Consumer Search and Social Learning in Agentic Markets

Brendan Lucier[1] **Nicole Immorlica**[1,2] **Markus Mobius**[1] **Aleksandrs Slivkins**[1]
Daniel G. Goldstein[1] Jake M. Hofman[1] Sonia Jaffe[1] David M. Rothschild[1]

[1]Microsoft Research [2]Yale University

Abstract

Motivated by *agentic markets*—where AI tools facilitate search—we propose a model of consumer search and market-wide learning. A sequence of consumers engage in costly search to acquire progressively refined signals of product fit prior to purchase. The market observes acquired signals and post-transaction feedback thereby impacting future searches. We solve for the per-consumer optimal search policy and characterize the steady-state of the learning dynamics. Search technologies (agentic or otherwise) that reduce costs at fixed inspection depth improve long-run learning and consumer surplus. In a special case with screen-level feedback, more informative search also improves these outcomes. In contrast, we show through counterexamples and numerical simulations that more informative search can degrade long-run learning and consumer surplus when the market observes only aggregate inspection and transaction feedback, highlighting the importance of feedback design.

1 Introduction

Recent advances in AI are enabling a new class of *agentic markets*: markets in which consumers and firms delegate search, interaction, and even transactions to AI agents [Rothschild et al., 2025, Hadfield and Koh, 2025, Shahidi et al., 2025]. Consumer-facing systems such as OpenAI's Operator (now ChatGPT agent mode) and Amazon Rufus already allow consumers to offload complex search and comparison tasks, while similar tools on the firm side increasingly automate consumer interaction. In addition to simply reducing the cost of clicking and browsing, agentic interactions can enable new forms of inquiry. Consumers can engage in unscripted, adaptive questioning to assess whether a product or service meets their needs. At the same time, agents can lower the cost of extracting, summarizing, and organizing information generated after transactions, such as consumer experiences and feedback. Together, these capabilities stand to make search both *cheaper*—by enabling consumers to investigate more firms with the same amount of effort—and *more informative*—by allowing consumers to probe more detailed aspects of fit with any given firm prior to transacting.

These changes should be beneficial for individual transactions. Cheaper and more informative search helps consumers identify promising options and avoid poor matches. However, markets are not composed of isolated interactions. Outcomes emerge from the accumulation of many consumer-firm interactions over time, as information generated by early searches and transactions shapes the beliefs and choices of those who arrive later. When the technology governing search and interaction changes, it may therefore alter not only individual decisions, but also how information aggregates at the market level, with consequences for learning and welfare that are not immediately obvious.

As an example, consider a consumer hiring a caterer for an event (see Figure 1). The value of a transaction depends on caterer features. Some features, such as cuisine, dietary accommodations, or calendar availability, determine *fit* with the consumer's needs and in principle can be assessed pre-transaction. But, in practice, today's interfaces limit how many features can be efficiently checked

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

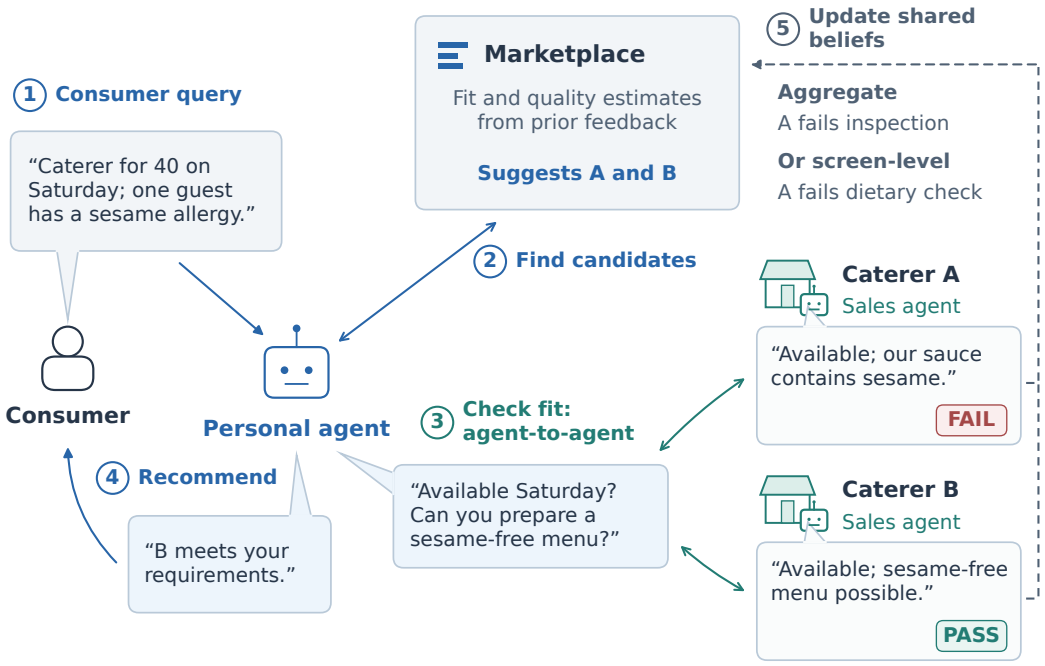


Figure 1: Illustrative agent-assisted search and market learning. A marketplace uses prior feedback to suggest caterers; agent-to-agent conversations check their fit. Inspection outcomes update beliefs for subsequent consumers. Aggregate feedback reports whether inspection passed, whereas screen-level feedback identifies which checks passed or failed. Post-transaction quality feedback is omitted.

in advance, often pushing more detailed assessment into costly back-and-forth conversations or even into the post-transaction phase. An AI agent could ask the caterer, or its agent, whether a particular menu can accommodate a guest's dietary restriction not addressed in the listing. Automating a check that would otherwise require a phone call reduces search costs; resolving a mismatch that would otherwise emerge only when the food is served makes search more informative. Other features are harder to verify in advance and tend to be revealed or evaluated only after a transaction takes place; we refer to these as *quality* features. This might include taste, execution, and reliability, which are typically only learned through experience.

Consumers leverage information generated by others to help estimate quality and the probability of a fit prior to a more costly investigation. Early interactions produce signals through feedback that guide later consumers' choices. Over time, some caterers become effectively well-understood by the market, while others may be effectively *lost* if pessimistic beliefs discourage further exploration. Returning to our example, if the platform observes the answer to each individual query, it can distinguish, e.g., a dietary mismatch from a lack of availability. If it sees only whether the caterer passed inspection, these failures are pooled. Thus, the effects on future consumers depend not only on what the agent learns, but also on what the platform observes and shares.

Our contributions are two-fold: a tractable model that we hope can be the basis for future research, and an analysis of how search costs, pre-transaction information, and feedback observability jointly shape market learning and consumer surplus. These mechanisms are motivated by conversational agents but are not specific to AI; our results also apply to other search technologies with the same information structure.

Model. We develop a tractable model of consumer search in which consumers inspect products to determine fit, generate feedback (through both inspections and transactions), and the market aggregates this information over time. We assume a satisficing model of utility, where each consumer has her own fixed private value for any product that satisfies her needs [Simon, 1956]. Consumer inspection of a product is modeled as a sequence of costly *screens*: pass/fail tests. The search

technology determines the maximum depth of inspection – which screens can be revealed pre-transaction – and the cost of resolving a particular screen. A consumer may stop searching, abandon a firm, or switch between partially inspected firms and return to them later. A consumer that stops searching can transact with any firm whose inspectable screens have been completed.

Optimal search and learning. Consumers enter the market sequentially and (prior to search) update their beliefs based on past consumers' feedback. Despite the resulting complex dynamics, we show that this model converges to a distribution over steady-states and we characterize the long-run learning outcomes. To analyze these dynamics, we characterize optimal search using an index-based policy, building on the Pandora's box literature [Weitzman, 1978, Kleinberg et al., 2016]. This policy, which is efficiently computable, induces an endogenous search order over products and screens that is responsive to consumers' beliefs and values.

Comparative statics. We prove that lowering search costs, at fixed inspection depth, robustly improves market-level learning and consumer surplus in the long-run. We also study more informative search, which shifts information about fit into the pre-purchase stage without changing the underlying probability that a firm's product meets the consumer's need. When all pre-purchase screens can be revealed for a single inspection cost, this improves long-run learning and consumer surplus if the market observes each individual screen outcome. By contrast, if the market observes only whether a firm's inspection was successful, more informative search can reduce both. Pooling failures across screens changes the evidence available to future consumers, potentially causing useful firms to be prematurely abandoned.

Simulations. The analytical results above concern long-run market outcomes; our simulations examine finite-horizon dynamics. We find that cheaper and more informative search improve consumer utility, but learning failures (i.e., useful firms abandoned prematurely) need not be monotone as search becomes more informative. With transcripts, this observed inversion reverses in longer runs; without transcripts, the non-monotonicity persists over the extended horizon and across market variants.

These results identify risks and opportunities from search improvements motivated by agentic interaction. Long-run learning depends critically on how information generated during search and transactions is surfaced and aggregated. These findings highlight the importance of market and platform design choices in determining whether agentic tools ultimately lead to more efficient and effective markets.

Related Work. Our motivation draws on work on AI-mediated search. Shahidi et al. [2025] discuss reductions in search, communication, and preference-elicitation costs, while Kumar et al. [2026] model the matching benefits and communication costs of follow-up questions. Yan et al. [2026] document pre-purchase use of a platform AI assistant for exploratory, hard-to-keyword tasks. In retail field experiments, Fang et al. [2026] find conversion gains consistent with reduced search and information frictions. We study how such changes interact with the feedback available for social learning.

Our model of consumer search is a variant of the *Pandora's box problem* [Weitzman, 1978].[1] Many researchers have investigated variants of Weitzman's highly tractable model; see Beyhaghi and Cai [2024] for a recent survey. Particularly relevant are sequential partial inspections where the contents of box can be partially revealed [Aouad et al., 2025, Gibbard, 2022, Kleinberg et al., 2016] and correlated Pandora's Box models where the consumer can learn about the market as they search [Bikhchandani and Sharma, 1996, Adam, 2001]. In contrast, learning in our model occurs at the platform level; each consumer learns from previous consumers rather than from their own search.

Our setting features consumers that learn from past consumers. It thus connects to social learning with myopic short-lived agents, most notably to the models where both past actions and outcomes are observed [e.g., Kannan et al., 2018, Acemoglu et al., 2022, Banihashem et al., 2023]. In essence, the agents collectively face explore-exploit tradeoff, while individually favoring exploitation. This

[1] Weitzman's model is itself a (non-obvious) special case of Markovian multi-armed bandits [Gittins, 1979, Whittle, 1980].

literature tends to focus on simple models such as multi-armed bandits and study asymptotic "success" vs "failure" of the learning dynamics. It does not address more complex models such as ours.[2]

Other social learning models feature agents who learn from *privately observed signals* as well as the actions taken by past agents [Golub and Sadler, 2016]. Banerjee [1992], Bikhchandani et al. [1992] introduce this model and show that herds result when signals and/or action spaces are coarse. Subsequent work described a variety of conditions necessary for such results, including the richness of the action sets [Ali, 2018b, Lee, 1993] and the signal spaces [Smith and Sørensen, 2000], and costly signal acquisition [Ali, 2018a, Burguet and Vives, 2000, Hendricks et al., 2012, Mueller-Frank and Pai, 2016]. Relative to these literatures, our action and signal spaces are different – being motivated by sequential consumer search – and we focus on quantifying and comparing what is learned in the long-run limit, rather than determining necessary/sufficient conditions for asymptotic success.

2 Model and Preliminaries

We model a two-sided market with sequential consumer arrivals, individual consumer search (as a variant of the Pandora's box problem), and bandit-like feedback across consumers.[3]

Consumers and firms. We have an infinite stream of *consumers* $i \in \mathbb{N}$ and a fixed pool of m *firms* $j \in [m]$. The consumer has a need and receives value $V_i \geq 0$ if this need is met. Each consumer i can transact with at most one firm j at a fixed and known price $P_j \geq 0$, yielding *transaction utility* $U_i^{\mathtt{tr}} = V_i - P_j$ if the need is met by firm j or $U_i^{\mathtt{tr}} = -P_j$ if it is not. Value V_i is private to consumer i and drawn i.i.d. from a known distribution $\mathcal{D}^{\mathtt{v}}$ over $[0, \overline{V}]$, for some finite $\overline{V}$.

Whether consumer i's need is met by firm j is determined by a *screen sequence* $\vec{f}_{ij} = (\, f_{ijk},\, k \in [r]\,)$ of r binary random variables, called *screens*. We think of $\vec{f}_{ij}$ as the outcomes of progressively deeper consumer inspection. Each screen is pass/fail (resp., value 1/0). E.g., f_{ij1} might indicate whether the firm sells the right product; f_{ij2} whether basic product features listed on the webpage are suitable; and so on. The need is met ex post if and only if all screens pass.

The screen sequence $\vec{f}_{ij}$ is drawn independently for each consumer i and firm j, from some distribution that depends on the firm but is the same for all consumers. The screens can be correlated with each other. We write $\pi_{jk} := \Pr[\, f_{ijk} = 1 \mid f_{i,j,1} = \cdots = f_{i,j,k-1} = 1\,]$ for the conditional probability that the k-th screen passes for firm j (and a given consumer), given that all previous screens have passed. [4] We posit $\pi_{jk} \in (0,1)$, i.e., the screen outcome is not deterministic.

Thus, firm j is characterized by the parameters π_{jk}, $k \in [r]$. The consumers do not know these parameters, but they know prior distributions $\mathcal{P}_{jk}$ from which they are independently drawn. The common prior over the firm parameters is given by the tuple $\mathcal{P} := (\, \mathcal{P}_{jk} :\ j \in [m], k \in [r]\,)$.

