N_{\mathbf{z}}) = \bar{h}$.

To prove Proposition C.5, we introduce the revenue maximizing price of selling to a customer with valuation $\Theta + x$ for a non-negative scalar $x \geq 0$ i.e. $p^*(\Theta + x) = \arg \max_{p \in \mathbb{R}} p \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + x \geq p]$.

We also introduce the function $g(x) := p^*(\Theta + x) - x$ which intuitively captures the smallest idiosyncratic valuation a customer needs to have to purchase the product and is also a proxy for the purchase probability of a customer with valuation $\Theta + x$ (since $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + x \geq p] = \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta \geq p - x]$).

Lemma C.8. *For any well-behaved customer-specific value distribution Θ , $g(x)$ is weakly decreasing for $x \geq 0$.*

Proof of Proposition C.5. Proposition 4.2 implies that the optimal dynamic pricing policy under σ^{RANDOM} is $\rho^{\text{RANDOM}}(\mathbf{z}) = p^*(\Theta + h(N_{\mathbf{z}}))$ for $\mathbf{z} \in \{0, 1\}^c$. Theorem 4.3 implies that the optimal dynamic pricing policy under σ^{NEWEST} is $\rho^{\text{NEWEST}}(\mathbf{z}) = h(N_{\mathbf{z}}) + p^*(\Theta + \bar{h}) - \bar{h}$. As a result:

$$\rho^{\text{NEWEST}}(\mathbf{z}) - \rho^{\text{RANDOM}}(\mathbf{z}) = (p^*(\Theta + \bar{h}) - \bar{h}) - (p^*(\Theta + h(N_{\mathbf{z}})) - h(N_{\mathbf{z}})) = g(\bar{h}) - g(h(N_{\mathbf{z}})).$$

By Lemma C.8, g is monotonically decreasing which implies the result of the proposition. \square

To prove Lemma C.8, we use an auxiliary lemma which shows that for large enough x , the revenue-maximizing price $p^*(\Theta + x)$ is equal to $\underline{\theta} + x$ (and thus induces a purchase probability of one). Given that Θ is well-behaved, we let its support be $[\underline{\theta}, \bar{\theta}]$.

Lemma C.9. *For any well-behaved customer-specific value distribution Θ and $x \geq 0$, there exists some threshold $M \geq 0$ such that $p^*(\Theta + x) \in (\underline{\theta} + x, \bar{\theta} + x)$ for $x \in [0, M)$ and $p^*(\Theta + x) = \underline{\theta} + x$ for $x \geq M$.*

Proof of Lemma C.9. First observe that $p^*(\Theta + x) \geq \underline{\theta} + x$ since setting a price below $\underline{\theta} + x$ is never optimal. Indeed, prices $p < \underline{\theta} + x$ induce a purchase probability of one. Setting a price of $p + \epsilon$ for a small enough ϵ still induces a purchase probability of one and achieves a strictly higher revenue.

Second, observe that any price $p \geq \bar{\theta} + x$ induces a purchase probability of zero and thus zero revenue. As Θ is well-behaved, then so is $\Theta + x$ (Lemma C.6). This implies that $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + x > 0] > 0$ and thus the optimal revenue is strictly positive (as one can find a price $p > 0$ such that $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + x \geq p] > 0$). Thus, $p^*(\Theta + x) < \bar{\theta} + x$ for all $x \geq 0$.

Finally, let $S = \{x \geq 0 \mid p^*(\Theta + x) = \underline{\theta} + x\}$ be the set of increments $x \geq 0$ such that the revenue-maximizing price of selling to a customer with valuation $\Theta + x$ is equal to $\underline{\theta} + x$. Observe that setting a price of $\underline{\theta} + x$ induces a purchase probability of one and thus a revenue of $\underline{\theta} + x$. Thus, $x \in S$ if and only if the optimal revenue of selling to a customer with valuation $\Theta + x$ is $\underline{\theta} + x$. Given that $p^*(\Theta + x) \in [\underline{\theta} + x, \bar{\theta} + x)$, to conclude the proof it is sufficient to show that there exists some $M \geq 0$ such that $S = [M, +\infty)$. To prove this, it suffices to show that S satisfies two properties: (a) If $x \in S$ then $x' \in S$ for all $x' \geq x$; (b) If $\{x_i\}_{i=1}^\infty$ is a decreasing sequence with $x_i \in S$ and $\lim_{i \rightarrow \infty} x_i = x_\infty$, then $x_\infty \in S$.

For an increment x and price p , let $R(x, p) = p\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + x \geq p]$ be the revenue of offering a price p to a customer with valuation $\Theta + x$. For property (a), let $x \in S$ and $x' \geq x$. To show that $x' \in S$, it is sufficient to show $R(x', p) \leq \underline{\theta} + x'$ for all prices p . Expanding we have

$$\begin{aligned} R(x', p) &= p\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + x' \geq p] = (p - (x' - x) + (x' - x))\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + x + (x' - x) \geq p] \\ &= (p - (x' - x))\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + x \geq p - (x' - x)] + (x' - x)\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + x' \geq p] \\ &= R(x, p - (x' - x)) + (x' - x)\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + x' \geq p] \\ &\leq x + \underline{\theta} + (x' - x) = \underline{\theta} + x' \Rightarrow x' \in S \end{aligned}$$

where the last inequality uses $R(x, p - (x' - x)) \leq \underline{\theta} + x$ since $x \in S$ and $\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + x' \geq p] \leq 1$.

For property (b), let $\{x_i\}_{i=1}^\infty$ be a decreasing sequence with $x_i \in S$ and such that $\lim_{i \rightarrow \infty} x_i = x_\infty$. We show that $x_\infty \in S$. The continuity of Θ implies that revenue function $R(x, p)$ is continuous in x for any fixed price p . Combining this with the fact that $R(x_i, p) \leq \underline{\theta} + x_i$ for all i , we obtain

$$R(x_\infty, p) = \lim_{i \rightarrow \infty} R(x_i, p) \leq \lim_{i \rightarrow \infty} \underline{\theta} + x_i = \underline{\theta} + x_\infty \Rightarrow x_\infty \in S.$$

□

Proof of Lemma C.8. Lemma C.9 shows that $p^*(\Theta + x) \in (\underline{\theta} + x, \bar{\theta} + x)$ for $x \in [0, M)$ and $p^*(\Theta + x) = \underline{\theta} + x$ for $x \geq M$ for some threshold $M \geq 0$. Thus, $g(x) \in (\underline{\theta}, \bar{\theta})$ for $x \in [0, M)$ and $g(x) = \underline{\theta}$ for $x \in [M, +\infty)$. It thus suffices to show that $g(x)$ is monotonically decreasing for $x \in [0, M)$. The continuity of Θ implies that the function $p\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + x \geq p]$ is differentiable for $p \in (\underline{\theta} + x, \bar{\theta} + x)$ and thus the first-order conditions must be satisfied at $p = p^*(\Theta + x) \in (\underline{\theta} + x, \bar{\theta} + x)$. Denoting the survival and density functions of Θ by \bar{F} and f respectively, this yields

$$\frac{d}{dp} p\mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + x \geq p] = \frac{d}{dp} p\bar{F}(p - x) = \bar{F}(p - x) - pf(p - x) = 0$$

for $p = p^*(\Theta + x)$. Rearranging the last equation yields

$$p^*(\Theta + x) - x - \frac{\bar{F}(p^*(\Theta + x) - x)}{f(p^*(\Theta + x) - x)} = -x \iff g(x) - \frac{\bar{F}(g(x))}{f(g(x))} = -x \quad (23)$$

for any $x \in [0, M)$. The strict regularity of Θ implies that the function $u \rightarrow u - \frac{\bar{F}(u)}{f(u)}$ is strictly increasing. Thus, the left-hand side of (23) is increasing in $u = g(x)$ while the right-hand side is strictly decreasing in x . Hence, $u = g(x)$ is decreasing for $x \in [0, M)$ concluding the proof. \square

C.11 Platforms unaware of state-dependent behavior (Corollary 4.2)

Proof of Corollary 4.2. Suppose that \mathcal{F} has support on $[0, \bar{\theta}]$. Then Theorem 4.1 gives

$$\text{REV}(\sigma^{\text{RANDOM}}, \Pi^{\text{STATIC}}) \geq \frac{\mu^c h(c)}{h(0) + \bar{\theta}} \text{REV}(\sigma^{\text{NEWEST}}, \Pi^{\text{STATIC}}). \quad (24)$$

Since \mathcal{F} is non-negative, Theorem 4.2 gives

$$\text{REV}(\sigma^{\text{NEWEST}}, \Pi^{\text{DYNAMIC}}) \geq \frac{1}{2} \text{REV}(\sigma^{\text{RANDOM}}, \Pi^{\text{DYNAMIC}}). \quad (25)$$

Combining (24) and (25) with the fact that $\text{REV}(\sigma^{\text{RANDOM}}, \Pi^{\text{DYNAMIC}}) \geq \text{REV}(\sigma^{\text{RANDOM}}, \Pi^{\text{STATIC}})$,

$$\frac{\text{REV}(\sigma^{\text{NEWEST}}, \Pi^{\text{DYNAMIC}})}{\text{REV}(\sigma^{\text{NEWEST}}, \Pi^{\text{STATIC}})} \geq \frac{\mu^c h(c)}{2(h(0) + \bar{\theta})}.$$

For any $M > 0$ there exist $\epsilon(M) > 0$ such that when $\bar{\theta}, h(0) < \epsilon(M)$, $\frac{\text{REV}(\sigma^{\text{NEWEST}}, \Pi^{\text{DYNAMIC}})}{\text{REV}(\sigma^{\text{NEWEST}}, \Pi^{\text{STATIC}})} > M$. \square

D Supplementary material for Section 5

D.1 Higher revenue of σ^{random} than σ^{newest} even with non-stationarity (Proposition 5.1)

Our proof relies on analyzing the Markov Chain $M_t = (r_t, \mu^{(t)})$ where r_t is the newest review the customer sees at round t and $\mu^{(t)}$ is the product quality at round t , under the σ^{NEWEST} review ordering policy. Let $\pi^{\text{NEWEST}}(r, \mu; \xi)$ be the stationary distribution of the process M_t at a review $r \in \{0, 1\}$ and product quality $\mu \in \{\mu_L, \mu_H\}$ when the probability of change is ξ . We use the following lemma which shows that σ^{NEWEST} spends strictly more time at a negative review than σ^{RANDOM} .

Lemma D.1. *For any product qualities μ_L and μ_H satisfying $\mu_L < \mu_H$, probability of change $\xi \in (0, 1]$, and price p satisfying Assumption 3.1, σ^{NEWEST} spends more time at a negative review than σ^{RANDOM} , i.e.,*

$$\pi^{\text{NEWEST}}(0, \mu_L; \xi) + \pi^{\text{NEWEST}}(0, \mu_H; \xi) > \pi^{\text{RANDOM}}(0).$$

Proof of Proposition 5.1. The revenue of σ^{NEWEST} is given by

$$\begin{aligned} \text{REV}(\sigma^{\text{NEWEST}}, p) &= p \cdot [q_0(\pi^{\text{NEWEST}}(0, \mu_L; \xi) + \pi^{\text{NEWEST}}(0, \mu_H; \xi)) \\ &\quad + q_1(\pi^{\text{NEWEST}}(1, \mu_L; \xi) + \pi^{\text{NEWEST}}(1, \mu_H; \xi))]. \end{aligned}$$

Letting $\pi^{\text{RANDOM}}(0) = 1 - \frac{\mu_L + \mu_H}{2}$ and $\pi^{\text{RANDOM}}(1) = \frac{\mu_L + \mu_H}{2}$ be the probabilities of a positive and a negative review under σ^{RANDOM} respectively, the revenue of σ^{RANDOM} is given by

$$\text{REV}(\sigma^{\text{RANDOM}}, p) = p \cdot [q_0 \pi^{\text{RANDOM}}(0) + q_1 \pi^{\text{RANDOM}}(1)].$$

Lemma D.1 yields that σ^{NEWEST} spends more time at a negative review than σ^{RANDOM} , i.e.,

$$\pi^{\text{NEWEST}}(0, \mu_L; \xi) + \pi^{\text{NEWEST}}(0, \mu_H; \xi) > \pi^{\text{RANDOM}}(0).$$

Combined with the fact that $q_1 > q_0$ (as p is non-degenerate) the proof follows. \square

To prove Lemma D.1, we first prove two auxiliary lemmas about the probability of a negative review under σ^{NEWEST} for a probability of change ξ by $\pi^{\text{NEWEST}}(0; \xi)$, i.e.,

$$\pi^{\text{NEWEST}}(0; \xi) := \pi^{\text{NEWEST}}(0, \mu_L; \xi) + \pi^{\text{NEWEST}}(0, \mu_H; \xi).$$

Lemma D.2 shows that $\pi^{\text{NEWEST}}(0; \xi)$ is non-decreasing in the switching probability ξ . Lemma D.3 shows that as the switching probability ξ approaches 0, σ^{NEWEST} spends more time at a negative review than σ^{RANDOM} .

Lemma D.2. *For any product qualities μ_L and μ_H satisfying $\mu_L < \mu_H$, and any price p satisfying Assumption 3.1, the probability of a negative review $\pi^{\text{NEWEST}}(0; \xi)$ is strictly increasing in the probability of change ξ .*

Lemma D.3. *For any product qualities μ_L and μ_H satisfying $\mu_L < \mu_H$, and any price p satisfying Assumption 3.1, σ^{NEWEST} spends strictly more time at a negative review than σ^{RANDOM} as the probability of change ξ converges to zero, i. e., $\lim_{\xi \rightarrow 0} \pi^{\text{NEWEST}}(0; \xi) > \pi^{\text{RANDOM}}(0)$.*

Proof of Lemma D.1. The proof follows directly from Lemma D.2 and Lemma D.3. \square

To prove Lemma D.2 and Lemma D.3, we characterize the stationary distribution of the Markov Chain $M_t = (r_t, \mu^{(t)})$ consisting of the newest review r_t and the current product quality $\mu^{(t)}$ at round t under the review ordering policy σ^{NEWEST} . To state the characterization it is useful to define the quantities $A_H = 1 - ((1 - \mu_H)q_1 + \mu_H q_0)$ and $A_L = 1 - ((1 - \mu_L)q_1 + \mu_L q_0)$.

Lemma D.4. *For any price p satisfying Assumption 3.1, the stationary distribution of the Markov Chain M_t at the states with a negative review is given by*

$$\begin{aligned} \pi^{\text{NEWEST}}(0, \mu_H; \xi) &= \frac{q_1 [2(1 - \mu_H)A_L(\xi - 1) + (1 - \mu_L)\xi + (1 - \mu_H)(2 - \xi)]}{(2 - (2 - \xi)A_H)(2 - (2 - \xi)A_L) - \xi^2 A_H A_L} \quad \text{and} \\ \pi^{\text{NEWEST}}(0, \mu_L; \xi) &= \frac{q_1 [2(1 - \mu_L)A_H(\xi - 1) + (1 - \mu_H)\xi + (1 - \mu_L)(2 - \xi)]}{(2 - (2 - \xi)A_H)(2 - (2 - \xi)A_L) - \xi^2 A_H A_L} \end{aligned}$$

Proof of Lemma D.2. Using Lemma D.4, we can rewrite

$$\frac{\pi^{\text{NEWEST}}(0; \xi)}{q_1} = \frac{((1 - \mu_L)A_H + (1 - \mu_H)A_L)\xi + (1 - \mu_L)(1 - A_H) + (1 - \mu_H)(1 - A_L)}{((1 - A_H)A_L + (1 - A_L)A_H)\xi + 2(1 - A_H)(1 - A_L)}$$

We will use the fact that a rational function $\xi \rightarrow \frac{A\xi+B}{C\xi+D}$ is strictly increasing if and only if $AD > BC$. To conclude the proof it suffices to show that this fact holds true for $A = (1 - \mu_L)A_H + (1 - \mu_H)A_L$,

$B = (1 - \mu_L)(1 - A_H) + (1 - \mu_H)(1 - A_L)$, $C = (1 - A_H)A_L + (1 - A_L)A_H$, and $D = 2(1 - A_H)(1 - A_L)$. Expanding the product AD yields

$$\begin{aligned} AD &= 2((1 - \mu_L)A_H + (1 - \mu_H)A_L)(1 - A_H)(1 - A_L) \\ &= 2(1 - \mu_L)(1 - A_H)(1 - A_L)A_H + 2(1 - \mu_H)(1 - A_L)(1 - A_H)A_L. \end{aligned} \quad (26)$$

Expanding the product BC yields

$$\begin{aligned} BC &= ((1 - \mu_L)(1 - A_H) + (1 - \mu_H)(1 - A_L))((1 - A_H)A_L + (1 - A_L)A_H) \\ &= (1 - \mu_L)(1 - A_H)^2 A_L + (1 - \mu_H)(1 - A_L)^2 A_H \\ &\quad + (1 - \mu_L)(1 - A_H)(1 - A_L)A_H + (1 - \mu_H)(1 - A_L)(1 - A_H)A_L. \end{aligned} \quad (27)$$

Therefore, canceling the terms $(1 - \mu_L)(1 - A_H)(1 - A_L)A_H$ and $(1 - \mu_H)(1 - A_L)(1 - A_H)A_L$, the inequality $AD > BC$ is equivalent to showing

$$\begin{aligned} &(1 - \mu_L)(1 - A_H)(1 - A_L)A_H + (1 - \mu_H)(1 - A_L)(1 - A_H)A_L \\ &> (1 - \mu_L)(1 - A_H)^2 A_L + (1 - \mu_H)(1 - A_L)^2 A_H \\ \iff &(1 - \mu_L)(1 - A_H)((1 - A_L)A_H - (1 - A_H)A_L) \\ &\quad + (1 - \mu_H)(1 - A_L)((1 - A_H)A_L - (1 - A_L)A_H) > 0 \\ \iff &(1 - \mu_L)(1 - A_H)(A_H - A_L) + (1 - \mu_H)(1 - A_L)(A_L - A_H) > 0 \text{ (dividing by } A_H - A_L > 0) \\ \iff &(1 - \mu_L)(1 - A_H) > (1 - \mu_H)(1 - A_L) \text{ (substituting the expressions for } A_H \text{ and } A_L) \\ \stackrel{(1)}{\iff} & \\ \iff &(1 - \mu_L)((1 - \mu_H)q_1 + \mu_H q_0) > (1 - \mu_H)((1 - \mu_L)q_1 + \mu_L q_0) \\ \iff &(1 - \mu_L)\mu_H q_0 > (1 - \mu_H)\mu_L q_0 \\ \stackrel{(2)}{\iff} &\mu_H > \mu_L \end{aligned}$$

where in (1) we divided by $A_H - A_L = (\mu_H - \mu_L)(q_1 - q_0) > 0$ since $q_1 > q_0$ (as p satisfies Assumption 3.1) and $\mu_H > \mu_L$ and in (2) we divided by $q_0 > 0$ as p satisfies Assumption 3.1. The last inequality holds as $\mu_H > \mu_L$ finishing the proof. \square

Proof of Lemma D.3. Using Lemma D.4,

$$\pi^{\text{NEWEST}}(0; \xi) = \frac{q_1 \left[((1 - \mu_L)A_H + (1 - \mu_H)A_L)\xi + (1 - \mu_L)(1 - A_H) + (1 - \mu_H)(1 - A_L) \right]}{((1 - A_H)A_L + (1 - A_L)A_H)\xi + 2(1 - A_H)(1 - A_L)}$$

Taking the limit $\xi \rightarrow 0$ yields

$$\lim_{\xi \rightarrow 0} \pi^{\text{NEWEST}}(0; \xi) = \frac{\frac{1}{2}q_1 [(1 - \mu_L)(1 - A_H) + (1 - \mu_H)(1 - A_H)]}{(1 - A_H)(1 - A_L)} = \underbrace{\frac{1}{2} \frac{q_1(1 - \mu_L)}{1 - A_L}}_{\text{TERM}_L} + \underbrace{\frac{1}{2} \frac{q_1(1 - \mu_H)}{1 - A_H}}_{\text{TERM}_H}.$$

By substituting the expression for $A_L = 1 - q_1(1 - \mu_L) - q_0\mu_L$, and using that $q_1 > q_0$ (as p is non-degenerate as it satisfies Assumption 3.1), we can lower bound TERM_L as

$$\text{TERM}_L = \frac{1}{2} \frac{q_1(1 - \mu_L)}{(1 - \mu_L)q_1 + \mu_L q_0} > \frac{1}{2} \frac{q_1(1 - \mu_L)}{q_1} = \frac{1 - \mu_L}{2}. \quad (28)$$

By substituting the expression for $A_H = 1 - q_1(1 - \mu_H) - q_0\mu_H$, and using that $q_1 > q_0$ (as p is non-degenerate as it satisfies Assumption 3.1), we can lower bound TERM_H as

$$\text{TERM}_H = \frac{1}{2} \frac{q_1(1 - \mu_H)}{(1 - \mu_H)q_1 + \mu_H q_0} > \frac{1}{2} \frac{q_1(1 - \mu_H)}{q_1} = \frac{1 - \mu_H}{2}. \quad (29)$$

Combining (28) and (29) with the fact that $\pi^{\text{RANDOM}}(0) = \frac{1 - \mu_L}{2} + \frac{1 - \mu_H}{2}$ finishes the proof. \square

To prove Lemma D.4, we first establish that M_t is irreducible and aperiodic, which are conditions which guarantee the existence of a unique stationary distribution.

Lemma D.5. *The Markov Chain M_t is irreducible and aperiodic.*

Proof of Lemma D.5. We first show that the Markov Chain M_t is irreducible and aperiodic. For states $M^{(1)} = (r_1, \mu_1), M^{(2)} = (r_2, \mu_2) \in \{0, 1\} \times \{\mu_L, \mu_H\}$, M_t can reach state $M^{(2)}$ from $M^{(1)}$ after a state transition of the product quality (with probability ξ) and a new state of μ_1 (with probability $\frac{1}{2}$), a purchase (with probability at least $\min(q_1, q_0) = q_0$), and a review r_2 from μ_1 (with probability $\mu_1^{r_2}(1 - \mu_1)^{1 - r_2}$). Therefore, the one-step transition probability from $M^{(1)}$ to $M^{(2)}$ is at least

$$\mathbb{P}[M_t = M^{(2)} | M_{t-1} = M^{(1)}] \geq \frac{\xi}{2} q_0 \mu_1^{r_2} (1 - \mu_1)^{1 - r_2}. \quad (30)$$

Therefore, using that $\mu_L < 1$ and $\mu_H > 0$ (as $\mu_L < \mu_H$) and applying this formula when $(M^{(1)}, M^{(2)}) \in \left\{ ((0, \mu_L), (0, \mu_H)), ((0, \mu_H), (1, \mu_H)), ((1, \mu_H), (1, \mu_L)), ((1, \mu_L), (0, \mu_L)) \right\}$,

$$\begin{aligned} \mathbb{P}[M_t = (0, \mu_H) | M_{t-1} = (0, \mu_L)] &\geq \frac{\xi}{2} q_0 (1 - \mu_L) > 0, \\ \mathbb{P}[M_t = (1, \mu_H) | M_{t-1} = (0, \mu_H)] &\geq \frac{\xi}{2} q_0 \mu_H > 0, \\ \mathbb{P}[M_t = (1, \mu_L) | M_{t-1} = (1, \mu_H)] &\geq \frac{\xi}{2} q_0 \mu_H > 0, \quad \text{and} \\ \mathbb{P}[M_t = (0, \mu_L) | M_{t-1} = (1, \mu_L)] &\geq \frac{\xi}{2} q_0 (1 - \mu_L) > 0. \end{aligned}$$

Thus, any state of M_t can be reached from any other state, i.e., M_t is irreducible.

We next show that M_t is aperiodic. Applying (30) to $M^{(1)} = M^{(2)} = (0, \mu_L)$ we obtain $\mathbb{P}[M_t = (0, \mu_L) | M_{t-1} = (0, \mu_L)] \geq \frac{\xi}{2} q_0 (1 - \mu_L) > 0$. Thus, M_t has a self-loop, i.e., it is aperiodic. \square

Proof of Lemma D.4. By Lemma D.5, M_t is aperiodic and irreducible. Thus, M_t has a unique stationary distribution $\pi(\cdot, \cdot; \xi)$. We next show that $\pi(0, \mu_L; \xi)$ and $\pi(0, \mu_H; \xi)$ are given by the expressions in the statement. Given that the product quality $\mu^{(t)}$ spends an equal fraction of rounds in each of μ_L and μ_H in steady-state, then

$$\pi(0, \mu_L; \xi) + \pi(1, \mu_L; \xi) = \pi(0, \mu_H; \xi) + \pi(1, \mu_H; \xi) = \frac{1}{2}. \quad (31)$$

We first write down the steady-state equations of π for the state $(0, \mu_L)$. This state can be reached from the state $(0, \mu_L)$ if the product quality stays the same (with probability $(1 - \frac{\xi}{2})$) and there is either a purchase and a negative review (with probability $q_0(1 - \mu_L)$) or a non-purchase (with probability $1 - q_0$). Thus, the state $(0, \mu_L)$ can be reached from the state $(0, \mu_L)$ with probability of $(1 - \frac{\xi}{2})(q_0(1 - \mu_L) + 1 - q_0)$. The state $(0, \mu_L)$ can be reached from the state $(0, \mu_H)$ if the quality

changes (with probability $\frac{\xi}{2}$) and there is either a purchase and a negative review (with probability $q_0(1 - \mu_H)$) or a non-purchase (with probability $1 - q_0$). Thus, the state $(0, \mu_L)$ can be reached from the state $(0, \mu_H)$ with probability $\frac{\xi}{2}(q_0(1 - \mu_H) + 1 - q_0)$. The state $(0, \mu_L)$ can be reached from the state $(1, \mu_L)$ if the quality stays the same and there is a purchase with a negative review (with probability $(1 - \frac{\xi}{2})q_1(1 - \mu_L)$). The state $(0, \mu_L)$ can be reached from the state $(1, \mu_H)$ if the quality stays the same and there is a purchase with negative review (with probability $\frac{\xi}{2}q_1(1 - \mu_H)$). Therefore, the steady-state equation at the state $\pi(0, \mu_L)$ is given by

$$\begin{aligned}\pi(0, \mu_L; \xi) &= (1 - \frac{\xi}{2})(q_0(1 - \mu_L) + 1 - q_0)\pi(0, \mu_L; \xi) + \frac{\xi}{2}(q_0(1 - \mu_H) + 1 - q_0)\pi(0, \mu_H; \xi) \\ &\quad + (1 - \frac{\xi}{2})q_1(1 - \mu_L)\pi(1, \mu_L; \xi) + \frac{\xi}{2}q_1(1 - \mu_H)\pi(1, \mu_H; \xi).\end{aligned}$$

Using (31) the expressions for A_L and A_H , the above is rewritten as

$$\pi(0, \mu_L; \xi) = \pi(0, \mu_L; \xi)(1 - \frac{\xi}{2})A_L + \pi(0, \mu_H; \xi)\frac{\xi}{2}A_H + \frac{1}{2}q_1((1 - \frac{\xi}{2})(1 - \mu_L) + \frac{\xi}{2}(1 - \mu_H)).$$

Rearranging terms, we obtain

$$\pi(0, \mu_L; \xi) = \frac{\pi(0, \mu_H; \xi)\xi A_H + \frac{1}{2}q_1((2 - \xi)(1 - \mu_L) + \xi(1 - \mu_H))}{2 - (2 - \xi)A_L}. \quad (32)$$