Each screen $k \in [r]$ has an associated inspection cost $c_k \in [0, \infty]$. Costs are public knowledge and uniform across firms. Costs can be infinite: $c_k = \infty$ means that screen k cannot be inspected, only experienced upon a transaction. The search technology fixes an *inspection depth* $L \in [r]$: the first L screens have finite inspection costs, and any remaining screens can only be experienced upon a transaction.[5]

Definition 2.1. *A* market $\mathcal{M}$ *is defined by firm parameters* $(\, \pi_{jk} :\ j \in [m], k \in [r]\,)$*, prices* P_j*, value distribution* $\mathcal{D}^{\mathtt{v}}$*, common prior* $\mathcal{P}$*, and search costs* $\vec{c} = (\, c_k :\ k \in [r]\,)$.

Remark 2.2. *All of our results will hold* ***pointwise for each market****, as defined above. That is, for each feasible realization of the firm parameters* (π_{jk}).

Market Timing and Consumer Search. The market evolves according to the following protocol:

(P1) Time proceeds in rounds $i \in \mathbb{N}$. Consumer i arrives in round i and leaves when this round is over. All firms are present in all rounds.

[2] One exception is Immorlica et al. [2019] who also study a market where consumers perform optimal search given posterior beliefs shaped by prior consumers' inspections. Their utility and information structures are different from ours, motivated by a focus on diversity rather than search efficacy. Moreover, our model does not fully reveal firm's quality upon inspection.

[3] The technical connections to Weitzman's Pandora's box and multi-armed bandits are spelled out in Appendix A.

[4] π_{jk} can be interpreted as the fraction of consumers for which this conditional event holds.

[5] Many of our results extend to a setting where each consumer strategically chooses their inspection depth; see Section 7.

(P2) Within round i, the consumer engages in sequential search. At each step, she chooses some firm to *inspect*, depending on the outcome of the previous steps. Each firm j can be inspected at most L times: the k-th time, the consumer observes screen f_{ijk} and pays cost $c_k \geq 0$. She may interleave inspections of different firms, abandon a firm, or return to a partially inspected firm. She can stop searching at any point and optionally transact with any "fully inspected" firm, meaning one whose first L screens have been revealed.

(P3) After each round i, some feedback is collected: (i) inspection feedback: the sequence of inspected screens f_{ijk} and their values; (ii) transaction feedback: firm j transacted with (if any), whether consumer's need was satisfied, or why not: first screen k that failed. The current history (the feedback over all rounds so far) is then revealed to the next consumer.

The (total) utility of consumer i, U_i, is her transaction utility $U_i^{\texttt{tr}}$ minus her total search cost, $C_i^{\texttt{tot}}$. Each consumer is Bayesian-rational and seeks to maximize her expected utility $\mathbb{E}[U_i]$. We assume consumers break ties in favor of not inspecting a firm and not transacting.

Preliminaries. A *search policy* is (any) policy used by a consumer to implement the search in (P2).

All expectations are by default over prior $\mathcal{P}$ and value distribution $\mathcal{D}^{\texttt{v}}$. If there's a possibility of confusion, we'll explicitly condition on history, consumer values, or firm parameters.

Fixing market $\mathcal{M}$, a *history* is a sequence $H_n^{\mathcal{M}} = (h_1, h_2, \ldots, h_n)$ for some $n \leq \infty$, where h_i is the feedback collected in round i, as per (P3). Thus, $H_n^{\mathcal{M}}$ is a random variable determined by the V_i's and f_{ijk}'s. Given a history realization $H_{i-1}^{\mathcal{M}} = H$, the posterior beliefs of customer i are denoted $\mathcal{P}_{jk}(H) := P(\pi_{jk} \mid H_{i-1}^{\mathcal{M}} = H)$. The respective *posterior means* are $\hat{\pi}_{jk}(H) = \mathbb{E}\left[\pi_{jk} \mid H_{i-1}^{\mathcal{M}} = H\right]$. As a shorthand, we define $\mathcal{P}_{ijk} := \mathcal{P}_{jk}(H_{i-1}^{\mathcal{M}})$ for the beliefs, and $\hat{\pi}_{ijk} := \hat{\pi}_{jk}(H_{i-1}^{\mathcal{M}})$ for the posterior means, as random variables over all randomness in $H^{\mathcal{M}}$.

Model Interpretation. In practice, a market platform collects and aggregates various consumer feedback signals (such as the ones discussed above) and shares them with future consumers.

Inspecting a firm can be interpreted as a purchase funnel where a consumer passes through increasingly selective screening before choosing to make a purchase. Inspection may happen in various modes, from a phone call to browsing a webpage to having an AI agent browse a webpage to having an AI agent interact with a firm's AI agent. Different screens may correspond to these different modes and/or to different costs for engaging in these modes (e.g., AI tokens or human time spent). We treat the stages as ordered: identifying a product category, for example, precedes checking details on a product page. Cheaper information gathering within an inspectable stage is modeled by a lower c_k when the screen and feedback remain unchanged.

Inspection outcomes in our model are binary, random, and fully revealed to the platform. A failed screen rules out a satisfactory match, while any firm meeting the consumer's need yields the same gross value V_i. Such need-based utility specification is common in the advertising literature (see, e.g., Athey and Ellison [2011]). Consumers differ in value and realized screen outcomes, but the probabilities π_{jk} are common across consumers; these can be viewed as the likelihood that firm j satisfies the needs of a randomly-chosen consumer.[6]

Together, this utility specification, binary screens, and the requirement in (P2) to complete the first L screens before transacting support the persistent policy in Section 3. Obligatory inspection is also common in sequential search [Beyhaghi and Cai, 2024].

Long-Term Outcomes. We are interested in long-term outcomes of a given market, in terms of learning and consumer utility. Specifically, we focus on *limit beliefs* $\lim_{i\to\infty}(\ \mathcal{P}_{ijk}\ )$ and *limit consumer utility* $\lim_{i\to\infty} \mathbb{E}\left[U_i \mid H_{i-1}^{\mathcal{M}}\right]$. (We'll establish in Section 4 that these limits exist.) We ask whether a firm's parameters are learned in the following asymptotic sense:

Definition 2.3. *Firm j is* learned*, given history realization $H_\infty^{\mathcal{M}} = H$, if for all $k \in [r]$ beliefs $\mathcal{P}_{jk}(H_{i-1})$ weakly converge to π_{jk} as $i \to \infty$, where H_{i-1} is the prefix of H visible to consumer i.*

The set of all learned firms will be a key object in our results, with direct utility implications.

[6] In particular, as discussed in Section 1, the first L screens may correspond to (inspectable) features that determine the *fit* of a firm, whereas the latter $r - L$ screens may reflect the inherent *quality*.

ALGORITHM 1: Optimal search policy for customer i

For each firm j, define the ***index*** as $\sigma_{ij} := V_i \cdot \prod_{\ell > L} \hat{\pi}_{j\ell} - P_j - \sum_{\ell \le L} \left(c_\ell \,/\, \prod_{k \in [\ell, L]} \hat{\pi}_{jk} \right)$;
for *each firm j with $\sigma_{ij} > 0$, in order of decreasing index σ_{ij},* **do**
  for *each* $\ell = 1, 2, \ldots, L$ **do**
    Pay c_ℓ to inspect screen $f_{ij\ell}$;
    If $f_{ij\ell} = 0$, break and continue to the next firm.
  Transact with firm j and leave the market *// We must have $f_{ij\ell} = 1$ for all $\ell \in [L]$*
Leave the market without transacting *// No firms with positive index were a fit*

3 Optimal Consumer Search Policy

In this section we describe each consumer's optimal search policy given their beliefs. The sequential screens represent a search funnel and distinguish the cost of inspection from the information revealed before purchase. Our analysis builds on the index-rule approach to search [Weitzman, 1978, Kleinberg et al., 2016] to characterize the feedback that drives market learning.

Our search policy is Algorithm 1, which relies on the current posterior estimates $\hat{\pi}_{jk} = \hat{\pi}_{jk}(H_{i-1}^{\mathcal{M}})$. The policy orders the firms j according to a certain score σ_{ij}, traditionally called *index*. For each firm in this order, the consumer investigates until either observing a failure, in which case the firm is permanently abandoned, or until L successful searches, in which case the consumer transacts.

Theorem 3.1. *Fix customer i with value V_i and history $H_{i-1}^{\mathcal{M}}$. Algorithm 1 maximizes expected utility $\mathbb{E}[U_i \mid V_i, H_{i-1}^{\mathcal{M}}]$ over all search policies (and respects our tie-breaking).*

Although the model permits interleaving inspections across firms, an optimal policy can be chosen to inspect each firm in contiguous steps. The proof first establishes the Pandora reduction within this class of *persistent* policies, then shows that restricting attention to this class is without loss; details appear in Appendix B. This characterization lets us analyze the feedback generated by consumer search and its consequences for market learning.

Under this optimal policy, the feedback generated by inspecting a firm has the following form.

Corollary 3.2. *In each round i, if a given firm j is inspected at all, then all screens $k \in [L]$ are inspected until the first one that fails (if any), and no other screens are inspected.*

4 Feedback and Learning

This section studies the evolution of market history and consumer beliefs. We show that the market converges to a steady-state where some firms are learned and the rest are never again searched. We characterize the distribution over steady-states, including the probability of being learned. Finally, we show that expected consumer utility converges to the utility of a consumer with steady-state beliefs.

To see how posterior beliefs evolve over time, consider their impact on the index in Algorithm 1. Let $\hat{\pi}_{ij}^{\texttt{vec}} := (\, \hat{\pi}_{ijk} : k \in [r] \,)$ be the vector of posterior mean estimates for consumer i and firm j. Let $\sigma_j(\hat{\pi}_{ij}^{\texttt{vec}}, V)$ be the index assigned to firm j by a consumer with value V and posterior mean estimates $\hat{\pi}_{ij}^{\texttt{vec}}$. After some round i, the estimates $\hat{\pi}_{ij}^{\texttt{vec}}$ may fall low enough that $\sigma_j \left(\hat{\pi}_{ij}^{\texttt{vec}}, \overline{V} \right) \le 0$. Then even the highest-valued consumers have non-positive index for firm j, so the firm would never be inspected after round i and its associated posterior distribution would never update. It turns out that this is the only barrier to learning: we show that as long as $\sigma_j \left(\hat{\pi}_{ij}^{\texttt{vec}}, \overline{V} \right) > 0$ for all consumers i, the posterior distributions for firm j will converge in expectation to ground truth.

Proposition 4.1. *Fix market $\mathcal{M}$ and firm j. Almost surely over randomness in history $H^{\mathcal{M}}$, the following happens. Beliefs $\mathcal{P}_{ijk}$ weakly converge to some $\mathcal{P}_{jk}^{\texttt{lim}} \in \Delta(\mathbb{R})$ as $i \to \infty$, for all $k \in [r]$. Moreover, either firm j is learned, in which case $\mathcal{P}_{ijk} \Rightarrow \pi_{jk}$ for all $k \in [r]$, or else $\sigma_j \left(\hat{\pi}_{ij}^{\texttt{vec}}, \overline{V} \right) \le 0$ for some consumer i (and every consumer afterwards), in which case we say that firm j is* lost.

Thus, posterior beliefs converge to *limit beliefs* $\mathcal{P}^{\texttt{lim}} := \left(\mathcal{P}_{jk}^{\texttt{lim}} : j \in [m], k \in r \right)$. In every execution of the market dynamics, each firm is either learned or lost. Learned firms are inspected and transacted with infinitely often, whereas lost firms are never inspected after some finite time.

Proposition 4.1 follows by noting that, due to our assumption that screen probabilities lie strictly between 0 and 1, each firm that is not lost will be inspected and transacted with infinitely often. In this case the market observes an infinite sequence of independent Bernoulli draws for each screen, and hence the associated posterior beliefs weakly converge to their respective true means.

Proposition 4.1 and Corollary 3.2 imply that whether firm j is lost depends only on the progression of its posterior mean estimates, which are driven solely by inspection outcomes and feedback for firm j. This can be used to show that each firm has an independent probability of being lost.

Proposition 4.2. *Fix any market $\mathcal{M}$. Over all randomness in $H^{\mathcal{M}}$, each firm j is lost independently with a probability $\omega_j \in [0,1]$ that depends only on $(\pi_{jk})_k$, $\vec{c}$, P_j and $\overline{V}$. Probability ω_j is weakly increasing in each (finite) coordinate of c and in P_j and weakly decreasing in each π_{jk} and in $\overline{V}$.*

Write $S^{\texttt{lnd}}(\mathcal{M})$ for the set of firms learned in market $\mathcal{M}$; call it the *learned set*. Note that it is a random variable, with randomness coming from the history. Since the precise posterior beliefs for lost firms are not payoff-relevant to any consumer or firm (as they are never searched nor transacted with), $S^{\texttt{lnd}}(\mathcal{M})$ effectively determines the steady-state to which the market converges.

Utility implications. Given a set of firms $S \subseteq [m]$, consider a hypothetical consumer who (a) knows the features of all firms in S and (b) is not allowed to transact with any firm outside S. We write $U^*(S)$ to denote the expected utility of following the optimal search policy on firm set S. This is a static comparator: it is defined for a fixed subset of firms, independently of the learning dynamics.

We show that expected consumer utility converges to this comparator, evaluated at $S = S^{\texttt{lnd}}(\mathcal{M})$. More precisely, we consider $\mathbb{E}\left[\, U_i \mid H_{i-1}^{\mathcal{M}} \,\right]$, conditioning on history up to round i.

Lemma 4.3. $\mathbb{E}\left[\, U_i \mid H_{i-1}^{\mathcal{M}} \,\right] \to U^*(S^{\texttt{lnd}}(\mathcal{M}))$ *as $i \to \infty$, almost surely over random history $H^{\mathcal{M}}$.*

Fixing history $H^{\mathcal{M}}$, the comparator $U^*(S^{\texttt{lnd}}(\mathcal{M}))$ is a fixed real value, as is each $\mathbb{E}\left[\, U_i \middle|\, H_{i-1}^{\mathcal{M}} \,\right]$. Thus, Lemma 4.3 says expected consumer utilities almost surely converge in our market for a random history. Moreover, they converge to the utility obtained by a hypothetical consumer who holds the limit posterior beliefs from Proposition 4.1 (which are determined by the learned set $S^{\texttt{lnd}}(\mathcal{M})$).