Due to an analogous argument, the steady-state equation at the state $(0, \mu_H)$ is

$$\pi(0, \mu_H; \xi) = \pi(0, \mu_L; \xi)\frac{\xi}{2}A_L + \pi(0, \mu_H; \xi)(1 - \frac{\xi}{2})A_H + \frac{1}{2}q_1(\frac{\xi}{2}(1 - \mu_L) + (1 - \frac{\xi}{2})(1 - \mu_H)).$$

which can similarly be rewritten as

$$\pi(0, \mu_H; \xi) = \frac{\pi(0, \mu_L; \xi)\xi A_L + \frac{1}{2}q_1((2 - \xi)(1 - \mu_H) + \xi(1 - \mu_L))}{2 - (2 - \xi)A_H}. \quad (33)$$

Substituting (32) into (33) yields that

$$\begin{aligned}\pi(0, \mu_H; \xi) &= \frac{\frac{\pi(0, \mu_H; \xi)\xi A_H + \frac{1}{2}q_1((2 - \xi)(1 - \mu_L) + \xi(1 - \mu_H))}{2 - (2 - \xi)A_L}\xi A_L + \frac{1}{2}q_1((2 - \xi)(1 - \mu_H) + \xi(1 - \mu_L))}{2 - (2 - \xi)A_H} \\ &= \frac{\pi(0, \mu_H; \xi)\xi^2 A_L A_H + \frac{1}{2}q_1 X}{(2 - (2 - \xi)A_H)(2 - (2 - \xi)A_L)}\end{aligned} \quad (34)$$

where

$$\begin{aligned}X &= \xi A_L((2 - \xi)(1 - \mu_L) + \xi(1 - \mu_H)) + (2 - (2 - \xi)A_L)((2 - \xi)(1 - \mu_H) + \xi(1 - \mu_L)) \\ &= A_L(\xi(2 - \xi)(1 - \mu_L) + \xi^2(1 - \mu_H) - (2 - \xi)^2(1 - \mu_H) - \xi(2 - \xi)(1 - \mu_L)) \\ &\quad + 2((2 - \xi)(1 - \mu_H) + \xi(1 - \mu_L)) \\ &= A_L(1 - \mu_H)(\xi^2 - (2 - \xi)^2) + 2((2 - \xi)(1 - \mu_H) + \xi(1 - \mu_L)) \\ &= 4(1 - \mu_H)A_L(\xi - 1) + 2((2 - \xi)(1 - \mu_H) + \xi(1 - \mu_L))\end{aligned}$$

Substituting the expression for X in (34) yields

$$\begin{aligned}\pi(0, \mu_H; \xi) &= \frac{\pi(0, \mu_H; \xi) \xi^2 A_L A_H + q_1 [2(1 - \mu_H) A_L (\xi - 1) + (2 - \xi)(1 - \mu_H) + \xi(1 - \mu_L)]}{(2 - (2 - \xi) A_H)(2 - (2 - \xi) A_L)} \\ \iff \pi(0, \mu_H; \xi) &= \frac{q_1 [2(1 - \mu_H) A_L (\xi - 1) + (1 - \mu_L) \xi + (1 - \mu_H)(2 - \xi)]}{(2 - (2 - \xi) A_H)(2 - (2 - \xi) A_L) - \xi^2 A_H A_L}\end{aligned}$$

which is exactly the expression for $\pi^{\text{NEWEST}}(0, \mu_H; \xi)$ given in the theorem statement. Due an analogous argument substituting (33) into (32) and solving for $\pi(0, \mu_L; \xi)$ yields

$$\pi^{\text{NEWEST}}(0, \mu_L; \xi) = \frac{q_1 [2(1 - \mu_L) A_H (\xi - 1) + (1 - \mu_H) \xi + (1 - \mu_L)(2 - \xi)]}{(2 - (2 - \xi) A_H)(2 - (2 - \xi) A_L) - \xi^2 A_H A_L}$$

concluding the proof. \square

D.2 Belief error is always smaller for σ^{newest} under non-stationarity (Proposition 5.2)

Our proof relies on analyzing the Markov Chain $M_t = (r_t, \mu^{(t)})$ where r_t is the newest review the customer sees at round t and $\mu^{(t)}$ is the product quality at round t , under the σ^{NEWEST} review ordering policy. Let $\pi^{\text{NEWEST}}(r, \mu; \xi)$ be the stationary distribution of the process M_t at a review $r \in \{0, 1\}$ and product quality $\mu \in \{\mu_L, \mu_H\}$ when the probability of change is ξ . To prove the proposition we will use the characterization the stationary distribution of process M_t from Lemma D.4.

Proof of Proposition 5.2. For a review ordering policy σ and probability of change ξ , let $\pi^\sigma(r, \mu; \xi)$ be the fraction of time this policy spends at a state where the current review is $r \in \{0, 1\}$ and the current product quality is $\mu \in \{\mu_L, \mu_H\}$. The belief error $\text{BELIEFERROR}(\sigma, p; \xi)$ is given by:

$$\begin{aligned} & (\mathbb{E}[\text{Beta}(a, b + 1)] - \mu_L)^2 \pi^\sigma(0, \mu_L; \xi) + (\mathbb{E}[\text{Beta}(a + 1, b)] - \mu_L)^2 \pi^\sigma(1, \mu_L; \xi) \\ & + (\mathbb{E}[\text{Beta}(a, b + 1)] - \mu_H)^2 \pi^\sigma(0, \mu_H; \xi) + (\mathbb{E}[\text{Beta}(a + 1, b)] - \mu_H)^2 \pi^\sigma(1, \mu_H; \xi). \end{aligned}$$

Given that $\mu_L = \mathbb{E}[\text{Beta}(a, b + 1)]$ and $\mu_H = \mathbb{E}[\text{Beta}(a + 1, b)]$, the belief error can be rewritten as $\text{BELIEFERROR}(\sigma, p; \xi) = (\mu_H - \mu_L)^2 (\pi^\sigma(1, \mu_L; \xi) + \pi^\sigma(0, \mu_H; \xi))$. Given that the product quality $\mu^{(t)}$ spends an equal amount of time each of the states μ_L and μ_H , $\pi^\sigma(1, \mu_L; \xi) = \frac{1}{2} - \pi^\sigma(0, \mu_L; \xi)$, the belief error can be further rewritten as

$$\begin{aligned} \text{BELIEFERROR}(\sigma, p; \xi) &= (\mu_H - \mu_L)^2 \left(\frac{1}{2} - \pi^\sigma(0, \mu_L; \xi) + \pi^\sigma(0, \mu_H; \xi) \right) \\ &= (\mu_H - \mu_L)^2 (\pi^\sigma(0, \mu_H; \xi) - \pi^\sigma(0, \mu_L; \xi)) + \frac{(\mu_H - \mu_L)^2}{2}. \end{aligned} \quad (35)$$

Taking $\sigma = \sigma^{\text{RANDOM}}$, $\pi^\sigma(0, \mu_H; \xi) = \pi^\sigma(0, \mu_L; \xi) = \frac{1}{2}(1 - \frac{\mu_H + \mu_L}{2})$, and thus (35) implies that the belief error of σ^{RANDOM} is $\text{BELIEFERROR}(\sigma^{\text{RANDOM}}, p; \xi) = \frac{(\mu_H - \mu_L)^2}{2}$.

Taking $\sigma = \sigma^{\text{NEWEST}}$ and using Lemma D.4 and the expressions for A_L and A_H , the difference

$\pi^{\text{NEWEST}}(0, \mu_H; \xi) - \pi^{\text{NEWEST}}(0, \mu_L; \xi)$ equals

$$\begin{aligned}
& \frac{q_1[2(1 - \mu_H)A_L(\xi - 1) + (1 - \mu_L)\xi + (1 - \mu_H)(2 - \xi)]}{(2 - (2 - \xi)A_H)(2 - (2 - \xi)A_L) - \xi^2 A_H A_L} \\
& - \frac{q_1[2(1 - \mu_L)A_H(\xi - 1) + (1 - \mu_H)\xi + (1 - \mu_L)(2 - \xi)]}{(2 - (2 - \xi)A_H)(2 - (2 - \xi)A_L) - \xi^2 A_H A_L} \\
& = \frac{2q_1(\xi - 1)((1 - \mu_L)(1 - A_H) - (1 - \mu_H)(1 - A_L))}{(2 - (2 - \xi)A_H)(2 - (2 - \xi)A_L) - \xi^2 A_H A_L} \\
& = \frac{2q_1(\xi - 1)[(1 - \mu_L)((1 - \mu_H)q_1 + \mu_H q_0) - (1 - \mu_H)((1 - \mu_L)q_1 + \mu_L q_0)]}{(2 - (2 - \xi)A_H)(2 - (2 - \xi)A_L) - \xi^2 A_H A_L} \\
& = \frac{2q_1 q_0(\xi - 1)(\mu_H - \mu_L)}{(2 - (2 - \xi)A_H)(2 - (2 - \xi)A_L) - \xi^2 A_H A_L}.
\end{aligned}$$

Therefore (35) combined with the fact that $\text{BELIEFERROR}(\sigma^{\text{RANDOM}}, p; \xi) = \frac{(\mu_H - \mu_L)^2}{2}$ implies

$$\begin{aligned}
\text{BELIEFERROR}(\sigma^{\text{NEWEST}}, p; \xi) &= (\mu_H - \mu_L)^2 \frac{2q_1 q_0(\xi - 1)(\mu_H - \mu_L)}{(2 - (2 - \xi)A_H)(2 - (2 - \xi)A_L) - \xi^2 A_H A_L} \\
&+ \text{BELIEFERROR}(\sigma^{\text{RANDOM}}, p; \xi) \\
&= 2q_1 q_0(\mu_H - \mu_L)^3 g(\xi) + \text{BELIEFERROR}(\sigma^{\text{RANDOM}}, p; \xi)
\end{aligned}$$

where $g(\xi) = \frac{(\xi - 1)}{(2 - (2 - \xi)A_H)(2 - (2 - \xi)A_L) - \xi^2 A_H A_L} = \frac{\xi - 1}{2((1 - A_L)A_H + (1 - A_H)A_L)\xi + 4(1 - A_H)(1 - A_L)}$. Given that for any $\xi \in (0, 1)$, the term $g(\xi)$ is strictly negative and $\mu_H - \mu_L > 0$, $\text{BELIEFERROR}(\sigma^{\text{NEWEST}}, p; \xi) < \text{BELIEFERROR}(\sigma^{\text{RANDOM}}, p; \xi)$ showing the main claim of the proposition.

To show the second claim, it suffices to show that $g(\xi)$ is strictly increasing in ξ . A function of the form $\xi \rightarrow \frac{A\xi + B}{C\xi + D}$ is strictly increasing if and only if $AD - BC > 0$. The function g corresponds to parameters $A = 1$, $B = -1$, $C = 2((1 - A_L)A_H + (1 - A_H)A_L)$, and $D = 4(1 - A_H)(1 - A_L)$. The condition $AD - BC > 0$ is equivalent to $4(1 - A_H)(1 - A_L) + 2((1 - A_L)A_H + (1 - A_H)A_L) > 0$ which holds as $A_H > 0$ and $1 - A_L > 0$, showing the second claim of the proposition. Given that $\text{BELIEFERROR}(\sigma^{\text{RANDOM}}, p; \xi) = \frac{(\mu_H - \mu_L)^2}{2}$ does not depend on ξ , the third claim follows. \square

E Motivation behind using Random as benchmark (Section 2.2)

To motivate our choice of σ^{RANDOM} as a benchmark, we establish that it is the revenue-maximizing policy out of a class of review ordering policies which do not reveal review ratings, i.e., they

1. do not take the review ratings into account, and
2. do not take the actual times of posting of the reviews into account.

Without the first requirement, the platform can maximize revenue by always ordering all positive reviews first; this strategy provides a clearly biased set of reviews which makes it undesirable. Without the second requirement, the platform can use the time between consecutive reviews to infer whether a review is positive – if another review follows soon after, then it is more likely that the review is positive as positive reviews induce purchases faster than negative reviews. Incorporating the actual time therefore has a similar criticism with incorporating review ratings. As a result, we

restrict our attention to ordering policies that only take the relative time of posting of each of the reviews (which does not allow inference of whether a review is positive). We further restrict our attention to stationary policies: the exact ordering rule does not depend on the round.

We define a general class of ordering policies that satisfy these requirements. Let S_w be the family of all c -sized subsets of $\{1, \dots, w\}$. We denote by $\sigma^{\text{RANDOM}(w, \psi)}$ the review ordering policy that selects c reviews from the most recent w according to a fixed probability distribution $\psi \in \Delta(S_w)$. We denote $\sigma^{\text{RANDOM}(w, \psi)}$ as $\sigma^{\text{RANDOM}(w)}$ when $\psi = \mathcal{U}(S_w)$ is uniform distribution over S_w . Let $\mathcal{C} = \{\sigma^{\text{RANDOM}(w, \psi)}, w \geq c, \psi \in \Delta(S_w)\}$ be the class of review ordering policies $\sigma^{\text{RANDOM}(w, \psi)}$ for some $w \geq c$ and distribution $\psi \in \Delta(S_w)$. Note that σ^{RANDOM} is not strictly in the class \mathcal{C} but it can be defined as the limit of $\sigma^{\text{RANDOM}(w)}$ as $w \rightarrow \infty$ (see Proposition E.1, Appendix E.1). In Proposition E.2 (Appendix E.2), we show that (1) σ^{RANDOM} yields strictly larger revenue than any ordering policy in \mathcal{C} and (2) the revenue of σ^{RANDOM} can be approximated arbitrarily well by ordering policies in the class \mathcal{C} . Therefore, we can think of the revenue under σ^{RANDOM} as the maximum revenue achievable by any policy in this class.