A complication in the proof of Lemma 4.3 is that the expected (total) utility of a consumer is discontinuous in her beliefs, such as when updated beliefs lead to a change in firm inspection order. To prove Lemma 4.3, we use the convergence of consumer beliefs to the ground truth for learned firms to discipline the evolution of consumer utility and establish convergence in the limit.

Motivated by Lemma 4.3, we write $U(\mathcal{M}) := U^*(S^{\texttt{lnd}}(\mathcal{M}))$ and call it *limit consumer utility* in market $\mathcal{M}$. Like $S^{\texttt{lnd}}(\mathcal{M})$, $U(\mathcal{M})$ is a random variable with randomness coming from the history.

5 The Impact of Improving Consumer Search

We show that reducing the vector $\vec{c}$ of search costs, coordinate-wise, can only increase the learned set and the long-run consumer utility. While intuitive, this is not obvious: fixing a history, cheaper search clearly improves the current consumer's outcomes but it can also alter the dynamics of learning and impact future consumers. In principle this could impede long-run utility, but our proof rules this out.

On a technical level, our comparison considers two markets, with a specific difference between them, and holds almost surely over a *coupling* between their histories.[7] When and if the comparison is between real-valued metrics (stating that something is larger in one market than in the other), the same comparison holds in the sense of first-order stochastic dominance (FOSD).[8]

Theorem 5.1. *Suppose markets $\mathcal{M}, \widetilde{\mathcal{M}}$ are identical except $\widetilde{\mathcal{M}}$ has reduced finite search costs: $\tilde{c}_k \le c_k \; \forall k \le L$. Then for some coupling between their histories, the following holds almost surely:*

- $S^{\texttt{lnd}}(\widetilde{\mathcal{M}}) \supseteq S^{\texttt{lnd}}(\mathcal{M})$. *That is, more firms are learned under history $H^{\mathcal{M}}$ than under $H^{\widetilde{\mathcal{M}}}$, for any joint realization of the two histories.*

[7] A *coupling* between random variables X, Y (taking values in $\mathcal{X}, \mathcal{Y}$) is a joint distribution $\mathcal{D}$ over $\mathcal{X} \times \mathcal{Y}$ with correct marginals. More formally: for a random pair $(X', Y') \sim \mathcal{D}$, the marginal distribution of X' (resp., Y') coincides with that of X (resp., Y). Note that the original random variables X, Y do not need to lie in the same probability space.

[8] First-order stochastic dominance (FOSD) is defined as follows. Consider real-valued random variables X, X' (not necessarily lying in the same probability space). We say $X \ge_{\texttt{FOSD}} X'$ if it holds that $\Pr[X > x] \ge \Pr[X' > x]$ for every x. The relation is strict, denoted $X >_{\texttt{FOSD}} X'$, if additionally $\Pr[X > x] > \Pr[X' > x]$ for some x.

- $U(\widetilde{\mathcal{M}}) \geq U(\mathcal{M})$. *That is, the limit consumer utility is weakly higher under history $H^{\widetilde{\mathcal{M}}}$ than under $H^{\mathcal{M}}$, for any joint realization of the two histories.*

Only finite costs change in Theorem 5.1, so inspection depth and feedback are fixed.

The proof proceeds by coupling histories of $\mathcal{M}$ and $\widetilde{\mathcal{M}}$ on (a) the sequence of consumer values and (b) the outcome when screen k of firm j is revealed for the t'th time, for each k, j, and t.[9] Even under this coupling the resulting histories can generate very different posterior beliefs at any given round i, since different costs can lead to different firms being inspected in each round. Here we use the persistent structure of Algorithm 1: while firms might be inspected in different rounds, Corollary 3.2 implies similar information is revealed in markets $\mathcal{M}$ and $\widetilde{\mathcal{M}}$ on the t'th round that a firm is inspected. Lemma C.2 then implies that the posterior estimates for each firm j evolve similarly in markets $\mathcal{M}$ and $\widetilde{\mathcal{M}}$ *when evaluated only at rounds where those estimates change*. Search costs do not influence this progression, but they do change the condition for a firm being lost. This allows us to couple the set of firms lost across the two histories, which informs long-run consumer utility.

5.1 Search Informativeness and the Importance of Transcripts

The aggregate-feedback model below is a two-screen special case, so Theorem 5.1 also guarantees the benefits of cheaper search without transcripts. We now vary search informativeness rather than cost. For the special case defined below, screen-level feedback ensures weakly better learning and limit utility (Theorem 5.4). With aggregate feedback, both can instead deteriorate (Proposition 5.6). To this end, we consider a simple special case of our model where there are two screens and each firm can only be inspected once.

Definition 5.2. *A* fit-quality market *is a special case with $r = 2$ screens and costs $\vec{c} = (c, \infty)$.*

The inspection depth here is $L = 1$. The consumer can pay c to reveal f_{ij1}, whereas f_{ij2} can only be observed post-transaction. We think of $F_{ij} := f_{ij1}$ as an indicator of *fit* to the particular consumer, with mean $\pi_j := \pi_{j1}$. Whereas $q_{ij} := f_{ij2}$ is a *quality realization* that can only be observed post-transaction. Its mean is $Q_j := \pi_{j2}$, which may represent an intrinsic *quality* of firm j.

We think of this market as one *without* transcripts, in contrast to the following variant *with* transcripts.

Definition 5.3. *A fit-quality market* with transcripts *is special case of our model with $r \geq 2$ screens and cost vector $\vec{c}$ of the form $(c, 0, \ldots, 0, \infty, \ldots, \infty)$.*

Thus, $\vec{c}$ is parameterized by the cost $c \geq 0$ and the number $k \leq r$ of finite entries. Like in fit-quality markets, inspection is one-shot for each firm: pay c to reveal the first k screens. We call c, k, resp., the *cost* and the *informativeness* of the search. Increasing k at fixed cost c moves screens from post-transaction experience to pre-transaction inspection, leaving the underlying screen sequence unchanged. In this single-cost inspection setting, increasing informativeness with transcripts weakly improves market learning and long-run consumer utility.

Theorem 5.4. *The conclusions of Theorem 5.1 apply also to any pair $\mathcal{M}, \widetilde{\mathcal{M}}$ of fit-quality markets with transcripts that are identical except that $\widetilde{\mathcal{M}}$ has increased search informativeness.*

From the perspective of a single consumer with fixed beliefs, a fit-quality market with transcripts is equivalent to a fit-quality market (without transcripts) through a change of notation. One can take $F_{ij} := \prod_{1 \leq \ell \leq k} f_{ij\ell}$ for the information revealed by search, and $q_{ij} := \prod_{k < \ell \leq r} f_{ij\ell}$ for the *quality realization* observed post-transaction. With transcripts, the market observes individual screen outcomes, including which screen first fails; without transcripts, it observes only aggregate inspection and transaction outcomes. Here "why" refers to the failed screen's identity, not a separate causal explanation of consumer preferences. Consumers in both variants update Bayesian beliefs from the observed feedback.

This correspondence also allows us to define search informativeness in markets *without* transcripts. Fix a fit-quality market $\mathcal{M}$ with transcripts and corresponding fit-quality market $\mathcal{M}'$ without transcripts. Increasing k in $\mathcal{M}$ causes, in $\mathcal{M}'$, each fit probability π_j to decrease by some factor $\delta_j \geq 1$ while quality Q_j increases by the same factor δ_j. This motivates the following definition:

[9] Note that even though consumer values are not explicitly represented in a history, they influence the history through the order in which consumers inspect firms. Fixing both the consumer values and the random outcomes of screen inspections for each firm fully specifies each history, defining the coupling.

Definition 5.5. *Suppose $\mathcal{M}$ is a fit-quality market without transcripts, and $\widetilde{\mathcal{M}}$ is an identical market except that, for each firm j, $\tilde{\pi}_j = \pi_j/\delta_j$ and $\tilde{Q}_j = Q_j \times \delta_j$ for some $\delta_j \geq 1$. Then we say that $\widetilde{\mathcal{M}}$ is* more consumer-informative *than $\mathcal{M}$.*

A more consumer-informative market preserves the ex ante probability that the firm's product meets the consumer's need: $\tilde{\pi}_j \tilde{Q}_j = (\pi_j/\delta_j)(Q_j \delta_j) = \pi_j Q_j$. The comparison thus concerns this shift from post-transaction to pre-transaction information, which is a specific form of signal improvement.

Given Theorem 5.4, one might guess that increased consumer-informativeness improves learning and long-run consumer utility. But, as we now show, this is not always the case.

Proposition 5.6. *There exist fit-quality markets $\mathcal{M}, \widetilde{\mathcal{M}}$ without transcripts such that $\widetilde{\mathcal{M}}$ is more consumer-informative than $\mathcal{M}$, and yet it is strictly worse in the long-run: $S^{\texttt{lnd}}(\widetilde{\mathcal{M}}) \subseteq S^{\texttt{lnd}}(\mathcal{M})$ and $U(\widetilde{\mathcal{M}}) \leq U(\mathcal{M})$ almost surely over some coupling between the histories of $\mathcal{M}$ and $\widetilde{\mathcal{M}}$. This inclusion (resp., inequality) is strict with positive probability over the coupling.*

In the example behind Proposition 5.6, a firm's fit probability is either high or low, while its quality is known. More consumer-informative search moves a previously post-transaction screen into inspection, so an aggregate failure no longer identifies a failure of the original fit screen. Pooling these failures weakens the evidence about the firm's fit probability; even a high-fit firm now fails inspection more often and can be prematurely abandoned.

6 Numerical Simulations: a Summary

We complement our theoretical analysis with numerical simulations of market dynamics.[10] We focus on the impact of (i) improved consumer search and (ii) (un)observable transcripts, in the sense of Section 5. Unlike our limit comparisons, these experiments measure outcomes over a finite sequence of consumer arrivals.

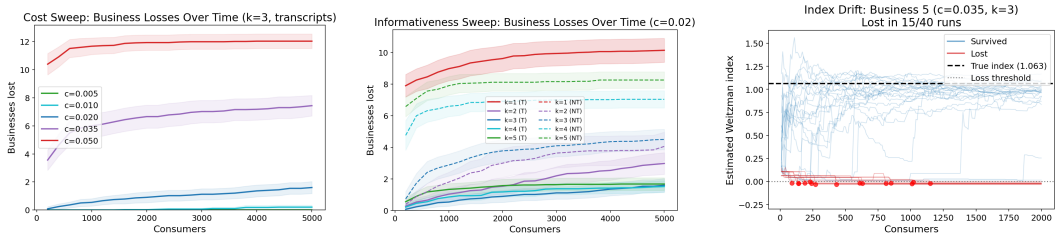


Figure 2: Firm losses in a numerical market simulation. (Left) Losses are high and rapid when costs are high, but accumulate more gradually at moderate costs. (Center) With transcripts (T), losses accumulate more slowly at moderate informativeness, producing a finite-horizon inversion. Without transcripts (NT), losses fall from $k = 1$ to $k = 2$ and then rise with k. (Right) Different trajectories of the index $\sigma_j\left(\hat{\pi}_{ij}^{\texttt{vec}}, \overline{V}\right)$ for a single firm (see Proposition 4.1), visualizing market histories in which the firm is lost (red).

We simulate a stylized market with 15 firms and 5 screens. Each simulation run has 5000 consumers; the firm parameters are redrawn independently from the prior each run.[11] We track several outcomes: consumer utility, firms lost, and belief error. We study how these outcomes change as one varies the underlying parameters of the market. E.g., Figure 2 illustrates the dynamics of firms lost. Cheaper search reduces losses and improves consumer utility in these simulations. In the baseline informativeness sweep, firm losses are non-monotone in k. With transcripts, this inversion reverses in the extended run; without transcripts, it persists through $20{,}000$ arrivals and is also observed across market variants. We also consider perturbations of the prior distributions, prices, and number of firms as robustness checks. All details are in Appendix E.

[10] All simulations were run on a laptop with an Intel Core i7-11370H processor (4 cores, 3.3 GHz) and 32 GB RAM.

[11] In our base simulation, values V_i are drawn from $U[0.5, 2]$; prices P_j are drawn from $\mathcal{N}(0.08, 0.04^2)$ clipped to $[0.01, 0.20]$; each π_{jk} is drawn independently from Beta(4,2) clipped to $[0.25, 0.95]$; and we take $R = 50$ runs.

7 Extensions and Further Directions

Extension: Endogenous Inspection Depth. We can allow each consumer i to strategically choose her inspection depth L_i before the round. This is a two-fold commitment: each firm can be inspected *at most* L_i times, and any firm transacted with must be inspected *exactly* L_i times. We restrict L_i to the range $[L_{\min}, L]$, for some $L_{\min} \in [L]$, so that our main model is a special case with $L_{\min} = L$. The chosen depth cannot vary across firms or be revised in response to inspection outcomes.

The consumer chooses L_i based on her value, beliefs, and inspection costs. For example, a low-value consumer may choose a smaller L_i to avoid an expensive last screen: e.g., inspect a webpage but avoid a phone call. Also, a consumer may choose a smaller L_i if she believes that the last screen(s) are very likely to pass; but future consumers may reconsider this if that belief shifts due to the history.

Our results in Sections 3-4 carry over (i.e., optimal consumer search and learning dynamics). The results in Section 3 need minor modifications: Algorithm 1 is for a chosen $L = L_i$, and Theorem 3.1 asserts optimality over search policies consistent with this depth. Proposition 4.1 and Lemma 4.3 carry over directly, as does the independence of loss probability across firms claimed in Proposition 4.2. We leave open whether/which comparative statics in Proposition 4.2 and Section 5 extend.

The optimal inspection depth L_i can be computed efficiently using the structure of Algorithm 1. It suffices to compute the expected utility of Algorithm 1 for every fixed L. Since the firm order is fixed, this can be done by backward induction. That is, consider firms j in the order of increasing index $\sigma_{ij} > 0$ and compute the expected utility collected by the algorithm conditional on reaching firm j.

For finitely many consumer types observed by the platform, one can extend our results by assuming independent priors over type-specific screen probabilities and maintaining separate posterior beliefs. If each type arrives infinitely often and its conditional model satisfies the baseline assumptions, its arrival subsequence can be analyzed as a separate market.

We assume in our model that screens are inspected in a fixed order. If one alternatively allows the consumer to choose the order of inspection, our analysis would not directly extend: with transcripts, reordering can change which screens are reached before failure and hence what is learned. That said, certain types of reordering, such as optimizing the order in which product features are considered on a website, could be captured in our model as a cost reduction within a single screen (i.e., viewing a detailed product page).

Reordering screens is not generally equivalent to reducing their costs. Any cost-reduction interpretation must preserve screen probabilities and feedback and meet the coordinate-wise comparison of Theorem 5.1. With transcripts, reordering can change which screens are reached before failure and hence what is learned, even when expected inspection cost falls.

Further Directions. Our stylized model can serve as a point of departure (and, hopefully, a basis for future research) in several directions. First, the screening model can be extended, e.g., to accommodate fractional outcomes or more complex aggregation of these outcomes (e.g., reflecting multi-dimensional preferences). Second, the firms can be strategic players, too, adjusting price or quality of their products in response to market incentives. Firms (or the platform) could also subsidize search costs (esp. for AI agents) to improve learning and/or attract customers. Third, richer heterogeneity with unobserved types or cross-type learning remains to be explored. Finally, one could incorporate more open-ended concerns such as consumers' privacy (w.r.t. the feedback they reveal) and misaligned incentives (between consumers/firms and their AI agents).

Acknowledgements and Disclosures

The authors are employees of Microsoft. This work received no third-party funding or support. Beyond their employment at Microsoft, the authors have no competing financial interests to disclose.

References

Daron Acemoglu, Ali Makhdoumi, Azarakhsh Malekian, and Asuman Ozdaglar. Learning From Reviews: The Selection Effect and the Speed of Learning. *Econometrica*, 2022. Working paper available since 2017.

Klaus Adam. Learning while searching for the best alternative. *Journal of Economic Theory*, 101(1): 252–280, 2001.

S Nageeb Ali. Herding with costly information. *Journal of Economic Theory*, 175:713–729, 2018a.

S Nageeb Ali. On the role of responsiveness in rational herds. *Economics Letters*, 163:79–82, 2018b.

Ali Aouad, Jingwei Ji, and Yaron Shaposhnik. The pandora's box problem with sequential inspections. *arXiv preprint arXiv:2507.07508*, 2025.

Susan Athey and Glenn Ellison. Position auctions with consumer search. *The Quarterly Journal of Economics*, 126(3):1213–1270, 2011.

Peter Auer, Nicolò Cesa-Bianchi, and Paul Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine Learning*, 47(2-3):235–256, 2002.

Abhijit V Banerjee. A simple model of herd behavior. *The quarterly journal of economics*, 107(3): 797–817, 1992.

Kiarash Banihashem, MohammadTaghi Hajiaghayi, Suho Shin, and Aleksandrs Slivkins. Bandit social learning under myopic behavior. In *37th Advances in Neural Information Processing Systems (NeurIPS)*, 2023.

Hedyeh Beyhaghi and Linda Cai. Recent developments in pandora's box problem: Variants and applications. *ACM SIGecom Exchanges*, 21(1):20–34, 2024.

Sushil Bikhchandani and Sunil Sharma. Optimal search with learning. *Journal of Economic Dynamics and Control*, 20(1-3):333–359, 1996.

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.

Roberto Burguet and Xavier Vives. Social learning and costly information acquisition. *Economic theory*, 15(1):185–205, 2000.

Lu Fang, Zhe Yuan, Kaifu Zhang, Dante Donati, and Miklos Sarvary. Generative AI and sales productivity: Field experiments in online retail. Working Paper 12201, CESifo, 2026.

Peter Gibbard. A model of search with two stages of information acquisition and additive learning. *Management science*, 68(2):1212–1217, 2022.

John C Gittins. Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 41(2):148–164, 1979.

Benjamin Golub and Evan D. Sadler. Learning in social networks. In Yann Bramoullé, Andrea Galeotti, and Brian Rogers, editors, *The Oxford Handbook of the Economics of Networks*. Oxford University Press, 2016.

Gillian K Hadfield and Andrew Koh. An economy of ai agents. *arXiv preprint arXiv:2509.01063*, 2025.

Kenneth Hendricks, Alan Sorensen, and Thomas Wiseman. Observational learning and demand for search goods. *American Economic Journal: Microeconomics*, 4(1):1–31, 2012.

Nicole Immorlica, Jieming Mao, and Christos Tzamos. Diversity and exploration in social learning. In *The World Wide Web Conference*, pages 762–772, 2019.

Sampath Kannan, Jamie Morgenstern, Aaron Roth, Bo Waggoner, and Zhiwei Steven Wu. A smoothed analysis of the greedy algorithm for the linear contextual bandit problem. In *Advances in Neural Information Processing Systems (NIPS)*, pages 2231–2241, 2018.

Robert Kleinberg, Alexandru Niculescu-Mizil, and Yogeshwer Sharma. Regret bounds for sleeping experts and bandits. In *21st Conf. on Learning Theory (COLT)*, pages 425–436, 2008.

Robert Kleinberg, Bo Waggoner, and E. Glen Weyl. Descending price optimally coordinates search. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, EC, page 23–24. Association for Computing Machinery, 2016. ISBN 9781450339360. doi: 10.1145/2940716.2940760.

Akshit Kumar, Vahideh Manshadi, and Akhilesh Tumu. How much should a conversational recommender system converse? Discussion Paper 2545, Cowles Foundation, Yale University, 2026. URL `https://elischolar.library.yale.edu/cowles-discussion-paper-series/2968/`.

Tor Lattimore and Csaba Szepesvári. *Bandit Algorithms*. Cambridge University Press, Cambridge, UK, 2020.

In Ho Lee. On the convergence of informational cascades. *Journal of Economic theory*, 61(2): 395–411, 1993.

Manuel Mueller-Frank and Mallesh M Pai. Social learning with costly search. *American Economic Journal: Microeconomics*, 8(1):83–109, 2016.

David M Rothschild, Markus Mobius, Jake M Hofman, Eleanor W Dillon, Daniel G Goldstein, Nicole Immorlica, Sonia Jaffe, Brendan Lucier, Aleksandrs Slivkins, and Matthew Vogel. The agentic economy. *arXiv preprint arXiv:2505.15799*, 2025.

Peyman Shahidi, Gili Rusak, Benjamin S Manning, Andrey Fradkin, and John J Horton. The coasean singularity? demand, supply, and market design with ai agents. Technical report, National Bureau of Economic Research, 2025.

Herbert A Simon. Rational choice and the structure of the environment. *Psychological review*, 63(2): 129, 1956.

Aleksandrs Slivkins. Introduction to multi-armed bandits. *Foundations and Trends® in Machine Learning*, 12(1-2):1–286, November 2019. Published with *Now Publishers* (Boston, MA, USA). Also available at `https://arxiv.org/abs/1904.07272`.

Lones Smith and Peter Sørensen. Pathological outcomes of observational learning. *Econometrica*, 68 (2):371–398, 2000.

Martin Weitzman. *Optimal search for the best alternative*, volume 78. Department of Energy, 1978.

Peter Whittle. Multi-armed bandits and the Gittins index. *J. Royal Statistical Society, Series B*, 42(2): 143–149, 1980.

Se Yan, Han Zhong, Zemin (Zachary) Zhong, and Wenyu Zhou. Shopping with a platform AI assistant: Who adopts, when in the journey, and what for. *arXiv preprint arXiv:2603.24947*, 2026.

A Technical Connection to Pandora's Box Problem and Multi-Armed Bandits

Let us spell out the technical connection between our model and these well-studied problems.

The optimal-search problem faced by a single consumer is related to Weitzman's *Pandora's box problem* [Weitzman, 1978]. Specifically, a special case of our model with a single inspectable screen (i.e., $L = 1$) is also a special case of this problem. Here, firms j correspond to "boxes". The boxes are inspected via the same process as described above, then one of the inspected boxes can be "opened" and its reward is realized as an independent random draw from some known distribution. For us, the reward distribution for box j is given by the current Bayesian posterior on "quality", i.e., the product of uninspectable screen probabilities. Within the scope of Pandora's box problem, we focus on the special case when the inspection outcome is binary. Moreover, we posit *obligatory inspection* (a common assumption that a box must be inspected before it is opened; see Beyhaghi and Cai [2024] for a survey of problem variants with and without this assumption) and an *outside option* (not opening any box is also a possibility). An optimal solution, for general inspection outcomes, is given by the celebrated Weitzman's algorithm [Weitzman, 1978].

Collectively, the consumers face a sequential learning problem with a notion of "bandit feedback": only the outcomes of consumer's actions (i.e., inspections and transaction) are revealed to future consumers, whereas the counterfactuals are not. Hence, the consumers collectively face a trade-off between exploration and exploitation, of which multi-armed bandits is a fundamental model.

Simple (and, in some sense, degenerate) cases of our model correspond to bandit problems studied in prior work. In a bandit problem, one "pulls" some arm in each round, observing a random reward drawn from some fixed but unknown distribution specific to this arm [Auer et al., 2002]. Now, firms correspond to "arms". Consider a variant of our model when all inspectable screens are known (so there's no point in inspecting any firm) and some screens are uninspectable. Moreover, assume all prices are $P_j \equiv 0$ and all values are $V_i \equiv 1$. If all fits are inspectable screens are $f_{ijk} = 1$, this model variant corresponds to the bandit problem described above, where reward of each arm j is the product of its uninspectable screens. [12] For general screens f_{ijk}, we obtain the *sleeping bandits* problem [Kleinberg et al., 2008], where the set of available arms changes from one round to another. For general values V_i (and going back to $f_{ijk} \equiv 1$) we have a special case of *contextual bandits*: a bandit problem in which a payoff-relevant *context* is revealed before each round.[13]

Of course, the crucial difference of our model compared to these bandit problems is that each consumer can "inspect" arms before "pulling" them (let alone inspect them multiple times).

While in bandit problems one typically designs an algorithm (which trades off exploration and exploitation in the long run), consumer behavior in our model is "greedy", optimizing for a particular round. Projecting this back to bandits, it corresponds to the "greedy" bandit algorithm: one that chooses the posterior-best arm in each round.

B Proof of Theorem 3.1: Optimal Consumer Search

We will actually consider a slightly generalized class of problem instances and prove optimality of a correspondingly generalized form of Algorithm 1. In our generalized problem, each firm j could have a firm-specific number of screens r_j, the consumer can choose a different inspection depth L_j for each firm, inspection costs are firm-specific and denoted c_{jk}, and the outside option value of not transacting with any firm is $Z \geq 0$. The original problem is a special case where the values r_j are identical across firms, as are the values L_j and the costs c_{jk} for each k, and the outside option value is $Z = 0$. We claim that the optimal policy for this more general class of problems is Algorithm 2, which is simply Algorithm 1 with the appropriate firm-specific quantities r_j, L_j, and c_{jk} used in place of r, L, and c_k respectively, and value Z used in place of 0 as the index threshold at which to stop inspecting firms.

We prove optimality of Algorithm 2 in two steps. First, we restrict attention to policies where, if the consumer inspects one firm j and then later inspects another firm $j' \neq j$, the consumer cannot return to firm j to inspect further or transact. We call such policies *persistent*.

[12] Notably, Bernoulli rewards is a paradigmatic special case in bandits (along with Gaussian rewards).

[13] Contextual bandit problems come in many variants [e.g., see books Slivkins, 2019, Lattimore and Szepesvári, 2020]. We have a rather simple special case, where the context is V_i and mean rewards are given by the transaction utility.

ALGORITHM 2: Optimal search policy at firm-specific inspection depths L_j and outside option Z

for *each firm* j **do**

$$\sigma_{ij} := V_i \cdot \prod_{L_j < \ell \leq r_j} \hat{\pi}_{j\ell} - P_j - \frac{\sum_{\ell \leq L_j} c_{j\ell} \prod_{\ell' < \ell} \hat{\pi}_{j\ell'}}{\prod_{\ell \leq L_j} \hat{\pi}_{j\ell}}.$$

for *each firm* j *with* $\sigma_{ij} > Z$*, in order of decreasing index* σ_{ij}**, do**
  for *each* $\ell = 1, 2, \ldots, L_j$ **do**
    Pay $c_{j\ell}$ to inspect $f_{ij\ell}$;
    If $f_{ij\ell} = 0$, break and continue to the next firm.
  Transact with firm j and leave the market *// We must have* $f_{ij\ell} = 1$ *for all* $\ell \in [L_j]$

Leave the market without transacting *// No firms with index exceeding* Z *were a fit*

Definition B.1. *We say that a search policy is* persistent *if, for every firm* j *and in every history, the set of inspection actions for firm* j *occur on contiguous steps, and any transaction with firm* j *occurs immediately following an inspection of firm* j*.*

Note that Algorithm 2 is a persistent policy. We first show that Algorithm 2 is optimal over the class of persistent policies.

Lemma B.2. *Given any market* $\mathcal{M}$*, history* $H_{i-1}^{\mathcal{M}}$*, and value* V_i *for consumer* i*, Algorithm 2 generates at least as much expected utility as any persistent policy.*

Proof. First observe that, without loss of generality, we can limit attention to persistent policies that stop inspecting a firm j after observing that some f_{ijk} is equal to 0, since any further inspection of firm j is suboptimal. We can also limit attention to policies that, after inspecting f_{ij1}, continue to inspect firm j until observing a 0 or until screen f_{ijL_j} has been inspected. This is because it is suboptimal to inspect a firm j unless there is at least one possible realization of outcomes in which firm j is transacted with, and this can only occur if L_j screens of firm j are inspected.