E.1 Formal definition of Random as the limit of $\sigma^{\text{random}(w)}$

We now show that the policy $\sigma^{\text{RANDOM}}(\mathcal{H}_t)$ can be defined as the limit of $\sigma^{\text{RANDOM}(w)}(\mathcal{H}_t)$ as $w \rightarrow \infty$. For a vector of review ratings $\mathbf{z} \in \{0, 1\}^c$, we denote $N(\mathbf{z}) := N_{\mathbf{z}} = \sum_{i=1}^c z_i$. Proposition E.1 implies that as $w \rightarrow \infty$, the draws from $\sigma^{\text{RANDOM}(w)}$ across different rounds are independent and the distribution $\sigma^{\text{RANDOM}(w)}(\mathcal{H}_t)$ approaches the distribution of σ^{RANDOM} .

Proposition E.1. *For any review rating vectors $\mathbf{z}^{(1)}, \mathbf{z}^{(2)}, \dots, \mathbf{z}^{(k)} \in \{0, 1\}^c$ and rounds $t^{(1)} < t^{(2)} < \dots < t^{(k)}$, the distribution of reviews by $\sigma^{\text{RANDOM}(w)}$ approaches the one of σ^{RANDOM} as $w \rightarrow \infty$, i.e.,*

$$\lim_{w \rightarrow \infty} \mathbb{P}[\forall i \in \{1, \dots, k\} : \sigma^{\text{RANDOM}(w)}(\mathcal{H}_{t^{(i)}}) = \mathbf{z}^{(i)}] = \prod_{i=1}^k \mathbb{P}[\sigma^{\text{RANDOM}}(\mathcal{H}_{t^{(i)}}) = \mathbf{z}^{(i)}].$$

We first show that as w goes to infinity, the c reviews selected by $\sigma^{\text{RANDOM}(w)}(\mathcal{H}_{t^{(i)}})$ at each round $t^{(i)}$ come from the infinite pool of reviews $\{X_{-i}\}_{i=1}^{\infty}$ with high probability. Let $n_w = \lfloor \sqrt{w - t^{(k)}} \rfloor$ be the largest integer such that $n_w^2 \leq w - t^{(k)}$ and let w be large enough so that $n_w \geq 1$. We define $\mathcal{E}^{\text{ORDERINF}}$ as the event that for every $i \in \{1, \dots, k\}$ the c reviews selected by $\sigma^{\text{RANDOM}(w)}(\mathcal{H}_{t^{(i)}})$ come from the set $\{X_{-1}, \dots, X_{-n_w^2}\}$. Lemma E.1 shows that $\mathcal{E}^{\text{ORDERINF}}$ happens with high probability.

Lemma E.1. $\mathbb{P}[\mathcal{E}^{\text{ORDERINF}}] \geq \left[\frac{\binom{n_w^2}{c}}{\binom{w}{c}} \right]^k$. Furthermore, $\lim_{w \rightarrow \infty} \mathbb{P}[\mathcal{E}^{\text{ORDERINF}}] = 1$.

Given that $\sigma^{\text{RANDOM}(w)}$ selects from the infinite negative pool of reviews with high probability (if $\mathcal{E}^{\text{ORDERINF}}$ holds), we now derive a concentration bound for the reviews in this pool. Recall that $X_{-\ell} \sim_{i.i.d.} \text{Bern}(\mu)$ for $\ell \in \{1, \dots, n_w^2\}$. We partition those reviews into n_w groups each containing n_w reviews. Let $\mathcal{E}^{\text{CONC}}$ be the event that, for each group, the average review rating concentrates around the group's mean, i.e.

$$\mathcal{E}^{\text{CONC}} = \left\{ \forall j \in \{1, \dots, n_w\} : \sum_{\ell=(j-1) \cdot n_w + 1}^{j \cdot n_w} X_{-\ell} \in [\mu(n_w - n_w^{2/3}), \mu(n_w + n_w^{2/3})] \right\}.$$

Our next lemma shows that the concentration event $\mathcal{E}^{\text{CONC}}$ happens with high probability.

Lemma E.2. $\mathbb{P}[\mathcal{E}^{\text{CONC}}] \geq 1 - 2n_w \exp\left(-\frac{n_w^{1/3}\mu}{3}\right)$. Furthermore, $\lim_{w \rightarrow \infty} \mathbb{P}[\mathcal{E}^{\text{CONC}}] = 1$.

In order to decompose the left-hand-side of Proposition E.1 into a product of probabilities, we show the following independence lemma.

Lemma E.3. Let $(y_1, \dots, y_{n_w^2}) \in \{0, 1\}^{n_w^2}$. Then conditioned on $(X_{-1}, \dots, X_{-n_w^2}) = (y_1, \dots, y_{n_w^2})$ and $\mathcal{E}^{\text{ORDERINF}}$, the events $\{\sigma^{\text{RANDOM}}(\mathcal{H}_{t(i)}) = \mathbf{z}^{(i)}\}$ for $i = 1, \dots, k$ are independent.

Having shown independence across rounds, we prove that, for any round $t^{(i)}$, $\sigma^{\text{RANDOM}(w)}(\mathcal{H}_{t(i)})$ is close to σ^{RANDOM} . Let \mathcal{A} be the set of review rating sequences for which event $\mathcal{E}^{\text{CONC}}$ holds i.e.

$$\mathcal{A} = \left\{ (y_1, \dots, y_{n_w^2}) \in \{0, 1\}^{n_w^2} : \forall j \in \{1, \dots, n_w\}, \sum_{\ell=(j-1) \cdot n_w + 1}^{j \cdot n_w} y_\ell \in [\mu(n_w - n_w^{2/3}), \mu(n_w + n_w^{2/3})] \right\}$$

Our next lemma shows that if the events $\mathcal{E}^{\text{ORDERINF}}$ and $\mathcal{E}^{\text{CONC}}$ hold, then for any round $t^{(i)}$, the distribution of $\sigma^{\text{RANDOM}(w)}(\mathcal{H}_{t(i)})$ is close to the distribution of $\sigma^{\text{RANDOM}}(\mathcal{H}_{t(i)})$. To ease notation, we let $X_{-1:n_w^2} := (X_{-1}, \dots, X_{-n_w^2})$ and $y_{1:n_w^2} := (y_1, \dots, y_{n_w^2})$.

Lemma E.4. There exists some function $f : \mathbb{N} \rightarrow \mathbb{R}$ satisfying $\lim_{n \rightarrow \infty} f(n) = 0$ such that assuming that $\mathcal{E}^{\text{ORDERINF}}$ and $\mathcal{E}^{\text{CONC}}$ hold, for any $\mathbf{z} = (z_1, \dots, z_c) \in \{0, 1\}^c$ and $(y_1, \dots, y_{n_w^2}) \in \mathcal{A}$:

$$\left| \mathbb{P}[\sigma^{\text{RANDOM}(w)}(\mathcal{H}_{t(i)}) = \mathbf{z} | X_{-1:n_w^2} = y_{1:n_w^2}] - \mathbb{P}[\sigma^{\text{RANDOM}}(\mathcal{H}_{t(i)}) = \mathbf{z}] \right| \leq f(n_w).$$

Proof of Proposition E.1. By Lemma E.1 and Lemma E.2, we can upper bound the probability that neither of $\mathcal{E}^{\text{ORDERINF}}$ nor $\mathcal{E}^{\text{CONC}}$ holds by

$$\mathbb{P}[(\neg \mathcal{E}^{\text{ORDERINF}}) \cup (\neg \mathcal{E}^{\text{CONC}})] \leq 1 - \left[\frac{\binom{n_w^2}{c}}{\binom{w}{c}} \right]^k + 2n_w \exp\left(-\frac{n_w^{1/3}\mu}{3}\right) \rightarrow_{w \rightarrow \infty} 0. \quad (36)$$

Hence, we assume that $\mathcal{E}^{\text{ORDERINF}}$ and $\mathcal{E}^{\text{CONC}}$ hold and thus focus on sequences $(X_{-1}, \dots, X_{-n_w^2}) \in \mathcal{A}$. By the law of total probability and the independence of $\sigma^{\text{RANDOM}(w)}$ across rounds (i.e. Lemma E.3):

$$\begin{aligned} & \mathbb{P}[\forall i \in \{1, \dots, k\} : \sigma^{\text{RANDOM}(w)}(\mathcal{H}_{t(i)}) = \mathbf{z}^{(i)} | \mathcal{E}^{\text{ORDERINF}}, \mathcal{E}^{\text{CONC}}] \\ &= \sum_{(y_1, \dots, y_{n_w^2}) \in \mathcal{A}} \underbrace{\prod_{i=1}^k \mathbb{P}[\sigma^{\text{RANDOM}(w)}(\mathcal{H}_{t(i)}) = \mathbf{z}^{(i)} | \mathcal{E}^{\text{ORDERINF}}, X_{-1:n_w^2} = y_{1:n_w^2}]}_{(\text{DEC})} \cdot \underbrace{\mathbb{P}[X_{-1:n_w^2} = y_{1:n_w^2} | \mathcal{E}^{\text{ORDERINF}}, \mathcal{E}^{\text{CONC}}]}_{\text{SEQ}(y_{1:n_w^2})}. \end{aligned} \quad (37)$$

By Lemma E.4, we can upper and lower bound this decomposition by

$$\prod_{i=1}^k \left(\mathbb{P}[\sigma^{\text{RANDOM}}(\mathcal{H}_{t(i)}) = \mathbf{z}^{(i)}] - f(n_w) \right) \leq (\text{DEC}) \leq \prod_{i=1}^k \left(\mathbb{P}[\sigma^{\text{RANDOM}}(\mathcal{H}_{t(i)}) = \mathbf{z}^{(i)}] + f(n_w) \right). \quad (38)$$

Recall that \mathcal{A} consists of all sequences $(y_1, \dots, y_{n_w^2})$ that satisfy the event $\mathcal{E}^{\text{CONC}}$. Summing across those sequences and using the independence between the choice of $\sigma^{\text{RANDOM}(w)}$ (which determines $\mathcal{E}^{\text{ORDERINF}}$) and the values of $X_{-1:n_w^2}$ (which determine $\mathcal{E}^{\text{CONC}}$), we obtain

$$\sum_{(y_1, \dots, y_{n_w^2}) \in \mathcal{A}} \text{SEQ}(y_{1:n_w^2}) = \mathbb{P}[\mathcal{E}^{\text{CONC}} | \mathcal{E}^{\text{ORDERINF}}] = \underbrace{\mathbb{P}[\mathcal{E}^{\text{CONC}}]}_{\text{Lemma E.2}} \rightarrow_{w \rightarrow \infty} 1. \quad (39)$$

We conclude the proof by combining (36), (37), (38), (39) and using the fact that $f(n_w) \rightarrow_{w \rightarrow \infty} 0$. \square

E.1.1 High probability bound on $\mathcal{E}^{\text{OrderInf}}$ (Proof of Lemma E.1)

Proof of Lemma E.1. Fix a particular round $t^{(i)}$. Note that since $n_w = \lfloor \sqrt{w - t^{(k)}} \rfloor$, the w most recent reviews that $\sigma^{\text{RANDOM}(w)}(\mathcal{H}_{t^{(i)}})$ considers contain $(X_{-1}, \dots, X_{-n_w^2})$. The probability that all c reviews come from that set is $\frac{\binom{n_w^2}{c}}{\binom{w}{c}}$. For each round in $t^{(1)}, \dots, t^{(k)}$, the subset of c reviews is independently drawn by $\sigma^{\text{RANDOM}(w)}$. Thus, $\mathbb{P}[\mathcal{E}^{\text{ORDERINF}}] \geq \left[\frac{\binom{n_w^2}{c}}{\binom{w}{c}} \right]^k$. Next notice that $n_w^2 \geq w - t^{(k)} - 1$ and therefore, $\frac{\binom{n_w^2}{c}}{\binom{w}{c}} \geq \frac{(w - t^{(k)} - 1)^c}{w^c} \rightarrow_{w \rightarrow \infty} 1$ showing the second part. \square

E.1.2 Concentration for infinite pool (Proof of Lemma E.2)

Proof of Lemma E.2. For a fixed $j \in \{1, \dots, n_w\}$, by Chernoff bound, it holds that

$$\mathbb{P}\left[\sum_{\ell=(j-1)n_w+1}^{jn_w} X_{-\ell} \in [\mu(n_w - n_w^{2/3}), \mu(n_w + n_w^{2/3})]\right] \geq 1 - 2 \exp\left(-\frac{n_w^{1/3} \mu}{3}\right).$$

By union bound on the n_w groups, the event $\mathcal{E}^{\text{CONC}}$ has probability at least $1 - 2n_w \exp\left(-\frac{n_w^{1/3} \mu}{3}\right)$. Note that $n_w \exp\left(-\frac{n_w^{1/3} \mu}{3}\right) \rightarrow_{w \rightarrow \infty} 0$, which yields the second part. \square

E.1.3 Independence of $\sigma^{\text{RANDOM}(w)}$ across rounds (Proof of Lemma E.3)

Proof of Lemma E.3. By definition, $\sigma^{\text{RANDOM}(w)}$ samples a c -sized subset of reviews independently at every round. Thus, conditioning on the subset being sampled from $(X_{-1}, \dots, X_{-n_w^2})$, i.e. $\mathcal{E}^{\text{ORDERINF}}$, and also conditioning on the exact values of these ratings, i.e. $(X_{-1}, \dots, X_{-n_w^2}) = (y_1, \dots, y_{n_w^2})$, the draws of $\sigma^{\text{RANDOM}(w)}(\mathcal{H}_{t^{(i)}})$ are independent for $i = 1, \dots, k$. \square

E.1.4 Single round approximation (Proof of Lemma E.4)

Let $\sigma^{\text{RANDOM}(w)}(\mathcal{H}_{t^{(i)}}) = (Z_1, \dots, Z_c) \in \{0, 1\}^c$ be the c review ratings chosen from the most recent w reviews. Recall that when $\mathcal{E}^{\text{ORDERINF}}$ holds, $\sigma^{\text{RANDOM}(w)}$ selects reviews from $(X_{-1}, \dots, X_{-n_w^2})$ and that $(X_{-1}, \dots, X_{-n_w^2}) = (y_1, \dots, y_{n_w^2})$. Let $S_j = \{y_\ell\}_{\ell=(j-1)n_w+1}^{jn_w}$ for $j \in \{1, \dots, n_w\}$ be a partition of all reviews $(y_1, \dots, y_{n_w^2})$ into n_w groups of n_w reviews each. We first show that the reviews drawn by $\sigma^{\text{RANDOM}(w)}(\mathcal{H}_{t^{(i)}})$ come from different groups S_j with high probability. Let $\mathcal{E}^{\text{DIFF_GROUPS}}$ be the event that no two indices $m_1 < m_2 \in \{1, \dots, c\}$ are such that Z_{m_1} and Z_{m_2} belong to the same group S_j for some $j \in \{1, \dots, n_w\}$. Our next lemma lower bounds the probability of $\mathcal{E}^{\text{DIFF_GROUPS}}$.