Under these assumptions, a persistent policy is determined entirely by the (possibly adaptive) order in which firms are inspected. Given this observation, there is a natural reduction from the problem of finding an optimal persistent policy to the Weitzman Pandora's Box problem. Indeed, given the choice to begin inspection of firm j, the process of inspection will either fail (a 0 is observed on some screen) or succeed (all L_j inspected screens are equal to 1). Write $\tilde{\pi}_j$ for the probability of passing all inspections for firm j from 1 to L_j: $\tilde{\pi}_j = \prod_{1 \leq \ell \leq L_j} \hat{\pi}_{j\ell}$. If they pass and the transaction occurs, the expected value from the transaction is $\tilde{V}_i := V_i \cdot \prod_{L_j < \ell \leq r_j} \hat{\pi}_{j\ell}$. And the expected inspection cost is the total sum of costs from each step ℓ between 1 and L_j, which is $c_{j\ell}$ times the probability that inspection step ℓ is reached. Call this total inspection cost $\tilde{c}_j$. Then the process of inspecting firm j is equivalent in expectation to a single search operation of cost $\tilde{c}_j$ that reveals a transaction utility of $\tilde{V}_i - P_j$ with probability $\tilde{\pi}_j$, otherwise $-P_j$.

A persistent policy can perform this inspection operation on firms, in any (possibly adaptive) order over firms. The optimal such policy can therefore be described via Weitzman's algorithm [Weitzman, 1978], a known optimal solution to the Pandora's Box problem. Weitzman's algorithm inspects firms according to an index σ_{ij}, in the same sense as in Algorithm 2, but for a more generic notion of index. Namely (fixing the history $H_{i-1}^{\mathcal{M}}$), the index σ_{ij} is the solution to

$$\tilde{c}_j = \mathbb{E}\left[\left(U_{ij} - \sigma_{ij} \right)^+ \mid H_{i-1}^{\mathcal{M}} \right],$$

where U_{ij} is the transaction utility and the expectation is over the random outcome of inspection. Since U_{ij} is $\tilde{V}_i - P_j$ with probability $\tilde{\pi}_j$ and otherwise 0,[14] this expression simplifies to

$$\tilde{c}_j = \tilde{\pi}_j \left(\tilde{V}_i - P_j - \sigma_{ij} \right).$$

[14] If firm j is not a fit then the transaction utility, if a transaction with firm j were forced, is $-P_j$. But since we assume tie-breaking in favor of not transacting, the consumer would never choose to make such a transaction. So we can take the box reward to be 0 in this case. That said, all subsequent analysis would still hold without change if we defined $U_{ij} = -P_j$ in the case of a non-fit.

Rearranging yields $\sigma_{ij} = \tilde{V}_i - P_j - \tilde{c}_j/\tilde{\pi}_j$. Substituting our expressions for $\tilde{V}_i$, $\tilde{c}_j$, and $\tilde{\pi}_j$ yields precisely the index rule from Algorithm 2.

We are not done yet, since our definition of persistent policies requires also that the policy transact with firm j immediately after observing the first L_j screens to be 1. In general the Weitzman search algorithm has no such constraint: it continues to search until the maximum value observed (or the outside option value of Z) exceeds the index of all remaining firms, then takes the best value revealed so far (or the outside option). We note, however, that in our setting a successful search of firm j (i.e., where the first L_j screens are all observed to be 1) always reveals a transaction value weakly higher than the index σ_{ij} for firm j, and hence higher than the index of all remaining firms. Thus, upon a successful inspection of firm j up to screen L_j, the Weitzman optimal search policy will always immediately terminate and transact with firm j, as required of a persistent policy.

We conclude that Algorithm 2 is exactly the Weitzman optimal search policy for the corresponding Pandora's Box problem instance described above, and is therefore optimal over persistent policies. □

We are now ready to prove the optimality of Algorithm 2 over all (possibly non-persistent) policies. We will proceed by induction on $x := \sum_j L_j$, the total number of inspectable screens over all firms.

For the base case, if $x = 1$ then there is only a single firm j. This means that any policy must be persistent, and hence by Lemma B.2 Algorithm 2 is an optimal policy.

Likewise, if $x > 1$ and $n = 1$, then there is again only one available firm j and every policy must be persistent by definition. Thus, once again, Algorithm 2 must be optimal by Lemma B.2.

Now suppose $n \geq 2$ and hence $x \geq 2$. To prove the existence of an optimal policy that is persistent, we will show that no violation of the persistence conditions occurs along different potential execution outcomes of the policy. To this end, consider an optimal policy and suppose it begins by inspecting screen f_{ij1} of firm j. If that inspection reveals that $f_{ij1} = 0$, then (without loss of generality) firm j will not be explored further nor transacted with. We are therefore left with a reduced problem instance with one fewer firms and hence a reduced x. By induction, Algorithm 2 is an optimal policy for this downstream subproblem. Since Algorithm 2 is persistent for the subproblem, and the single inspection of firm j trivially occurred contiguously, we conclude that the overall policy is also persistent on this path of its execution.

If, alternatively, the inspection of firm j reveals $f_{ij1} = 1$, we are left with one of two possible reduced problem instances depending on L_j. If $L_j = 1$, meaning that we have now performed all necessary inspections to transact with firm j, this reduced subproblem corresponds to removing firm j and increasing the outside option Z to the expected value of transacting with firm j. Taking the outside option in this subproblem corresponds to transacting with firm j. Alternatively, if $L_j > 1$, then firm j is still present in the reduced problem instance but has one fewer screens; that is, we lower L_j and r_j by one and perform a change of variables $c_{jk} \leftarrow c_{j(k+1)}$ and $\pi_{jk} \leftarrow \pi_{j(k+1)}$ so that the remaining screens are indexed from 1 to $r_j - 1$. Induction then implies that, in both cases, Algorithm 2 is optimal for this downstream subproblem.

Since Algorithm 2 is persistent, the only potential violation of persistence over all steps (including the initial inspection of firm j) is if firm j is neither immediately transacted with (if $L_j = 1$) nor immediately inspected again (if $L_j > 1$). Suppose this violation of persistence occurs, as otherwise we are done.

Note that, from the definition of Algorithm 2, the order in which other firms (besides j) are searched, after the first inspection of j, follows an index rule and is therefore independent of whether the first inspection of firm j revealed $f_{ij\ell} = 0$ or $f_{ij\ell} = 1$. Given this, we claim that our supposedly optimal policy can be weakly improved in a manner that recovers persistence. The improvement is to delay the first inspection of j until just prior to the step in which the second inspection of j occurs (if $L_j > 1$) or the step in which the outside option is taken (if $L_j = 1$) in the resulting inspection order conditional on $f_{ij1} = 1$. This modified policy is well-defined, since the subsequent ordering (up to the next inspection of firm j, if any) follows an index rule and therefore is not determined by the result of the inspection of firm j.

The reason this modified policy is an improvement is that the outcome of the initial inspection of j does not influence which firms are inspected, and in what order, up until the next step in which firm j would be inspected or transacted with. Indeed, if the initial inspection of j reveals $f_{ij1} = 0$, then

firm j is removed from consideration and the remaining firms are simply considered in index order. If instead the initial inspection of j reveals $f_{ij1} = 1$, then the only change is that firm j remains in the consideration set and would be inserted into the inspection order according to its index (or as an outside option). This is where we use the inductive hypothesis that our policy uses an index rule: since the indices of other firms are not affected by this change, all other firms are searched in the same order up to the point at which firm j is inserted. In either case, following the initial inspection of firm j, the subsequent policy makes the same inspections, in the same order, up until the next time that firm j is inspected or transacted with. We emphasize that an observation that $f_{ij1} = 0$ does not terminate the search procedure, it only removes further inspection of firm j from consideration.

As a result, the only payoff impact of delaying the inspection of firm j is that, in the event that the policy does not reach the step in which firm j is to be inspected further or transacted with, the consumer saves a payment of c_{j1} from inspecting j the first time, which is a weak improvement.

We conclude that, in all cases, the optimal policy performs no better than a persistent policy. Thus, by Lemma B.2, our Algorithm 2 is optimal for all problem instances. This completes the proof of Theorem 3.1.

C Proofs from Section 4: Feedback and Learning

C.1 Projection of a History onto a firm

The first step in proving Proposition 4.1 is Lemma C.2 below. It states that the posterior beliefs $(\mathcal{P}_{ijk})$ depend only on the screen realizations from firm j in $H_{i-1}^{\mathcal{M}}$, and not any other information in the history.

Definition C.1. *For a history realization H, the* projection of H onto firm j*, written $H|_j$, contains all the screen realizations for firm j present in H. The projection $H|_j$ preserves the order in which these data appear in H, but not the specific rounds from H in which they were observed.*

We can think of $H|_j$ as r sequences of binary variables: one for each level of screen. Element t of each sequence corresponds to the tth time that screen was revealed in the history, either through inspection or post-transaction feedback.

Lemma C.2. *For any finite history realization $H_{i-1}^{\mathcal{M}} = H$ of market $\mathcal{M}$ and any firm j, posterior beliefs $\mathcal{P}_{ijk}(H)$ are determined by $H|_j$.*

Proof. Fix any partial history realization $H_{i-1}^{\mathcal{M}} = H$ and suppose consumer i with value V_i arrives in round i. The consumer's choice of which firms to inspect, and in which order, is fully determined by c, V_i, and by $(\mathcal{P}_{ijk}(H_{i-1}))_{jk}$, so in particular does not reveal any further information about π_{jk} beyond what is in H_{i-1}. The updated posterior beliefs $\hat{\pi}_{jk}(H_i)$ are therefore determined entirely by the previous beliefs plus the inspection outcomes and post-transaction feedback. As only the inspection outcomes and post-transaction feedback for firm j are correlated with π_{jk}, the updated posteriors are determined entirely by information contained in $H|_j$. □

C.2 Completing the Proof of Proposition 4.1

Proof. Consider the following procedure for generating an infinite history. We first construct, for each firm j and screen $k \in [r]$ in advance, an infinite sequence Ω_{jk} of independent Bernoulli signals of mean π_{jk}. Then, whenever a screen realization is needed during the course of infinite history H, it is pulled from the corresponding pre-determined sequence Ω_{jk}. Note then that, for each j, the projected history $H|_j$ consists of prefixes of the streams $\Omega_{j1}, \ldots, \Omega_{jr}$.

With probability 1, the empirical average of each (infinite) stream Ω_{jk} converges to its true mean π_{jk} in the limit over its sequence, so in particular a posterior estimate of the mean will converge to π_{jk} over infinitely many observations. We will condition on this event, which is determined entirely by the sequences Ω_{jk} and occurs with probability 1.

Since all screen probabilities are bounded away from 0, any firm that is inspected infinitely often will be transacted with infinitely often and will therefore generate, as part of the history, a signal from each of the r screens infinitely often. Thus, by Lemma C.2, for any firm j that is inspected infinitely often, with probability 1 the posterior beliefs converge to the means from which these signals are

drawn. That is, $\mathcal{P}_{ijk}$ converges to π_{jk} in the limit as $i \to \infty$ for each k. This implies that any firm that is inspected infinitely often is learned.

It remains to consider firms that are inspected only finitely often; say A is the set of such firms. If $j \in A$ then there is a time after which j is no longer selected in the optimal consumer search process. Let T be the maximum such time over all $j \in A$. We claim that $\mathcal{P}_{ijk}(H) = \mathcal{P}_{ijk}(H_T)$ for all $k \in [r]$, all $j \in A$ and all $i \geq T$. This follows because each such j is not searched at any time $t > T$, so Lemma C.2 implies that $\mathcal{P}_{ijk}(H_{i-1}) = \mathcal{P}_{ijk}(H_T)$ for all $i > T$, and hence the limits exist and equal the claimed quantities.

Finally, since $\pi_{jk} < 1$ for all j and k, any firm j for which $\sigma_j((\hat{\pi}_{ijk}), \overline{V}) > 0$ would be inspected by consumer i with positive probability bounded away from 0. As $(\hat{\pi}_{ijk})$ does not change until firm j is inspected, this implies that firm j will be inspected in some round at or after round i with probability 1. Thus, if $j \in A$, then there must exist some i for which $\sigma_j((\hat{\pi}_{ijk}), \overline{V}) \leq 0$, as otherwise j is inspected infinitely often. This implies that every firm $j \in A$ is lost, as required. □

C.3 Proof of Proposition 4.2

Proof. For any market $\mathcal{M}$ and firm j, let $H|_j^{\mathcal{M}}$ denote the projection of $H^{\mathcal{M}}$ onto firm j. Then the marginal distribution over $H|_j^{\mathcal{M}}$, the projection of history $H^{\mathcal{M}}$ onto firm j, has an especially simple form. Each inspection/feedback outcome in $H|_j^{\mathcal{M}}$ is an independent draw from a Bernoulli distribution of mean π_{jk}, for some k. In particular, $H|_j^{\mathcal{M}}$ and $H|_{j'}^{\mathcal{M}}$ are independent for any two firms j and j'.

Let $(\hat{\pi}_{jtk})$ denote the posterior mean estimates held by a consumer given the first t entries of $H|_j^{\mathcal{M}}$. Note that this is not the same as the estimates held by consumer t. Let ω_j be the probability that there exists any t such that $\sigma_{ij}((\hat{\pi}_{jtk}), \overline{V}) \leq 0$. Then, by Lemma C.2, ω_j is precisely the probability that j is lost in history $H^{\mathcal{M}}$. And since ω_j depends only on the marginal distribution over $H|_j^{\mathcal{M}}$, which are independent across firms, we conclude that each firm is lost independently with probability ω_j.