Lemma E.5. *Assume that $\mathcal{E}^{\text{ORDERINF}}$ and $\mathcal{E}^{\text{CONC}}$ hold. The probability of any two selected indices being of different groups is: $\mathbb{P}[\mathcal{E}^{\text{DIFF_GROUP}}] \geq 1 - c! \cdot c^2 \frac{n_w^{2c-1}}{(n_w^2 - c)^c}$ and thus $\lim_{w \rightarrow \infty} \mathbb{P}[\mathcal{E}^{\text{DIFF_GROUP}}] = 1$.*

Proof. Let $\mathcal{E}_{m_1, m_2, j}$ be the event that $Z_{m_1}, Z_{m_2} \in S_j$. Notice that $\mathcal{E}^{\text{DIFF_GROUP}}$ is exactly the event that none of the $\mathcal{E}_{m_1, m_2, j}$ hold. Further, note that $Z_{m_1}, Z_{m_2} \in S_j$ implies that $Z_{m_1}, Z_{m_1+1}, \dots, Z_{m_2} \in S_j$ since we output an ordered set of c reviews by recency. There are thus $\binom{n_w}{m_2-m_1+1}$ ways to choose $Z_{m_1}, Z_{m_1+1}, \dots, Z_{m_2}$ to be in the same group S_j . Hence the probability of $\mathcal{E}_{m_1, m_2, j}$ is at most

$$\mathbb{P}[\mathcal{E}_{m_1, m_2, j}] \leq \frac{\binom{n_w}{m_1-1} \binom{n_w}{m_2-m_1+1} \binom{n_w}{c-m_2}}{\binom{n_w}{c}}$$

as there are at most $\binom{n_w}{m_1-1}$ ways to choose Z_1, \dots, Z_{m_1-1} and at most $\binom{n_w}{m_2-m_1}$ ways to choose Z_{m_2+1}, \dots, Z_c . Using the inequalities $\frac{(n-k)^k}{k!} \leq \binom{n}{k} \leq n^k$, we obtain that

$$\mathbb{P}[\mathcal{E}_{m_1, m_2, j}] \leq c! \cdot \frac{n_w^{2(m_1-1)} n_w^{m_2-m_1+1} n_w^{2(c-m_2)}}{(n_w^2 - c)^c} = c! \cdot \frac{n_w^{2c-(m_2-m_1+1)}}{(n_w^2 - c)^c} \leq c! \cdot \frac{n_w^{2c-2}}{(n_w^2 - c)^c}$$

since $m_2 - m_1 + 1 \geq 2$. Thus, via union bound over $m_1, m_2 \in \{1, \dots, c\}$ and $j \in \{1, \dots, n_w\}$, i.e., $c^2 n_w$ events, the probability of $\mathcal{E}^{\text{DIFF_GROUP}}$ is lower bounded as follows

$$\mathbb{P}[\mathcal{E}^{\text{DIFF_GROUP}}] \geq 1 - c^2 n_w \cdot c! \cdot \frac{n_w^{2c-2}}{(n_w^2 - c)^c} = 1 - c! \cdot c^2 \frac{n_w^{2c-1}}{(n_w^2 - c)^c} \xrightarrow{n_w \rightarrow \infty} 1.$$

□

When the event $\mathcal{E}^{\text{DIFF_GROUPS}}$ holds then the reviews Z_m for $m = 1, \dots, c$ come from different groups $\{S_j\}_{j=1}^{n_w}$. We show that when this happens the values of the reviews Z_m are independent. Let $\mathcal{E}_{j_1, \dots, j_c}$ be the event that review Z_m comes from group S_{j_m} for $m = 1, \dots, c$. The next lemma shows that conditioned on the event $\mathcal{E}_{j_1, \dots, j_c}$, the values of Z_1, \dots, Z_c are independent.¹⁰ Recall that $\mathcal{E}^{\text{CONC}}$ implies that $(y_1, \dots, y_{n_w^2}) \in \mathcal{A}$.

Lemma E.6. *Let $j_1, j_2, \dots, j_c \in \{1, \dots, n_w\}$ be group indices with $j_1 < j_2 < \dots \leq j_c$. Conditioned on $\mathcal{E}_{j_1, \dots, j_c}, \mathcal{E}^{\text{CONC}}, \mathcal{E}^{\text{ORDERINF}}$, and $(y_1, \dots, y_{n_w^2}) \in \mathcal{A}$, for any vector of review ratings $\mathbf{z} \in \{0, 1\}^c$, the events $\{Z_m = z_m\}$ for $m = 1, \dots, c$ are independent. Furthermore,*

$$\mathbb{P}[Z_m = z_m] = \left(\frac{\sum_{\ell \in S_{j_m}} y_\ell}{n_w} \right)^{z_m} \left(1 - \frac{\sum_{\ell \in S_{j_m}} y_\ell}{n_w} \right)^{1-z_m}.$$

Proof. For any specific review ratings $y_{\ell_1} \in S_{j_1}, \dots, y_{\ell_m} \in S_{j_m}$, applying Bayes rule

$$\mathbb{P}[Z_1 = y_{\ell_1}, \dots, Z_c = y_{\ell_c} | \mathcal{E}_{j_1, \dots, j_c}] = \frac{\mathbb{P}[Z_1 = y_{\ell_1}, \dots, Z_m = y_{\ell_m}, \mathcal{E}_{j_1, \dots, j_c}]}{\mathbb{P}[\mathcal{E}_{j_1, \dots, j_c}]}$$

There are $(n_w)^c$ choices for the set of c ratings (Z_1, \dots, Z_c) so that $Z_m \in S_{j_m}$ for $m = 1, \dots, c$ and $\binom{(n_w)^2}{c}$ total number of choices. Hence $\mathcal{E}_{j_1, \dots, j_c}$ holds with probability $\mathbb{P}[\mathcal{E}_{j_1, \dots, j_c}] = \frac{(n_w)^c}{\binom{(n_w)^2}{c}}$. Given that there is exactly one choice for the c reviews (Z_1, \dots, Z_c) which satisfies $Z_1 = y_{\ell_1}, \dots, Z_c = y_{\ell_c}$:

$$\mathbb{P}[Z_1 = y_{\ell_1}, \dots, Z_c = y_{\ell_c} | \mathcal{E}_{j_1, \dots, j_c}] = \frac{\frac{1}{\binom{(n_w)^2}{c}}}{\frac{(n_w)^c}{\binom{(n_w)^2}{c}}} = \frac{1}{(n_w)^c}. \quad (40)$$

¹⁰This is not the case without conditioning on $\mathcal{E}_{j_1, \dots, j_c}$. As an example, consider two ordered reviews (Z_1, Z_2) drawn uniformly from the three ordered review ratings $(1, 0, 1, 1)$. Conditioned on $Z_1 = 0$ then $Z_2 = 1$ deterministically while conditioned on $Z_1 = 1$ then $\mathbb{P}[Z_2 = 0 | Z_1 = 1] = \frac{1}{4}$. As a results Z_1 and Z_2 are correlated.

Since this holds for any $y_{\ell_1} \in S_{j_1}, \dots, y_{\ell_m} \in S_{j_m}$ we obtain independence of $\{Z_m = z_m\}$ for $m = 1, \dots, c$. By summing (40) over $y_{\ell_2}, \dots, y_{\ell_c}$, for any particular $y_{\ell_m} \in S_{j_m}$, we have $\mathbb{P}[Z_m = y_{\ell_m} | \mathcal{E}_{\ell_1, \dots, \ell_c}] = \frac{1}{n_w}$. Therefore, the probability that Z_m has a particular value $z_m \in \{0, 1\}$ is equal to the fraction of $y_{\ell_m} \in S_{j_m}$ which have value z_m i.e.

$$\mathbb{P}[Z_m = z_m] = \left(\frac{\sum_{\ell \in S_{j_m}} y_{\ell}}{n_w} \right)^{z_m} \left(1 - \frac{\sum_{\ell \in S_{j_m}} y_{\ell}}{n_w} \right)^{1-z_m}.$$

□

We next show that conditioned on the events $\mathcal{E}^{\text{DIFF_GROUP}}$, $\mathcal{E}^{\text{CONC}}$, and $\mathcal{E}^{\text{ORDERINF}}$ the distribution of $\sigma^{\text{RANDOM}(w)}(\mathcal{H}_{t(i)})$ is close to the distribution of σ^{RANDOM} . Given that the latter consists of i.i.d. $\text{Bern}(\mu)$, $\mathbb{P}(\sigma^{\text{RANDOM}} = \mathbf{z}) = \mu^{N(\mathbf{z})} (1 - \mu)^{c-N(\mathbf{z})}$ for vector of review ratings $\mathbf{z} \in \{0, 1\}^c$. Recall that $N(\mathbf{z}) := N_{\mathbf{z}} = \sum_{i=1}^c z_i$. Using the independence across different groups (Lemma E.6), the following lemma shows that $\mathbb{P}[\sigma^{\text{RANDOM}(w)}(\mathcal{H}_{t(i)}) = \mathbf{z}]$ has a similar decomposition as $\mathbb{P}[\sigma^{\text{RANDOM}} = \mathbf{z}]$.

Lemma E.7. *Conditioned on $\mathcal{E}_{j_1, \dots, j_c}$, $\mathcal{E}^{\text{CONC}}$, $\mathcal{E}^{\text{ORDERINF}}$, and $(y_1, \dots, y_{n_w}) \in \mathcal{A}$, for any vector of review ratings $\mathbf{z} \in \{0, 1\}^c$, $\mathbb{P}[\sigma^{\text{RANDOM}(w)}(\mathcal{H}_{t(i)}) = \mathbf{z}]$ is lower bounded by $(\mu(1 - n_w^{-\frac{1}{3}}))^{N(\mathbf{z})} (1 - \mu(1 + n_w^{-\frac{1}{3}}))^{c-N(\mathbf{z})}$ and upper bounded by $(\mu(1 + n_w^{-\frac{1}{3}}))^{N(\mathbf{z})} (1 - \mu(1 - n_w^{-\frac{1}{3}}))^{c-N(\mathbf{z})}$*

Proof. Using Lemma E.6:

$$\mathbb{P}[\sigma^{\text{RANDOM}(w)}(\mathcal{H}_{t(i)}) = \mathbf{z} | \mathcal{E}_{j_1, \dots, j_c}] = \prod_{m=1}^c \left(\frac{\sum_{\ell \in S_{j_m}} y_{\ell}}{n_w} \right)^{z_m} \left(1 - \frac{\sum_{\ell \in S_{j_m}} y_{\ell}}{n_w} \right)^{1-z_m}.$$

As $(y_1, \dots, y_{n_w}) \in \mathcal{A}$, we have that $\mu(1 - n_w^{-\frac{1}{3}}) \leq \frac{\sum_{\ell \in S_{j_m}} y_{\ell}}{n_w} \leq \mu(1 + n_w^{-\frac{1}{3}})$ for all $m = 1, \dots, c$. Applying these inequalities for each m yields

$$\begin{aligned} \mathbb{P}[\sigma^{\text{RANDOM}(w)}(\mathcal{H}_{t(i)}) = \mathbf{z} | \mathcal{E}_{j_1, \dots, j_c}] \\ \in \left[(\mu(1 - n_w^{-\frac{1}{3}}))^{N(\mathbf{z})} (1 - \mu(1 + n_w^{-\frac{1}{3}}))^{c-N(\mathbf{z})}, (\mu(1 + n_w^{-\frac{1}{3}}))^{N(\mathbf{z})} (1 - \mu(1 - n_w^{-\frac{1}{3}}))^{c-N(\mathbf{z})} \right]. \end{aligned}$$

for any $\mathcal{E}_{j_1, \dots, j_c}$. As $\mathcal{E}^{\text{DIFF_GROUPS}}$ is the union of events $\mathcal{E}_{j_1, \dots, j_c}$ for all indices j_1, \dots, j_c , this implies

$$\begin{aligned} \mathbb{P}[\sigma^{\text{RANDOM}(w)}(\mathcal{H}_{t(i)}) = \mathbf{z} | \mathcal{E}^{\text{DIFF_GROUPS}}] \\ \in \left[(\mu(1 - n_w^{-\frac{1}{3}}))^{N(\mathbf{z})} (1 - \mu(1 + n_w^{-\frac{1}{3}}))^{c-N(\mathbf{z})}, (\mu(1 + n_w^{-\frac{1}{3}}))^{N(\mathbf{z})} (1 - \mu(1 - n_w^{-\frac{1}{3}}))^{c-N(\mathbf{z})} \right]. \end{aligned}$$

□

Proof of Lemma E.4. The law of total probability yields the following decomposition

$$\begin{aligned} \mathbb{P}[(Z_1, \dots, Z_c) = \mathbf{z}] &= \mathbb{P}[(Z_1, \dots, Z_c) = \mathbf{z} | \mathcal{E}^{\text{DIFF_GROUP}}] \mathbb{P}[\mathcal{E}^{\text{DIFF_GROUP}}] \\ &\quad + \mathbb{P}[(Z_1, \dots, Z_c) = \mathbf{z} | \neg \mathcal{E}^{\text{DIFF_GROUP}}] \mathbb{P}[\neg \mathcal{E}^{\text{DIFF_GROUP}}]. \end{aligned}$$

By this decomposition and Lemma E.7, we can lower bound $\mathbb{P}[(Z_1, \dots, Z_c) = \mathbf{z}]$ by

$$\mathbb{P}[(Z_1, \dots, Z_c) = \mathbf{z} | \mathcal{E}^{\text{DIFF_GROUP}}] \mathbb{P}[\mathcal{E}^{\text{DIFF_GROUP}}] \geq \underbrace{(\mu(1 - n_w^{-\frac{1}{3}}))^{N(\mathbf{z})} (1 - \mu(1 + n_w^{-\frac{1}{3}}))^{c-N(\mathbf{z})} \mathbb{P}[\mathcal{E}^{\text{DIFF_GROUP}}]}_{\text{LB}(n_w, \mathbf{z})}$$

and upper bound the same probability by

$$\begin{aligned} \mathbb{P}[(Z_1, \dots, Z_c) = \mathbf{z} | \mathcal{E}^{\text{DIFF_GROUP}}] + 1 - \mathbb{P}[\mathcal{E}^{\text{DIFF_GROUP}}] \\ \leq \underbrace{(\mu(1 + n_w^{-\frac{1}{3}}))^{N(\mathbf{z})}(1 - \mu(1 - n_w^{-\frac{1}{3}}))^{c-N(\mathbf{z})} + 1 - \mathbb{P}[\mathcal{E}^{\text{DIFF_GROUP}}]}_{\text{UB}(n_w, \mathbf{z})} \end{aligned}$$

Therefore, $|\mathbb{P}[(Z_1, \dots, Z_c) = \mathbf{z}] - \mu^{N(\mathbf{z})}(1 - \mu)^{c-N(\mathbf{z})}| \leq f(\mathbf{z}, n_w)$ where

$$f(\mathbf{z}, n_w) = \max \left(|\text{LB}(n_w, \mathbf{z}) - \mu^{N(\mathbf{z})}(1 - \mu)^{c-N(\mathbf{z})}|, |\text{UB}(n_w, \mathbf{z}) - \mu^{N(\mathbf{z})}(1 - \mu)^{c-N(\mathbf{z})}| \right).$$

Recall that $\mathbb{P}[\sigma^{\text{RANDOM}} = \mathbf{z}] = \mu^{N(\mathbf{z})}(1 - \mu)^{c-N(\mathbf{z})}$. Since $\mathbb{P}[\mathcal{E}^{\text{DIFF_GROUP}}] \rightarrow_{n_w \rightarrow \infty} 1$ (Lemma E.5) and $n_w^{-\frac{1}{3}} \rightarrow 0$, we obtain that $f(\mathbf{z}, n_w) \rightarrow_{n \rightarrow \infty} 0$ for any $\mathbf{z} \in \{0, 1\}^c$ because

$$\text{UB}(n_w, \mathbf{z}), \text{LB}(n_w, \mathbf{z}) \rightarrow_{n_w \rightarrow \infty} \mu^{N(\mathbf{z})}(1 - \mu)^{c-N(\mathbf{z})}.$$

Hence, $f(n_w) = \max_{\mathbf{z} \in \{0, 1\}^c} f(\mathbf{z}, n_w)$ satisfies the desired property and concludes the proof. \square

E.2 Random maximizes revenue in a class of review-oblivious policies

Proposition E.2. *For any fixed price p satisfying Assumption 3.1 and any finite window w ,*

- $\mathcal{U}(S_w)$ maximizes revenue over $\Delta(S_w)$; $\text{REV}(\sigma^{\text{RANDOM}(w)}, p) = \max_{\psi \in \Delta(S_w)} \text{REV}(\sigma^{\text{RANDOM}(w, \psi)}, p)$
- $\text{REV}(\sigma^{\text{RANDOM}(w)}, p)$ is strictly increasing in w
- $\lim_{w \rightarrow \infty} \text{REV}(\sigma^{\text{RANDOM}(w)}, p) = \text{REV}(\sigma^{\text{RANDOM}}, p)$.

To show the proposition we show three useful lemmas. The first lemma (proof in E.2.1) shows that $\mathcal{U}(S_w)$ maximizes the revenue over all distributions ψ over S_w , which suggests for any fixed window w , more randomness in the ordering implies more revenue.

Lemma E.8. *For any fixed price p satisfying Assumption 3.1 the uniform distribution $\mathcal{U}(S_w)$ maximizes the revenue over $\Delta(S_w)$, i.e., $\text{REV}(\sigma^{\text{RANDOM}(w)}, p) = \max_{\psi \in \Delta(S_w)} \text{REV}(\sigma^{\text{RANDOM}(w, \psi)}, p)$.*

The next lemma (proof in Section E.2.3) shows the revenue of $\sigma^{\text{RANDOM}(w)}$ is strictly increasing in the window size w for any price p satisfying Assumption 3.1.

Lemma E.9. *For any fixed price p satisfying Assumption 3.1, $\text{REV}(\sigma^{\text{RANDOM}(w)}, p)$ is strictly increasing in the window w .*

The third lemma (proof in Section E.2.4) shows that as the window w goes to infinity the revenue of $\sigma^{\text{RANDOM}(w)}$ approaches the revenue of σ^{RANDOM} .

Lemma E.10. *For any fixed price p satisfying Assumption 3.1, $\lim_{w \rightarrow \infty} \text{REV}(\sigma^{\text{RANDOM}(w)}, p) = \text{REV}(\sigma^{\text{RANDOM}}, p)$.*

Proof of Proposition E.2. The proof follows by combining Lemmas E.8, E.9, and E.10. \square

E.2.1 Proof of Lemma E.8.

To prove Lemma E.8, we first characterize the stationary distribution of the reviews generated by $\sigma^{\text{RANDOM}(w,\psi)}$ in way analogous to Lemma 3.1. In particular for any state of the w most recent review ratings $\mathbf{z} \in \{0,1\}^w$, the purchase probability is $q_{\mathbf{z}}^{\psi} := \sum_{S \in \mathcal{S}_w} \psi(S) \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(\sum_{i \in S} z_i) \geq p]$. A useful quantity is the inverse purchase rate conditioned on k positive reviews:

$$\iota_k^{\psi} = \frac{1}{\binom{w}{k}} \sum_{\mathbf{z} \in \{0,1\}^w : N_{\mathbf{z}} = k} \frac{1}{q_{\mathbf{z}}^{\psi}}.$$

Intuitively, ι_k^{ψ} is the average number of rounds the process of the most recent w reviews spends at a review state with k positive reviews. The next lemma (proof in Appendix E.2.2) shows that the uniform distribution minimizes this purchase rate for any k and characterizes the equality conditions.

Lemma E.11. *For any $k \in \{0, \dots, c\}$ and any probability distribution ψ , the inverse purchase rate under ψ is at least the inverse purchase rate under $\mathcal{U}(\mathcal{S}_w)$, i.e., $\iota_k^{\psi} \geq \iota_k^{\mathcal{U}(\mathcal{S}_w)}$. Equality $\iota_k^{\psi} = \iota_k^{\mathcal{U}(\mathcal{S}_w)}$ is achieved if and only if $q_{\mathbf{z}}^{\psi} = q_{\mathbf{z}'}^{\psi}$ for all states $\mathbf{z}, \mathbf{z}' \in \{0,1\}^w$ with $N_{\mathbf{z}} = N_{\mathbf{z}'} = k$.*

Let $\mathbf{Z}_t = (Z_{t,1}, \dots, Z_{t,w})$ be the most recent w reviews at round t . We show that \mathbf{Z}_t is a time-homogenous Markov chain on $\{0,1\}^w$. If \mathbf{Z}_t is at state $\mathbf{z} \in \{0,1\}^w$, a purchase occurs with probability $q_{\mathbf{z}}^{\psi}$. If a purchase occurs, a new review is left and the state transitions to $\mathbf{Z}_{t+1} = (1, z_1, \dots, z_{w-1})$ if the review is positive (with probability μ) and to $\mathbf{Z}_{t+1} = (0, z_1, \dots, z_{w-1})$ if the review is negative (with probability $1 - \mu$). If there is no purchase, the state remains the same ($\mathbf{Z}_{t+1} = \mathbf{Z}_t$). Given that p is a non-absorbing price (Assumption 3.1), for every state of reviews $\mathbf{z} \in \{0,1\}^w$, the purchase probability $q_{\mathbf{z}}^{\psi}$ is positive and the probability of any new review is strictly positive (since $\mu \in (0,1)$). Then \mathbf{Z}_t can reach every state from every other state with positive probability (i.e. it is a single-recurrence-class Markov chain with no transient states), and hence \mathbf{Z}_t has a unique stationary distribution denoted by π . Our next lemma characterizes the form of this stationary distribution. For a vector of review ratings $\mathbf{z} \in \{0,1\}^w$, $N_{\mathbf{z}} = \sum_{i=1}^w z_i$ denotes the number of positive ratings.

Lemma E.12. *For any price p satisfying Assumption 3.1 and any distribution ψ , the stationary distribution of the Markov chain \mathbf{Z}_t under the review ordering policy $\sigma^{\text{RANDOM}(w,\psi)}$ is given by*

$$\pi_{\mathbf{z}}^{\psi} = \kappa^{\psi} \cdot \frac{\mu^{N_{\mathbf{z}}} (1 - \mu)^{w - N_{\mathbf{z}}}}{q_{\mathbf{z}}^{\psi}} \quad \text{where } \kappa^{\psi} = \frac{1}{\mathbb{E}_{K \sim \text{Binom}(w,\mu)} [\iota_K^{\psi}]}.$$

Proof. Similar to the proof of Lemma 3.1, we invoke Lemma B.2. Recall that Lemma B.2 starts with a Markov chain \mathcal{M} with state space \mathcal{S} , transition matrix M , and stationary distribution $\{\pi(s)\}_{s \in \mathcal{S}}$. For any function f on \mathcal{S} , it then transforms \mathcal{M} into a new Markov chain \mathcal{M}_f which remains at every state s with probability $1 - f(s)$ and transitions according to M otherwise. The lemma establishes that the stationary distribution of \mathcal{M}_f is given by $\pi_f = \{\kappa \cdot \frac{\pi(s)}{f(s)}\}_{s \in \mathcal{S}}$ where $\kappa = 1 / \sum_{s \in \mathcal{S}} \frac{\pi(s)}{f(s)}$.

In the language of Lemma B.2, \mathbf{Z}_t corresponds to \mathcal{M}_f , the state space \mathcal{S} to $\{0,1\}^w$, and f is a function that expresses the purchase probability at a given state, i.e., $f(\mathbf{z}) = q_{\mathbf{z}}^{\psi}$. Note that with probability $1 - f(\mathbf{z})$, \mathbf{Z}_t remains at the same state (as there is no purchase).

To apply Lemma B.2, we need to show that whenever there is a purchase, \mathbf{Z}_t transitions according to a Markov chain with stationary distribution $\mu^{N\mathbf{z}}(1-\mu)^{w-N\mathbf{z}}$. Consider the Markov chain \mathcal{M} which always replaces the w -th last review with a new $\text{Bern}(\mu)$ review. This process has stationary distribution equal to the above numerator and \mathbf{Z}_t transitions according to \mathcal{M} upon a purchase, i.e., with probability $f(\mathbf{z})$. Hence, by Lemma B.2, a stationary distribution for \mathbf{Z}_t is

$$\pi_{\mathbf{z}}^{\psi} = \kappa^{\psi} \cdot \frac{\mu^{N\mathbf{z}}(1-\mu)^{w-N\mathbf{z}}}{q_{\mathbf{z}}^{\psi}} \quad \text{where } \kappa^{\psi} = \frac{1}{\mathbb{E}_{Z_1, \dots, Z_w \sim i.i.d. \text{Bern}(\mu)} \left[\frac{1}{q_{(Z_1, \dots, Z_w)}^{\psi}} \right]}.$$

This is the unique stationary distribution as \mathbf{Z}_t is irreducible and aperiodic. Expanding over the number of positive reviews $k \in \{0, \dots, w\}$, the lemma follows as the expectation in the denominator of κ^{ψ} can be expressed as

$$\sum_{k=0}^w \left[\sum_{\mathbf{z} \in \{0,1\}^w : N\mathbf{z}=k} \frac{1}{q_{\mathbf{z}}^{\psi}} \right] \mu^k (1-\mu)^{w-k} = \sum_{k=0}^w \ell_k^{\psi} \binom{w}{k} \mu^k (1-\mu)^{w-k} = \mathbb{E}_{K \sim \text{Binom}(w, \mu)} [\ell_K^{\psi}].$$

□

Having established the stationary distribution of \mathbf{Z}_t we give an expression for the revenue of $\sigma^{\text{RANDOM}(w, \psi)}$ (similar to Proposition 3.2).