Finally, recall that $\sigma_{ij}((\hat{\pi}_{jtk}), \overline{V})$ is weakly increasing in $\overline{V}$ and each $\hat{\pi}_{jtk}$, and weakly decreasing in c. Since firm j is lost when this quantity becomes negative, we therefore have that the loss probability ω_j is weakly *decreasing* in $\overline{V}$ and each $\hat{\pi}_{jtk}$, and weakly *increasing* in c. Since random variables $\hat{\pi}_{jtk}$ are positively correlated with π_{jk}, respectively, we have that ω_j is weakly decreasing in π_{jk} as well. □

C.4 Proof of Lemma 4.3

Proof. We will condition on the event of Proposition 4.1 that every firm that is inspected infinitely often is learned and all others are lost, which occurs with probability 1. For each lost firm $j \notin S^{\texttt{lnd}}$, there is some finite time at which the firm is lost. As there are only finitely many firms, there is a fixed time t_0 after which no further firms are lost. We will restrict attention to consumers who arrive to the market following round t_0, so that only the firms in $S^{\texttt{lnd}}$ are in the consideration set of each consumer.[15]

Let $U(V; \mathcal{M}, H)$ denote the optimal utility obtained by a consumer with value V in market $\mathcal{M}$, following history H. Define

$$\tilde{U}(V, H) := \mathbb{E}[U(V; \mathcal{M}, H) \mid H]$$

to be the expected utility obtained by a consumer with value V given prior history H, evaluated with respect to the consumer's posterior beliefs over $\mathcal{M}$ given H. In other words, $\tilde{U}(V, H)$ is the consumer's expectation about their utility given their belief about $\mathcal{M}$.

Recall that Algorithm 1 optimizes expected utility over all search orders (i.e., permutations) of the firms. Moreover, for any given search order, the expected utility obtained by consumer i given a past history H is continuous in $\hat{\pi}_{jk}(H)$ for all j and k. Since $\tilde{U}(V, H)$ is a maximum over the utilities obtained from a finite set of search orders, we conclude that $\tilde{U}(V, H)$ is continuous in $(\hat{\pi}_{jk}(H))_j$ as

[15] Note that fixing the value of t_0 and the set $S^{\texttt{lnd}}$ imposes conditions on the progression of the market dynamics. However, the only property of the stochastic market dynamics that we will use in the remainder of the proof is the assumed convergence of posterior estimates to true means, which we have already conditioned upon.

well. Thus, since $\lim_{i\to\infty} \hat{\pi}_{ijk} = \pi_{jk}$ for all $j \in S^{\texttt{lnd}}$ by Proposition 4.1, and since each consumer that arrives after round t_0 only inspects firms in $S^{\texttt{lnd}}(H^{\mathcal{M}})$, we conclude that

$$\lim_{i\to\infty} \tilde{U}(V, H^{\mathcal{M}}_{i-1}) = U^*(S^{\texttt{lnd}}).$$

In other words, the consumer beliefs about their anticipated utility converge over time to the static comparator $U^*(S^{\texttt{lnd}})$.

It remains to relate these consumer beliefs to the true expected utility obtained by the consumer. Convergence of posterior mean estimates implies that $\lim_{i\to\infty} |\hat{\pi}_{ijk} - \pi_{jk}| = 0$. Continuity of $\tilde{U}(V, H)$ in $\hat{\pi}_{jk}(H)$ then implies

$$\lim_{i\to\infty} |\tilde{U}(V, H^{\mathcal{M}}_{i-1}) - U(V, \mathcal{M}, H^{\mathcal{M}}_{i-1})| = 0.$$

A standard sandwich argument therefore implies

$$\lim_{i\to\infty} U(V, \mathcal{M}, H^{\mathcal{M}}_{i-1}) = \lim_{i\to\infty} \tilde{U}(V, H^{\mathcal{M}}_{i-1}) = U^*(S^{\texttt{lnd}})$$

as required. □

D Proofs from Section 5: Improved Consumer Search

D.1 Proof of Theorem 5.1

Proof. We will couple the histories of $\mathcal{M}$ and $\widetilde{\mathcal{M}}$ as follows. First, we draw the infinite sequence of consumer values $V_1, V_2, \ldots$, which will be used for both histories. Second, for each firm j and screen $k \in [r]$, we independently draw an infinite sequence of screen realizations, each an independent Bernoulli random variable with mean π_{jk}. Write Ω_{jk} for each such sequence, and Ω_j for the collection of r value sequences for firm j.

To generate history $H^{\mathcal{M}}$, each consumer i will have value V_i and will engage in their consumer optimal search process. Whenever screen k of a firm j is inspected, say for the t'th time, the outcome is taken to be the t'th element of the sequence of screen realizations from Ω_{jk}. This fully determines the progression of history $H^{\mathcal{M}}$.

Write $H|_j^{\mathcal{M}}$ for the projection of history $H^{\mathcal{M}}$ onto firm j. By construction, $H|_j^{\mathcal{M}}$ consists of prefixes of the sequences in Ω_j. By Proposition 4.1, almost surely each firm is either learned or lost, and the projected history $H|_j^{\mathcal{M}}$ is finite (and a strict prefix of Ω_j) if firm j is lost, but infinite (and equal to Ω_j) if firm j is learned.

We generate history $H^{\widetilde{\mathcal{M}}}$ in the same way, noting that the consumer optimal search process may resolve differently due to the modified search costs. Nevertheless, it is still the case that $H|_j^{\widetilde{\mathcal{M}}}$ is a prefix of Ω_j, and equal to Ω_j if firm j is learned.

It will now be convenient to talk about posterior estimates for a firm j evaluated after a certain number of searches of that firm have occurred. For a given $z \geq 1$, consider the zth round in which firm j is inspected at least once. A feature of Algorithm 1 is that the set of signals generated for firm j is fully determined by Ω_j: a realization for each screen will be revealed, in sequence, until the first 0 is encountered (if any). In particular, the set of signals revealed is independent of the round index i, the consumer value V_i, the depth of inspection L, and whether or not a transaction occurs with firm j. In particular, this means that the projected histories $H|_j^{\mathcal{M}}$ and $H|_j^{\widetilde{\mathcal{M}}}$ will agree on the information revealed about firm j during the first z rounds in which firm j was inspected.

We now write $\hat{\pi}_{jk}(H, z)$ to denote $\hat{\pi}_{jk}(H_t)$ where t is the round in which firm j is searched for the zth time in history H. Then, since $H|_j^{\mathcal{M}}$ and $H|_j^{\widetilde{\mathcal{M}}}$ agree up to the minimum of their lengths, by Lemma C.2 we have that $\hat{\pi}_{jk}(H^{\mathcal{M}}, z) = \hat{\pi}_{jk}(H^{\widetilde{\mathcal{M}}}, z)$ for all z such that firm j is searched at least z times in each history.

By Proposition 4.1, firm j is lost in market $\widetilde{\mathcal{M}}$ if and only if there exists z such that

$$\tilde{\sigma}_{ij}((\hat{\pi}_{jk}(H^{\widetilde{\mathcal{M}}}, z)), \overline{V}) \leq 0. \tag{1}$$

But now, considering whether j is lost in history $H^{\mathcal{M}}$ of market $\mathcal{M}$, note that if firm j is searched at least z times in $H^{\mathcal{M}}$ then (1) would imply

$$\begin{aligned}\sigma_{ij}((\hat{\pi}_{jk}(H^{\mathcal{M}}, z)), \overline{V}) &= \sigma_{ij}((\hat{\pi}_{jk}(H^{\widetilde{\mathcal{M}}}, z)), \overline{V}) \\ &\leq \tilde{\sigma}_{ij}((\hat{\pi}_{jk}(H^{\widetilde{\mathcal{M}}}, z)), \overline{V}) \\ &\leq 0\end{aligned}$$

where the first equality follows because $\hat{\pi}_{jk}(H^{\widetilde{\mathcal{M}}}, z) = \hat{\pi}_{jk}(H^{\mathcal{M}}, z)$ for each j and k from our coupling, the second inequality follows because costs are pointwise lower in $\widetilde{\mathcal{M}}$ and this can only increase the index (recalling that the depth L of search is fixed and identical in $\mathcal{M}$ and $\widetilde{\mathcal{M}}$), and the third inequality is (1). Thus, if firm j is lost in history $H^{\widetilde{\mathcal{M}}}$, then firm j is lost in history $H^{\mathcal{M}}$ as well. We conclude that $S^{\texttt{lnd}}(\tilde{M}) \supseteq S^{\texttt{lnd}}(\mathcal{M})$ under our coupling.

Finally, since $S^{\texttt{lnd}}(\widetilde{\mathcal{M}}) \supseteq S^{\texttt{lnd}}(\mathcal{M})$ and $\tilde{c}_k \leq c_k$ for all k, each consumer performing optimal search in the steady-state of $\widetilde{\mathcal{M}}$ has only a larger set of learned firms to inspect, and a reduced cost of each such search. Thus, any chosen search order for learned firms in $\mathcal{M}$ could have been selected in $\widetilde{\mathcal{M}}$ as well at a weakly improved expected utility. Optimizing over all potential search orders, we conclude that the long-run expected consumer utility in $\widetilde{\mathcal{M}}$ under history $H^{\widetilde{\mathcal{M}}}$ is therefore only greater than that of $\mathcal{M}$ under $H^{\mathcal{M}}$, yielding $U(\widetilde{\mathcal{M}}) \geq U(\mathcal{M})$ over our coupling. □

D.2 Proof of Theorem 5.4

Proof. Suppose $\mathcal{M}$ has inspection depth k and $\widetilde{\mathcal{M}}$ has depth $\tilde{k} > k$. We will use the screen-stream coupling from the proof of Theorem 5.1, with the same sequence of consumer values. Whenever a firm is inspected, inspection and transaction feedback together reveal its screens up to the first failure (or all r passes), regardless of inspection depth. Thus, after the same number of visits to a firm, both markets have observed the same screen outcomes.

Let $p_{j\ell}(z)$ be the posterior mean for screen ℓ after the first $z \geq 0$ such visits, including the prior at $z = 0$, and put

$$a_j(z) := \prod_{\ell \leq k} p_{j\ell}(z), \qquad b_j(z) := \prod_{k < \ell \leq \tilde{k}} p_{j\ell}(z).$$

These posteriors are defined from the streams even if that prefix is never reached in one market. The indices at value $\overline{V}$ and the two inspection depths satisfy

$$\begin{aligned}\sigma_j^{(k)}(z) &= \overline{V} \prod_{\ell > k} p_{j\ell}(z) - P_j - \frac{c}{a_j(z)}, \\ \sigma_j^{(\tilde{k})}(z) &= \overline{V} \prod_{\ell > \tilde{k}} p_{j\ell}(z) - P_j - \frac{c}{a_j(z) b_j(z)}.\end{aligned}$$

Consequently,

$$\sigma_j^{(k)}(z) = b_j(z) \sigma_j^{(\tilde{k})}(z) - (1 - b_j(z)) P_j. \tag{2}$$

Since $P_j \geq 0$ and $0 < b_j(z) \leq 1$, a non-positive index at depth $\tilde{k}$ implies one at depth k. Firms whose parameters are all known a priori are learned in both markets. Otherwise, on the probability-one event of Proposition 4.1, if j is learned in $\mathcal{M}$, then every finite prefix is reached and $\sigma_j^{(k)}(z) > 0$ at every prefix. Equation (2) therefore forces $\sigma_j^{(\tilde{k})}(z) > 0$ at every prefix, so j cannot be lost in $\widetilde{\mathcal{M}}$ and must be learned there. Thus $S^{\texttt{lnd}}(\widetilde{\mathcal{M}}) \supseteq S^{\texttt{lnd}}(\mathcal{M})$ under this coupling.

For the utility comparison, fix a consumer value V and an optimal policy on the learned firms of $\mathcal{M}$. We consider the following (possibly suboptimal) policy for $\widetilde{\mathcal{M}}$. In $\widetilde{\mathcal{M}}$, ignore any additional learned firms and emulate the policy from $\mathcal{M}$ until it would transact with a firm j. First reveal the additional screens $k+1, \ldots, \tilde{k}$, at zero additional cost. If they all pass, make the same transaction; otherwise leave without transacting. This is a feasible policy at depth $\tilde{k}$. Coupling the consumer's screen outcomes, the two policies incur the same search costs. If an additional screen fails, the

original transaction would not have met the consumer's need, so avoiding its price $P_j \geq 0$ weakly improves realized utility; otherwise utility is unchanged. Optimizing in $\widetilde{\mathcal{M}}$ and averaging over V therefore gives $U(\widetilde{\mathcal{M}}) \geq U(\mathcal{M})$ on the same history coupling. □

D.3 Proof of Proposition 5.6

Proof. Take a single firm with price $P = 0$, consumers with value $V = 1$, and search cost $c = 1/5$. In $\mathcal{M}$, quality is known to be $Q = 1/2$, and the fit probability π is either $1/100$ or $99/100$ with equal prior probability. Fix the market realization $\pi = 99/100$. In $\widetilde{\mathcal{M}}$, rescale fit by $2/3$ and quality by $3/2$: quality is known to be $\tilde{Q} = 3/4$, the fit prior has equally likely values $1/150$ and $33/50$, and the realized fit probability is $33/50$. All screen probabilities are strictly between 0 and 1, and $\widetilde{\mathcal{M}}$ is more consumer-informative with $\delta = 3/2$. One can view this as moving a known $2/3$-pass screen into inspection, leaving a $3/4$-pass quality screen.

In $\mathcal{M}$, after s fit successes and f fit failures, the likelihood ratio of high to low fit is 99^{s-f}, so the posterior mean is

$$\hat{\pi}(d) = \frac{1}{100} + \frac{98}{100}\frac{99^d}{1+99^d}, \qquad d = s - f.$$

The index $1/2 - c/\hat{\pi}(d)$ is positive exactly when $d \geq 0$. Quality feedback does not affect beliefs because Q is known. Thus loss occurs when the success-minus-failure walk first hits -1. In the realized market, this walk steps up with probability $99/100$ and down with probability $1/100$, so its probability of ever hitting -1 is $1/99$.[16] The probability of learning the firm is therefore $p = 98/99$.

In $\widetilde{\mathcal{M}}$, the initial fit mean is $1/3$, giving positive index $3/4 - c/(1/3) = 3/20$. If the first inspection fails, Bayes' rule gives

$$\Pr(\tilde{\pi} = 33/50 \mid \text{failure}) = 51/200,$$
$$\mathbb{E}[\tilde{\pi} \mid \text{failure}] = 2599/15000 < 4/15.$$

The posterior mean is below $c/\tilde{Q} = 4/15$, so the index becomes negative and the firm is lost. In the realized market, the first inspection fails with probability $17/50$. Hence the probability $\tilde{p}$ of learning the firm satisfies $\tilde{p} \leq 33/50 < 98/99 = p$.

If the firm is learned, the optimal limiting policy inspects it, and its expected utility in either market is

$$u_* = \frac{99}{100}\frac{1}{2} - \frac{1}{5} = \frac{33}{50}\frac{3}{4} - \frac{1}{5} = \frac{59}{200} > 0.$$

If the firm is lost, limit utility is zero. Thus $U(\mathcal{M}) = u_* \mathbf{1}\{1 \in S^{\mathtt{lnd}}(\mathcal{M})\}$, and analogously for $\widetilde{\mathcal{M}}$.

Finally, draw $Z \sim \mathrm{Uniform}[0,1]$, prescribing learning in $\mathcal{M}$ when $Z \leq p$ and in $\widetilde{\mathcal{M}}$ when $Z \leq \tilde{p}$. Conditional on these indicators, sample a history from each market's correct conditional distribution. This gives a coupling of histories with the correct marginals. Since $\tilde{p} < p$, it satisfies $S^{\mathtt{lnd}}(\widetilde{\mathcal{M}}) \subseteq S^{\mathtt{lnd}}(\mathcal{M})$ and $U(\widetilde{\mathcal{M}}) \leq U(\mathcal{M})$ almost surely. Both comparisons are strict on $\{\tilde{p} < Z \leq p\}$, which has positive probability $p - \tilde{p}$. □

E Numerical Simulations

We complement our theoretical analysis with numerical simulations of market dynamics.

We simulate a market with $m = 15$ firms and $r = 5$ screens per firm. Probabilities π_{jk} are drawn independently from $\mathrm{Beta}(4,2)$, clipped to $[0.25, 0.95]$, yielding a skewed distribution with mean approximately 0.67. Platform posterior distributions are initialized with independent (unclipped) Beta priors.[17] Prices P_j are drawn from $\mathcal{N}(0.08, 0.04^2)$ clipped to $[0.01, 0.20]$. Consumer values V_i are drawn i.i.d. from $\mathrm{Uniform}(0.5, 2.0)$, giving $\bar{V} = 2.0$. These parameters were chosen to generate sufficient variability in the resulting market; we test some alternative choices at the end of this section.

[16] For $\rho = 1/99$, stopping the martingale ρ^d when the walk hits -1 or N gives the lower-boundary hitting probability $(1-\rho^N)/(\rho^{-1}-\rho^N)$; letting $N \to \infty$ gives ρ.

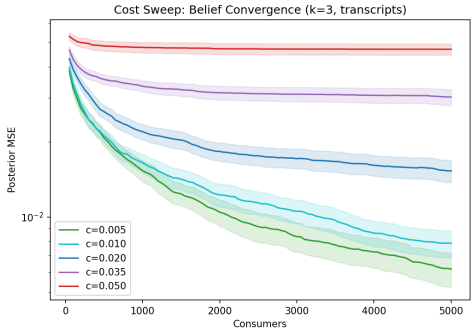
[17] This maintains tractability of Bayesian posterior estimates. This is consistent with our theoretical model, recalling that our results hold pointwise over market realizations (see Remark 2.2).

Each simulation run has $T = 5000$ consumers unless otherwise specified. We report averages over $R = 50$ independent runs (with firm realizations and other sources of randomness redrawn in each run). All shaded regions and error bars in plots represent 95% confidence intervals on the means.

We refer to firms not lost by T as *surviving*; the plot label "learned" denotes this finite-horizon count, not verified convergence. Utility trajectories show windowed realized averages, whereas panels labeled "steady-state utility" report terminal welfare estimates, not demonstrated limit outcomes.

Impact of Search Cost. We study a fit-quality model with transcripts as in Definition 5.3, fix the search informativeness to $k = 3$ and vary the search cost $c \in \{0.005, 0.01, 0.02, 0.035, 0.05\}$. See Figure 3 for the resulting dynamics and Figure 4 for final outcomes at the end of the simulation. Lower costs reduce belief error and firm losses and raise consumer utility over the simulated horizon. At low costs ($c \leq 0.01$), nearly all firms survive to $T = 5{,}000$, and belief error is small. As costs increase beyond $c \approx 0.02$, incorrect firm losses rise sharply. At $c = 0.05$, most firms are lost.

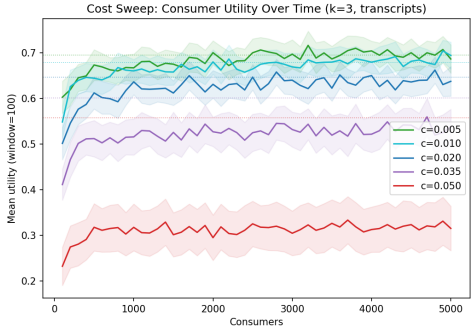
We distinguish between *correct* loss (where the index evaluated at the true screen probabilities is truly negative) versus *incorrect* loss (where the true index is positive and hence the loss is due to inaccurate posterior estimates). In all settings where losses occur, the majority of the observed losses are incorrect, driven by unlucky streaks of screen failures in initial inspections. As shown in Figure 4, higher costs lead to more correct losses (orange) but substantially more incorrect losses (red).



(a) Mean squared error of belief estimates.



(b) Cumulative firms lost over time.



(c) Consumer utility trajectory (window = 100).

Figure 3: Cost sweep dynamics. (a) MSE declines rapidly for low costs but stalls when firms are lost. (b) firm losses accumulate early and plateau; higher costs yield more losses. (c) Windowed realized utility rises early and then fluctuates around a plateau. Horizontal lines show expected optimal consumer utility with known screen probabilities and no firms lost, at each cost level.

An illustrating example for incorrect loss, showing the evolution of the index σ_j for the maximum consumer type $\overline{V}$ (i.e., $\sigma_j\left(\hat{\pi}_{ij}^{\mathsf{vec}}, \overline{V}\right)$) for a single firm j over a sequence of 50 runs, is given in Figure 5. The firm is lost for those traces where this index falls below zero, indicated in red. Surviving traces approach the true index over the plotted horizon. This plot uses the same realization of fit probabilities for all runs, so that the true index remains constant.

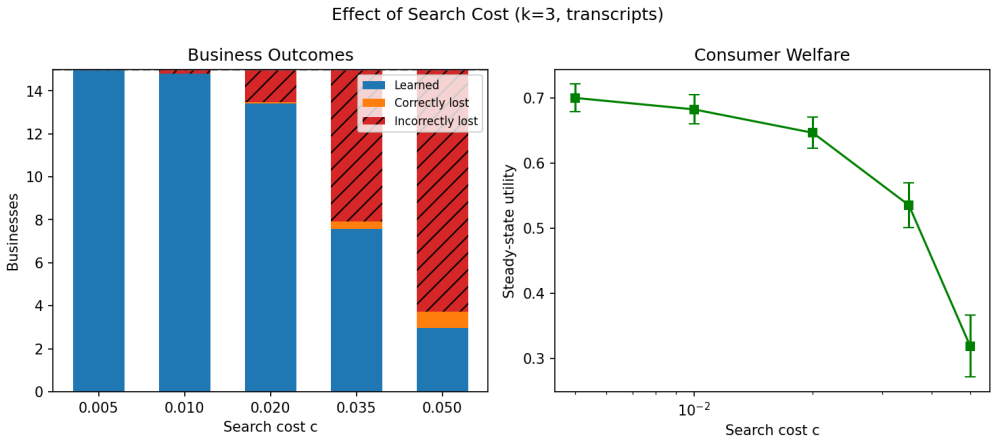


Figure 4: Cost sweep outcomes at $T = 5{,}000$: surviving firms and losses (left), and terminal welfare estimates (right). Losses are "correct" if the index evaluated at the true screen probabilities is negative, and "incorrect" if it is positive.

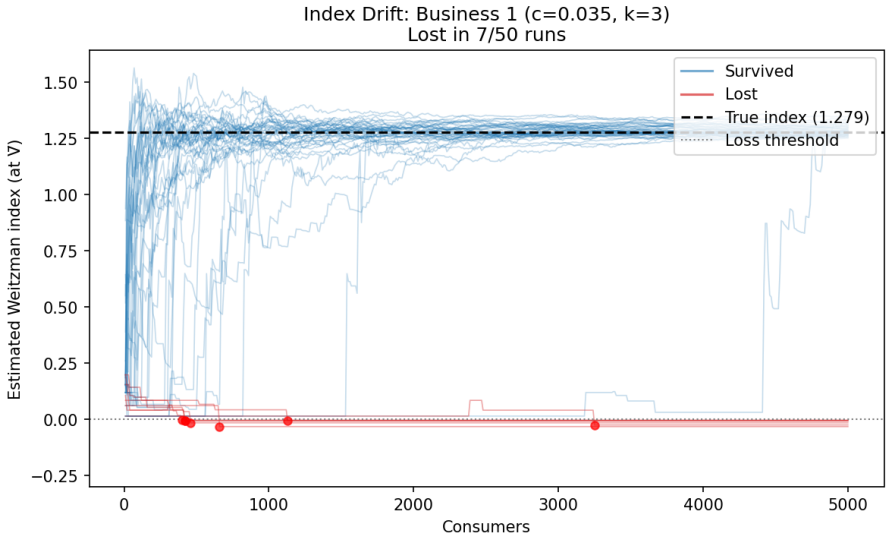
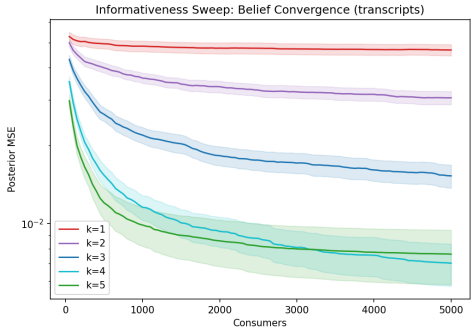


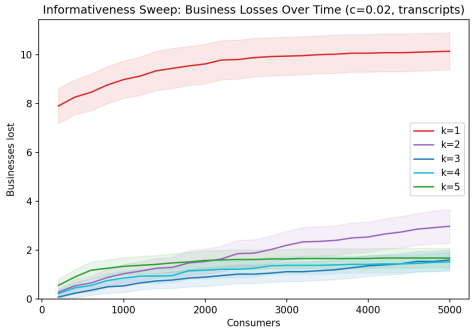
Figure 5: Surviving and lost trajectories for an example firm. Each trace plots a trajectory of $\sigma_j\left(\hat{\pi}_{ij}^{\text{vec}}, \overline{V}\right)$ (see Proposition 4.1). All runs use the same realizations of π_{jk} for all j and k. Red traces indicate runs where the business has been lost by round 5000.

Impact of Search Informativeness. We next fix cost $c = 0.02$ and vary informativeness $k \in \{1, 2, 3, 4, 5\}$. See Figure 6 for the resulting dynamics and Figure 7 for the final outcomes at the end of the simulation. Terminal welfare estimates rise from 0.31 at $k = 1$ to 0.92 at $k = 5$. In contrast, the number of surviving firms after $5{,}000$ rounds is non-monotone, peaking at $k = 4$ and declining slightly at $k = 5$. At $T = 2000$ this non-monotonicity is more pronounced, with the average number of lost firms minimized at $k = 3$ (0.90 lost) but higher for $k = 4$ and $k = 5$ (1.18 and 1.58 lost, respectively).

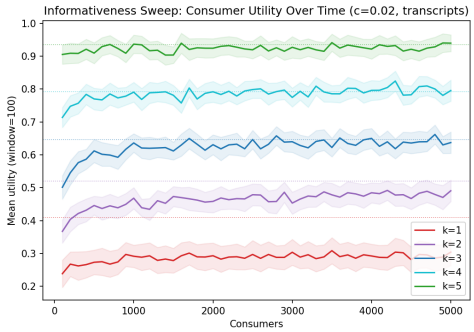
This finite-horizon non-monotonicity does not contradict Theorem 5.4, which compares limit outcomes. To investigate it, we fixed a market realization and examined an individual firm lost less frequently at $k = 3$ than at $k = 5$ (Figure 8). The plot covers $T = 2{,}000$ rounds: most surviving traces at $k = 5$ are close to the true index, whereas several at $k = 3$ remain near the loss threshold. This suggests a transient effect: losses accumulate more slowly under less informative search. Belief error generally falls faster at higher k (Figure 6(a)); we conjecture that the slower accumulation of losses reflects reduced sensitivity of the index to posterior changes when k is low.



(a) Mean squared error of belief estimates.



(b) Cumulative firms lost over time.



(c) Consumer utility trajectory (window = 100).

Figure 6: Informativeness sweep dynamics. (a) Belief error generally falls faster at higher k. (b) Loss trajectories show qualitatively different temporal patterns across k values: high-k losses occur early and plateau, while low-k losses accumulate gradually throughout the horizon. (c) Higher k yields higher windowed realized utility over the plotted horizon. Horizontal lines show expected optimal consumer utility with known screen probabilities and no firms lost, at each informativeness level.

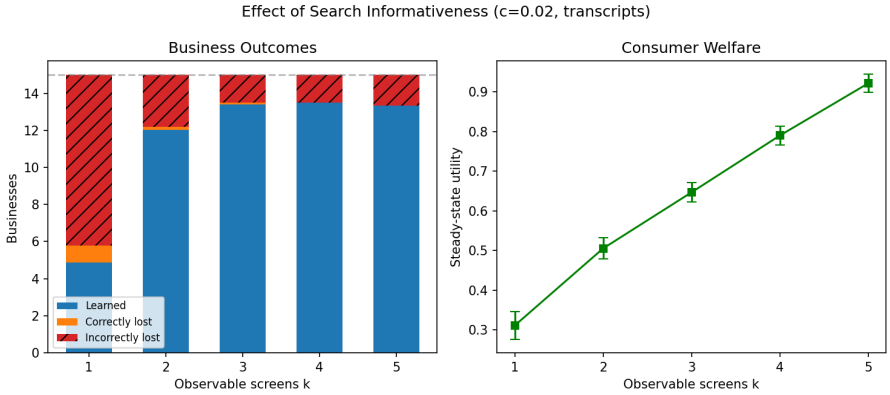


Figure 7: Informativeness sweep outcomes at $T = 5{,}000$: surviving firms and losses (left), and terminal welfare estimates (right).