Lemma E.13. *For any price p satisfying Assumption 3.1 and any distribution $\psi \in \Delta(\mathcal{S}_w)$, the revenue of $\sigma^{\text{RANDOM}(w, \psi)}$ is given by $\text{REV}(\sigma^{\text{RANDOM}(w, \psi)}, p) = p \cdot \kappa^{\psi}$ where $\kappa^{\psi} = \frac{1}{\mathbb{E}_{K \sim \text{Binom}(w, \mu)} [\ell_K^{\psi}]}$.*

Proof. Using Eq. (1) and the Ergodic theorem, the revenue of $\sigma^{\text{RANDOM}(w, \psi)}$ is

$$\text{REV}_1(\sigma^{\text{RANDOM}(w, \psi)}, p) = \liminf_{T \rightarrow \infty} \frac{\mathbb{E}[\sum_{t=1}^T p \mathbf{1}_{\Theta_t + h(\Phi_t) \geq p}]}{T} = p \liminf_{T \rightarrow \infty} \frac{\mathbb{E}[\sum_{t=1}^T q_{\mathbf{Z}_t}^{\psi}]}{T} = p \sum_{\mathbf{z} \in \{0,1\}^w} \pi_{\mathbf{z}}^{\psi} q_{\mathbf{z}}^{\psi}.$$

The second equality uses that the ex-ante purchase probability for review state \mathbf{Z}_t is $\mathbb{E}[\mathbf{1}_{\Theta_t + h(\Phi_t) \geq p}] = q_{\mathbf{Z}_t}^{\psi}$ and the law of iterated expectation. The third equality expresses the revenue of the stationary distribution via the Ergodic theorem. Expanding $\pi_{\mathbf{z}}^{\psi}$ by Lemma E.12, the $q_{\mathbf{z}}^{\psi}$ term cancels out:

$$\text{REV}_1(\sigma^{\text{RANDOM}(w, \psi)}, p) = p \cdot \kappa^{\psi} \cdot \left[\sum_{\mathbf{z} \in \{0,1\}^w} \mu^{N\mathbf{z}} (1-\mu)^{w-N\mathbf{z}} \right].$$

The proof is concluded by noting that the term in the square brackets equals 1, since it is the sum over all probabilities of $\text{Binom}(w, \mu)$. □

Proof of Lemma E.8. For $\kappa^{\psi} = \frac{1}{\mathbb{E}_{K \sim \text{Binom}(w, \mu)} [\ell_K^{\psi}]}$, the expectation in $\frac{1}{\kappa^{\psi}}$ can be expressed by summing over the number of positive review ratings $k \in \{0, 1, \dots, w\}$:

$$\frac{1}{\kappa^{\psi}} = \sum_{k=0}^w \ell_k^{\psi} \binom{w}{k} \mu^k (1-\mu)^{w-k} \geq \sum_{k=0}^w \ell_k^{\mathcal{U}(\mathcal{S}_w)} \binom{w}{k} \mu^k (1-\mu)^{w-k} = \frac{1}{\kappa^{\mathcal{U}(\mathcal{S}_w)}}.$$

where the inequality follows from Lemma E.11. Thus, $\kappa^{\psi} \leq \kappa^{\mathcal{U}(\mathcal{S}_w)}$ for any distribution ψ . Lemma E.13 then implies that $\text{REV}_1(\sigma^{\text{RANDOM}(w, \psi)}, p) = p \cdot \kappa^{\psi} \leq p \cdot \kappa^{\mathcal{U}(\mathcal{S}_w)} = \text{REV}_1(\sigma^{\text{RANDOM}(w)}, p)$. □

E.2.2 Proof of Lemma E.11.

The following lemma characterizes the inverse purchase rate for k positive review ratings under the uniform distribution, i.e., $\iota_k^{\mathcal{U}(\mathcal{S}_w)}$. To ease notation, let $q(n) := \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(n) \geq p]$ be the purchase probability given $n \in \{0, 1, \dots, c\}$ positive reviews.

Lemma E.14. *For any number of positive review ratings $k \in \{0, \dots, w\}$, the inverse purchase rate for k under the uniform distribution on \mathcal{S}_w is given by $\iota_k^{\mathcal{U}(\mathcal{S}_w)} = \frac{\binom{w}{k}}{\sum_{n=\max(0, c+k-w)}^{\min(c, k)} q(n) \binom{c}{n} \binom{w-c}{k-n}}$.*

Proof of Lemma E.14. Recall that $\mathcal{U}(\mathcal{S}_w)$ places a probability of $\frac{1}{\binom{w}{c}}$ on every subset of \mathcal{S}_w . Thus, by counting the number of occurrences of each term $q(n)$, the purchase probability at any state $\mathbf{z} \in \{0, 1\}^w$ with $N_{\mathbf{z}} = k$ can be expressed as:

$$q_{\mathbf{z}}^{\mathcal{U}(\mathcal{S}_w)} = \sum_{S \in \mathcal{S}_w} \frac{1}{\binom{w}{c}} q\left(\sum_{i \in S} z_i\right) = \sum_{n=\max(0, c+k-w)}^{\min(c, k)} q(n) \frac{\binom{k}{n} \binom{w-k}{c-n}}{\binom{w}{c}} = \sum_{n=\max(0, c+k-w)}^{\min(c, k)} q(n) \frac{\binom{c}{n} \binom{w-c}{k-n}}{\binom{w}{k}}. \quad (41)$$

The second equality uses that for any number of positive reviews $n \in [\max(0, c+k-w), \min(c, k)]$ there are $\binom{k}{n} \binom{w-k}{c-n}$ ways to choose $S \in \mathcal{S}_w$ with $\sum_{i \in S} z_i = n$: $\binom{k}{n}$ ways to choose the n indices $i \in S$ from the k indices i where $z_i = 1$, and $\binom{w-k}{c-n}$ ways to choose the remaining $c-n$ indices $i \in S$ from the $w-k$ indices where $z_i = 0$. The third equality uses the binomial coefficient identity

$$\frac{\binom{k}{n} \binom{w-k}{c-n}}{\binom{w}{c}} = \frac{\frac{k!}{n!(k-n)!} \cdot \frac{(w-k)!}{(c-n)!(w-k-c+n)!}}{\frac{w!}{c!(w-c)!}} = \frac{\frac{c!}{n!(c-n)!} \cdot \frac{(w-c)!}{(k-n)!(w-k-c+n)!}}{\frac{w!}{k!(w-k)!}} = \frac{\binom{c}{n} \binom{w-c}{k-n}}{\binom{w}{k}}.$$

Given that the right-hand side of Eq. (41) does not depend on the review ratings \mathbf{z} ,

$$\iota_k^{\mathcal{U}(\mathcal{S}_w)} = \frac{1}{\binom{w}{k}} \sum_{\mathbf{z} \in \{0, 1\}^w : N_{\mathbf{z}} = k} \frac{1}{q_{\mathbf{z}}^{\mathcal{U}(\mathcal{S}_w)}} = \frac{1}{\sum_{n=\max(0, c+k-w)}^{\min(c, k)} q(n) \frac{\binom{c}{n} \binom{w-c}{k-n}}{\binom{w}{k}}} = \frac{\binom{w}{k}}{\sum_{n=\max(0, c+k-w)}^{\min(c, k)} q(n) \binom{c}{n} \binom{w-c}{k-n}}$$

which completes the proof. \square

Proof of Lemma E.11. By Jensen's inequality applied to the convex function $x \rightarrow \frac{1}{x}$,

$$\iota_k^{\psi} = \frac{1}{\binom{w}{k}} \sum_{\mathbf{z} \in \{0, 1\}^w : N_{\mathbf{z}} = k} \frac{1}{q_{\mathbf{z}}^{\psi}} \geq \frac{\binom{w}{k}}{\sum_{\mathbf{z} \in \{0, 1\}^w : N_{\mathbf{z}} = k} q_{\mathbf{z}}^{\psi}}. \quad (42)$$

To show the lemma, it suffices to prove that for any probability distribution $\psi \in \Delta(\mathcal{S}_w)$, the lower bound term $\frac{\binom{w}{k}}{\sum_{\mathbf{z} \in \{0, 1\}^w : N_{\mathbf{z}} = k} q_{\mathbf{z}}^{\psi}}$ on the left-hand side is independent of ψ and equal to $\iota_k^{\mathcal{U}(\mathcal{S}_w)}$. We start by analyzing the sum in the denominator $\sum_{\mathbf{z} \in \{0, 1\}^w : N_{\mathbf{z}} = k} q_{\mathbf{z}}^{\psi}$. Letting $q(n) := \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(n) \geq p]$ be the purchase probability given $n \in \{0, 1, \dots, c\}$ positive reviews, recalling that $q_{\mathbf{z}}^{\psi} = \sum_{S \in \mathcal{S}_w} \psi(S) q(\sum_{i \in S} z_i)$, and rearranging, we express the sum of interest as:

$$\sum_{\mathbf{z} \in \{0, 1\}^w : N_{\mathbf{z}} = k} q_{\mathbf{z}}^{\psi} = \sum_{\mathbf{z} \in \{0, 1\}^w : N_{\mathbf{z}} = k} \sum_{S \in \mathcal{S}_w} \psi(S) q\left(\sum_{i \in S} z_i\right) = \sum_{S \in \mathcal{S}_w} \psi(S) \sum_{\mathbf{z} \in \{0, 1\}^w : N_{\mathbf{z}} = k} q\left(\sum_{i \in S} z_i\right).$$

By counting the number of occurrences of each term $q(n)$, the inner sum can be expressed as:

$$\sum_{\mathbf{z} \in \{0,1\}^w: N_{\mathbf{z}}=k} q\left(\sum_{i \in S} z_i\right) = \sum_{n=\max(0, c+k-w)}^{\min(c, k)} q(n) \binom{c}{n} \binom{w-c}{k-n} \quad (43)$$

since for any number of positive reviews $n \in [\max(0, c+k-w), \min(c, k)]$ there are $\binom{c}{n} \binom{w-c}{k-n}$ ways to choose sequences $\mathbf{z} \in \{0,1\}^w$ with $N_{\mathbf{z}} = k$ such that $\sum_{i \in S} z_i = n$: there are $\binom{c}{n}$ ways to choose the subsequence $(z_i)_{i \in S} \in \{0,1\}^c$ such that $\sum_{i \in S} z_i = n$ and $\binom{w-c}{k-n}$ ways to choose the complementary subsequence $(z_i)_{i \notin S} \in \{0,1\}^{w-c}$ such that $\sum_{i \notin S} z_i = k - n$.

Observing that the right-hand side of (43) is independent of S , using the fact that $\sum_{S \in \mathcal{S}_w} \psi(S) = 1$ (as ψ is a probability distribution) as well as inequality Eq. (42) and Lemma E.14, we obtain:

$$\iota_k^\psi \geq \frac{\binom{w}{k}}{\sum_{\mathbf{z} \in \{0,1\}^w: N_{\mathbf{z}}=k} q_{\mathbf{z}}^\psi} = \frac{\binom{w}{k}}{\sum_{n=\max(0, c+k-w)}^{\min(c, k)} q(n) \binom{c}{n} \binom{w-c}{k-n}} = \iota_k^{\mathcal{U}(\mathcal{S}_w)}.$$

This completes the proof of the inequality statement of the lemma. With respect to the equality statement, for any $k \in \{0, \dots, c\}$, the equality conditions of Jensen's inequality imply that Eq. (42) holds with equality if and only if $q_{\mathbf{z}}^\psi = q_{\mathbf{z}'}^\psi$ for any $\mathbf{z}, \mathbf{z}' \in \{0,1\}^w$ with $N_{\mathbf{z}} = N_{\mathbf{z}'} = k$. \square

E.2.3 Proof of Lemma E.9.

We first show the suboptimality of the distribution $\psi_w \in \Delta(\mathcal{S}_{w+1})$ which ignores the $(w+1)$ -th most recent review and places equal probability on every c -sized subset of $\{1, \dots, w\}$.

Lemma E.15. *For any price p satisfying Assumption 3.1, the revenue of ψ_w is strictly suboptimal, i.e., $\text{REV}(\sigma^{\text{RANDOM}(w+1, \psi_w)}, p) < \max_{\psi \in \Delta(\mathcal{S}_{w+1})} \text{REV}(\sigma^{\text{RANDOM}(w+1, \psi)}, p)$.*

Proof. By Lemma E.13, the revenue of ψ_w can be expressed as:

$$\text{REV}_1(\sigma^{\text{RANDOM}(w+1, \psi_w)}, p) = p \cdot \kappa^{\psi_w} = \frac{p}{\sum_{k=0}^w \iota_k^{\psi_w} \binom{w}{k} \mu^k (1-\mu)^{w-k}}.$$

By Lemma E.11, it is thus sufficient to show that there is some number of positive ratings $k \in \{0, \dots, c\}$ such that $\iota_k^{\psi_w} > \iota_k^{\mathcal{U}(\mathcal{S}_w)}$. For the sake of contradiction assume this is not the case, i.e., $\iota_k^{\psi_w} = \iota_k^{\mathcal{U}(\mathcal{S}_w)}$ for all $k \in \{0, \dots, c\}$. The equality condition of Lemma E.11 applied for each $k \in \{0, \dots, c\}$ implies

$$q_{\mathbf{z}}^{\psi_w} = q_{\mathbf{z}'}^{\psi_w} \text{ for any } \mathbf{z}, \mathbf{z}' \in \{0,1\}^w \text{ with } N_{\mathbf{z}} = N_{\mathbf{z}'} \quad (44)$$

Letting $q(n) := \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(n) \geq p]$ denote the purchase probability given n positive reviews, we show below (by induction on n) that (44) implies $q(n) = q(0)$ for any number of positive review ratings $n \in \{0, \dots, c\}$, which contradicts the fact that the price p is non-degenerate.

For the base case of the induction ($n = 1$), consider two settings with one positive review: in the first one the positive review is the $(w+1)$ -th most recent review, while in the second one the positive review is the w -th most recent review. Formally, we apply Eq. (44) for $\mathbf{z} = (0, \dots, 0, 0, 1)$ and $\mathbf{z}' = (0, \dots, 0, 1, 0)$. As ψ_w places zero mass on any subset containing the last review implies that, the purchase probability at \mathbf{z} is $q_{\mathbf{z}}^{\psi_w} = q(0)$. The purchase probability at \mathbf{z}' equals $q(1)$ when the

w -th review is chosen and $q(0)$ otherwise implying $q_{\mathbf{z}'}^{\psi_w} = q(1)\mathbb{P}_{S \sim \psi_w}[w \in S] + q(0)\mathbb{P}_{S \sim \psi_w}[w \notin S]$. Observing that $\mathbb{P}_{S \sim \psi_w}[w \in S] > 0$ and solving the equality $q_{\mathbf{z}'}^{\psi_w} = q_{\mathbf{z}'}^{\psi_w}$ implies $q(1) = q(0)$.

For the induction step ($n-1 \rightarrow n$), suppose that $q(n') = q(0)$ for all $n' < n$. We apply Eq. (44) for $\mathbf{z} = (0, \dots, 0, 1, \dots, 1)$ ($w-n$ zeros followed by n ones) and $\mathbf{z}' = (0, \dots, 0, 1, \dots, 1, 0)$ ($w-n-1$ zeros, followed by n ones, followed by one). If the state of the reviews is \mathbf{z} , as ψ_w never selects the $(w+1)$ -th most recent review, any c reviews selected by ψ_w contain at most $n-1$ positive review ratings and the induction hypothesis implies that the purchase probability is thus equal to $q(0)$ regardless of the choice of the c reviews. Thus, $q_{\mathbf{z}'}^{\psi_w} = q(0)$. Suppose the state of the review is \mathbf{z}' . If the selected set of the c selected reviews contains all the reviews at indices $I_n = \{w-n, w-n+1, \dots, w\}$, the number of positive reviews is n and the purchase probability $q(n)$. Otherwise, the c selected reviews contain at most $n-1$ positive review ratings and by the induction hypothesis the purchase probability is $q(0)$. Thus, $q_{\mathbf{z}'}^{\psi_w} = q(n)\mathbb{P}_{S \sim \psi_w}[I_n \subseteq S] + q(0)\mathbb{P}_{S \sim \psi_w}[I_n \not\subseteq S]$. Observing that $\mathbb{P}_{S \sim \psi_w}[I_n \subseteq S] > 0$ and solving the equality $q_{\mathbf{z}'}^{\psi_w} = q_{\mathbf{z}'}^{\psi_w}$ implies $q(n) = q(0)$, which finishes the induction step and the proof. \square

Proof of Lemma E.9. Note that $\sigma^{\text{RANDOM}(w+1, \psi_w)} = \sigma^{\text{RANDOM}(w)}$. By Lemma E.8 and Lemma E.15, $\text{REV}(\sigma^{\text{RANDOM}(w)}, p)$ is strictly increasing in w as for any window $w \geq c$,

$$\begin{aligned} \text{REV}(\sigma^{\text{RANDOM}(w)}, p) &= \text{REV}(\sigma^{\text{RANDOM}(w+1, \psi_w)}, p) \\ &< \max_{\psi \in \Delta(\mathcal{S}_{w+1})} \text{REV}(\sigma^{\text{RANDOM}(w+1, \psi)}, p) = \text{REV}(\sigma^{\text{RANDOM}(w+1)}, p). \end{aligned}$$

\square

E.2.4 Proof of Lemma E.10

Given that $\text{REV}(\sigma^{\text{RANDOM}(w)}, p)$ can be expressed as a function of a w i.i.d. $\text{Bern}(\mu)$ trials (Lemma E.13), we first show that $K \sim \text{Binom}(w, \mu)$ with w trials and success probability μ is well-concentrated around its mean. Formally, the event $\mathcal{E}^{\text{CONC}}(w) = \{K \in [\mu(w - w^{\frac{2}{3}}), \mu(w + w^{\frac{2}{3}})]\}$ occurs with high probability.

Lemma E.16. *For any $w \geq c$, $\mathbb{P}[\mathcal{E}^{\text{CONC}}(w)] \geq 1 - 2 \exp\left(-\frac{w^{\frac{1}{3}}\mu}{3}\right)$ and thus $\lim_{w \rightarrow \infty} \mathbb{P}[\mathcal{E}^{\text{CONC}}(w)] = 1$.*

Proof. By Chernoff bound, $\mathbb{P}_{K \sim \text{Binom}(w, \mu)}[|K - \mu w| \leq \mu w^{\frac{2}{3}}] \geq 1 - 2 \exp\left(-\frac{w^{\frac{1}{3}}\mu}{3}\right)$. Taking the limit as $w \rightarrow \infty$ yields the result. \square

Recall that $\iota_k^{\mathcal{U}(\mathcal{S}_w)} = \frac{1}{\binom{w}{c}} \sum_{\mathbf{z} \in \{0,1\}^w: N_{\mathbf{z}}=k} \frac{1}{q_{\mathbf{z}}^{\mathcal{U}(\mathcal{S}_w)}}$ is the inverse purchase rate conditioned on k positive review ratings and $q(n) := \mathbb{P}_{\Theta \sim \mathcal{F}}[\Theta + h(n) \geq p]$ denotes the purchase probability given $n \in \{0, 1, \dots, c\}$ positive reviews. The following lemma connects the inverse purchase rate of $\sigma^{\text{RANDOM}(w)}$ (assuming that the above concentration holds) with the revenue of σ^{RANDOM} .

Lemma E.17. *There exists a threshold $M > 0$ such that for any window $w > M$, it holds that*

$$\text{REV}(\sigma^{\text{RANDOM}}, p) = p \cdot \lim_{w \rightarrow \infty} \left(\frac{1}{\mathbb{E}_{K \sim \text{Binom}(w, \mu)} \left[\iota_K^{\mathcal{U}(\mathcal{S}_w)} \middle| \mathcal{E}^{\text{CONC}}(w) \right]} \right).$$

Proof. Let M be large enough so that for $w > M$, $\mu(w - w^{-\frac{2}{3}}) \geq c$ and $w \geq c + \mu(w + w^{-\frac{2}{3}})$, which implies that $\min(c, k) = c$ and $\max(0, c + k - w) = 0$ for $k \in [\mu(w - w^{-\frac{2}{3}}), \mu(w + w^{-\frac{2}{3}})]$. By Lemma E.14 for $w > M$ and $k \in [\mu(w - w^{-\frac{2}{3}}), \mu(w + w^{-\frac{2}{3}})]$, it holds that

$$\iota_k^{\mathcal{U}(S_w)} = \frac{\binom{w}{k}}{\sum_{n=\max(0, c+k-w)}^{\min(c, k)} q(n) \binom{c}{n} \binom{w-c}{k-n}} = \frac{1}{\sum_{n=\max(0, c+k-w)}^{\min(c, k)} q(n) \frac{\binom{c}{n} \binom{w-c}{k-n}}{\binom{w}{k}}} = \frac{1}{\sum_{n=0}^c q(n) \frac{\binom{c}{n} \binom{w-c}{k-n}}{\binom{w}{k}}}.$$

One of the terms in the denominator is the binomial coefficient ratio $\frac{\binom{w-c}{k-n}}{\binom{w}{k}}$. Letting $\underline{\mu}_w = \mu - w^{-\frac{1}{3}}$ and $\bar{\mu}_w = \mu + w^{-\frac{1}{3}}$, we show below

$$\left(\frac{w-c+1}{w}\right)^c (\bar{\mu}_w)^c (1 - \underline{\mu}_w)^{c-n} \geq \frac{\binom{w-c}{k-n}}{\binom{w}{k}} \geq (\underline{\mu}_w - \frac{n-1}{w})^n (1 - \bar{\mu}_w - \frac{c-n-1}{w})^{c-n} \quad (45)$$

which implies that $\mathbb{E}_{K \sim \text{Binom}(w, \mu)} [\iota_K^{\mathcal{U}(S_w)} | \mathcal{E}^{\text{CONC}}(w)]$ is lower and upper bounded respectively by

$$\frac{\left(\frac{w}{w-c+1}\right)^c}{\sum_{n=0}^c q(n) \binom{c}{n} (\bar{\mu}_w)^c (1 - \underline{\mu}_w)^{c-n}} \text{ and } \frac{1}{\sum_{n=0}^c q(n) \binom{c}{n} (\underline{\mu}_w - \frac{n-1}{w})^n (1 - \bar{\mu}_w - \frac{c-n-1}{w})^{c-n}}.$$

Given that $\underline{\mu}_w, \bar{\mu}_w \rightarrow_{w \rightarrow \infty} \mu$ (as $w^{-\frac{1}{3}} \rightarrow_{w \rightarrow \infty} 0$), $\frac{w-c+1}{w} \rightarrow_{w \rightarrow \infty} 1$ and $\frac{n-1}{w}, \frac{c-n-1}{w} \rightarrow_{w \rightarrow \infty} 0$, these lower and upper bounds converge to $\frac{1}{\mathbb{E}_{N \sim \text{Binom}(c, \mu)} [q(N)]}$ as $w \rightarrow \infty$ which is equal to $\frac{p}{\text{REV}(\sigma^{\text{RANDOM}}, p)}$ (by Proposition 3.1) and yields the result. We conclude the proof by showing Eq. (45). We expand the binomial coefficient ratio in the denominator of Eq. (45) as:

$$\frac{\binom{w-c}{k-n}}{\binom{w}{k}} = \frac{\frac{(w-c)!}{(k-n)!(w-c-k+n)!}}{\frac{w!}{k!(w-k)!}} = \frac{\prod_{i=1}^n (k-i+1) \prod_{j=1}^{c-n} (w-k-j+1)}{\prod_{l=1}^c (w-l+1)}.$$

Observing that $k \geq k-i+1 \geq k-n+1$ for $k \in [1, n]$, $w-k \geq w-k-j+1 \geq w-k-c+n+1$ for $j \in [1, c-n]$, and $w \geq w-l+1 \geq w-c+1$ for $l \in [1, c]$, we can bound the binomial coefficient

$$\underbrace{\left(\frac{k}{w-c+1}\right)^n \left(\frac{w-k}{w-c+1}\right)^{c-n}}_{=\left(\frac{w-c+1}{w}\right)^c \left(\frac{k}{w}\right)^n \left(1 - \frac{k}{w}\right)^{c-n}} \geq \frac{\binom{w-c}{k-n}}{\binom{w}{k}} \geq \underbrace{\left(\frac{k-n+1}{w}\right)^n \left(\frac{w-k-c+n+1}{w}\right)^{c-n}}_{=\left(\frac{k}{w} - \frac{n-1}{w}\right)^n \left(1 - \frac{k}{w} - \frac{c-n-1}{w}\right)^{c-n}}.$$

Given that $\frac{k}{w} \in [\underline{\mu}_w, \bar{\mu}_w]$ (as $k \in [\mu(w - w^{-\frac{2}{3}}), \mu(w + w^{-\frac{2}{3}})]$), the above proves (45). \square

Proof of Lemma E.10. Lemma E.13 connects the revenue of $\sigma^{\text{RANDOM}}(w)$ to $\mathbb{E}_{K \sim \text{Binom}(w, \mu)} [\iota_K^{\mathcal{U}(S_w)}]$. By the law of total expectation, the latter term can be expanded to:

$$\underbrace{\mathbb{E}_{K \sim \text{Binom}(w, \mu)} [\iota_K^{\mathcal{U}(S_w)} | \mathcal{E}^{\text{CONC}}(w)]}_{\rightarrow_{w \rightarrow \infty} \frac{p}{\text{REV}(\sigma^{\text{RANDOM}}, p)} \text{ (Lemma E.17)}} \underbrace{\mathbb{P}[\mathcal{E}^{\text{CONC}}(w)]}_{\rightarrow_{w \rightarrow \infty} 1 \text{ (Lemma E.16)}} + \underbrace{\mathbb{E}_{K \sim \text{Binom}(w, \mu)} [\iota_K^{\mathcal{U}(S_w)} | \neg \mathcal{E}^{\text{CONC}}(w)]}_{\rightarrow_{w \rightarrow \infty} 0 \text{ (Lemma E.16)}} \underbrace{\mathbb{P}[\neg \mathcal{E}^{\text{CONC}}(w)]}_{\rightarrow_{w \rightarrow \infty} 0 \text{ (Lemma E.16)}}$$

We now show that $\iota_k^{\mathcal{U}(S_w)} \leq \frac{1}{q(0)}$, which combined with Lemma E.16 implies that the second term goes to 0 as $w \rightarrow \infty$. As $h(n)$ is increasing, $q(n) \geq q(0)$ for all number of positive review ratings n

and thus $q_{\mathbf{z}}^{\mathcal{U}(S_w)} = \frac{1}{\binom{w}{c}} \sum_{S \in \mathcal{S}_w} q(\sum_{i \in S} z_i) \geq q(0)$ for all review rating states $\mathbf{z} \in \{0, 1\}^w$. Hence,

$$\iota_k^{\mathcal{U}(S_w)} = \frac{1}{\binom{w}{k}} \sum_{\mathbf{z} \in \{0, 1\}^w: N_{\mathbf{z}}=k} \frac{1}{q_{\mathbf{z}}^{\mathcal{U}(S_w)}} \leq \frac{1}{q(0)}.$$

The proof is concluded by invoking Lemma E.13 and bounding the limit as $w \rightarrow \infty$ for the expansion of $\mathbb{E}_{K \sim \text{Binom}(w, \mu)} \left[\iota_K^{\mathcal{U}(S_w)} \right]$ (using Lemmas E.16 and E.17 as well as the boundedness of $\iota_k^{\mathcal{U}(S_w)}$):

$$\text{REV}_1(\sigma^{\text{RANDOM}(w)}, p) = \frac{p}{\mathbb{E}_{K \sim \text{Binom}(w, \mu)} \left[\iota_K^{\mathcal{U}(S_w)} \right]} \xrightarrow{w \rightarrow \infty} \frac{p}{\overline{\text{REV}(\sigma^{\text{RANDOM}}, p)}} = \text{REV}(\sigma^{\text{RANDOM}}, p).$$

□