To investigate this finite-horizon explanation, we compare $k = 3$ and $k = 5$ over $T = 20{,}000$ rounds (Figure 9). The average loss count for $k = 3$ overtakes that for $k = 5$, reversing the earlier ordering.

Impact of Transcripts. Finally, we compare the transcript (T) and no-transcript (NT) information structures across levels of search informativeness. Recall that the NT scenario corresponds to a fit-quality market in which the platform only observes screen pass/fail outcomes collectively, rather than for each screen separately, maintaining an aggregate posterior over the success probability for each firm. For each fixed k, this aggregate posterior is maintained via Bayesian updating on the

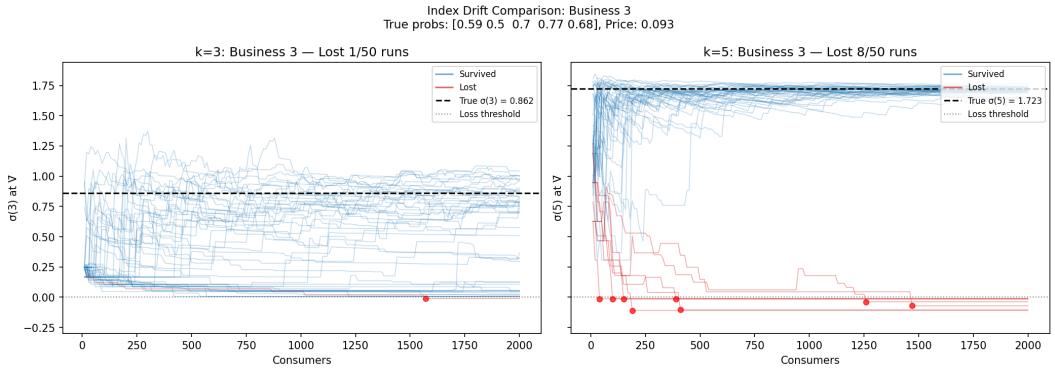


Figure 8: Weitzman index drift traces for a sample firm at $k = 3$ (left panel) and $k = 5$ (right panel). Blue traces have not been lost by $T = 2{,}000$; red traces are lost, with red circles marking the crossing point. At $k = 5$, losses occur early and most surviving traces approach the true index rapidly. At $k = 3$, several remain near the loss threshold, suggesting the possibility of further losses beyond the plotted horizon.

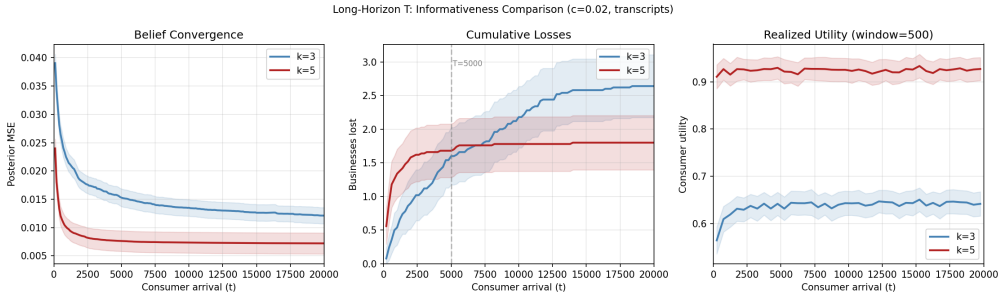


Figure 9: Extended transcript comparison ($T = 20{,}000$). Posterior MSE is shown on the left. The average loss count for $k = 3$ overtakes that for $k = 5$ (center); realized utility (right) uses a 500-consumer window.

Bernoulli signals received from the collective pass/fail outcomes.[18] See Figure 10. At $k = 5$ and $T = 5{,}000$, transcripts leave 13.3 surviving firms on average versus 6.7 under NT, with terminal welfare estimates of 0.92 versus 0.76.

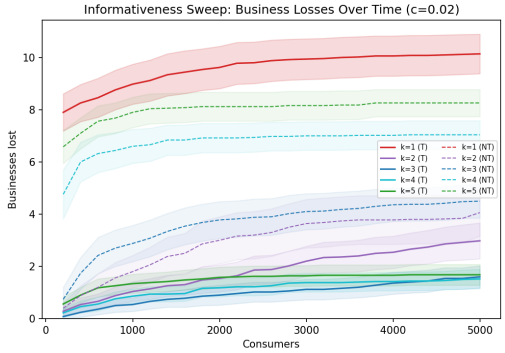
Why do transcripts help? Under NT, the aggregate signal "all k screens passed" has probability $\prod_{\ell=1}^{k} \pi_{j\ell}$, which for typical firms is quite low (approximately 0.67^k). At $k = 5$, this is roughly 0.13: the platform observes a success only 13% of the time, receiving a failure signal in the remaining 87% of visits that is uninformative about *which* screen failed. The aggregate posterior therefore converges slowly and with high variance. As aggregate successes become rarer at higher k, the gap between feedback regimes grows in this sweep. Surviving-firm counts fall from 10.5 at $k = 3$ to 6.7 at $k = 5$ under NT, while T remains roughly constant (13.4 to 13.3).

We also extend the NT simulation to $T = 20{,}000$ (Figure 11). Unlike the transcript comparison, the ordering of losses across $k = 3, 4, 5$ persists through this extended horizon.

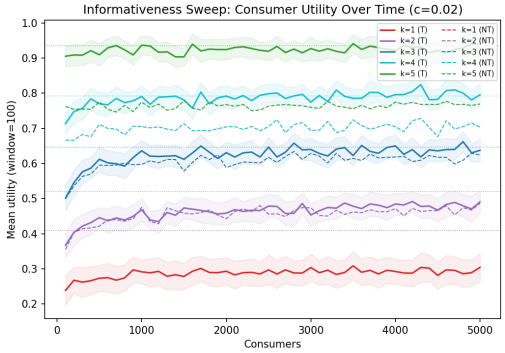
Robustness Checks We test the cost-sweep gains, utility gains from informativeness with transcripts, and non-monotone NT losses under five market variations. Each variant modifies one aspect of the baseline parameterization while holding all others fixed.

1. **More businesses** ($m = 30$)**:** Doubles the number of firms, reducing per-business observation frequency.
2. **Fewer businesses** ($m = 5$)**:** Reduces competition, concentrating observations.

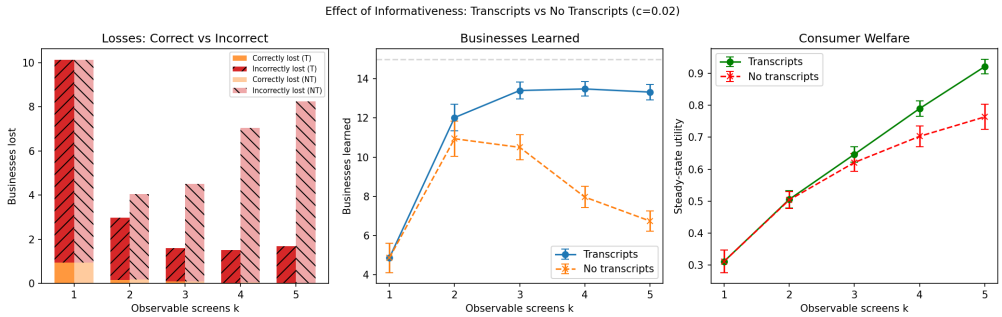
[18] To align with our analytical comparisons (Theorem 5.4 and Proposition 5.6), we keep the same initialization of the platform posterior estimates as we vary k, so that only consumer informativeness changes across treatments.



(a) firms lost over time (T and NT).



(b) Consumer utility trajectory (T and NT).



(c) Terminal survival, loss, and welfare estimates ($T = 5{,}000$).

Figure 10: Informativeness sweep dynamics comparing transcript (solid) and no-transcript (dashed) modes. (a) NT produces substantially more firm losses, particularly at higher k. (b) Both modes show windowed realized utility increasing in k, with an advantage for T at larger k.

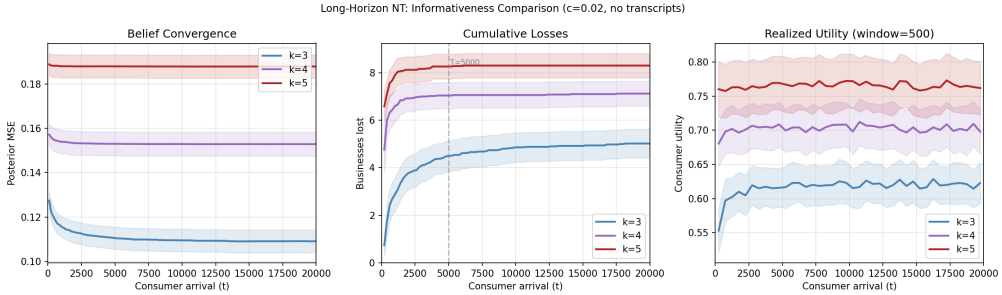


Figure 11: Extended no-transcript comparison ($T = 20{,}000$, $k = 3, 4, 5$). Posterior MSE (left) and cumulative losses (center) remain higher at higher k through this horizon; realized utility (right) uses a 500-consumer window.

3. **Symmetric screen probs (Beta(2,2)):** Replaces the skewed Beta(4,2) with a symmetric distribution (mean 0.5).
4. **Homogeneous values ($V = 1.25$):** Eliminates value heterogeneity by fixing V at the population mean.
5. **Higher prices ($P_j \sim \mathcal{N}(0.15, 0.04^2)$):** Nearly doubles average prices, squeezing transaction margins.

Terminal welfare estimates and surviving-firm counts are non-increasing in cost in all five variants (Figure 12). In the informativeness sweep (Figure 13), welfare estimates increase in all five variants, while survivor counts are non-monotone in two ($m = 30$ and $m = 5$). These finite-horizon inversions

are consistent with the transient pattern in the baseline experiment. The larger reversal at $m = 30$ may reflect fewer observations per firm; the $m = 5$ difference is small relative to the plotted uncertainty.

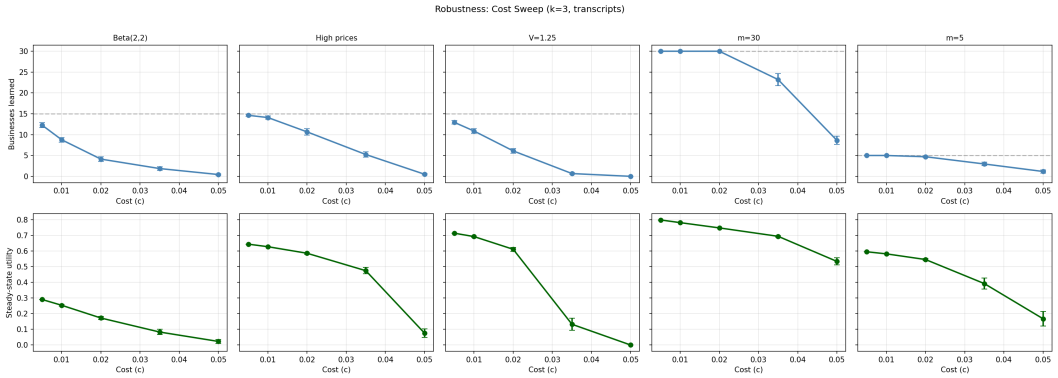


Figure 12: Robustness of cost sweep results across five market variants ($k = 3$, transcripts, $T = 5{,}000$). Top: surviving firms; bottom: terminal welfare estimates. Both are non-increasing in cost in all variants.

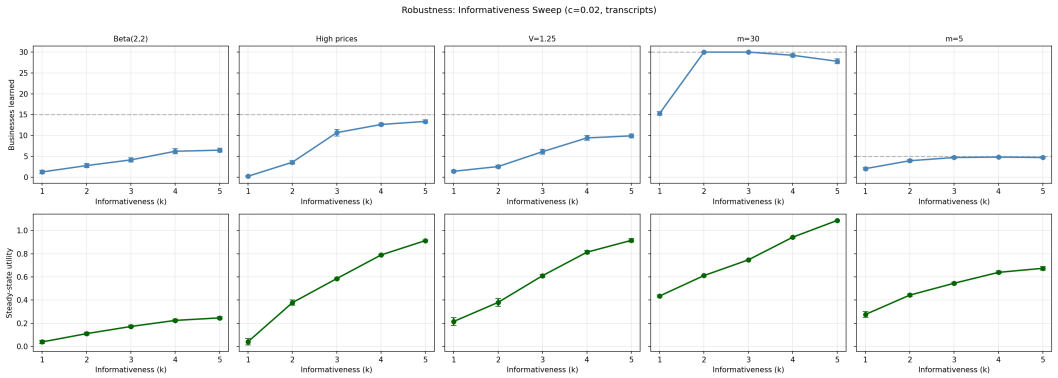


Figure 13: Robustness of informativeness sweep results across five market variants ($c = 0.02$, transcripts, $T = 5{,}000$). Terminal welfare estimates (bottom) increase in all variants. Surviving-firm counts (top) are monotone in three variants but decline at high k for $m = 30$ and $m = 5$.

We separately verify the robustness of the finding that businesses lost is non-monotone in consumer informativeness in the absence of transcripts. We run the informativeness sweep ($c = 0.02$, $k \in \{1, \ldots, 5\}$) in NT mode for each of the five variants above. See Figure 14. We find that all five variants exhibit a U-shaped relationship between consumer informativeness and business losses in the NT case. In four of the five scenarios this non-monotonicity is statistically significant at $p < 0.05$, but it falls within sampling noise for the variant with higher prices. Welfare estimates are also non-monotone in some NT variants (bottom row), unlike the baseline sweep.

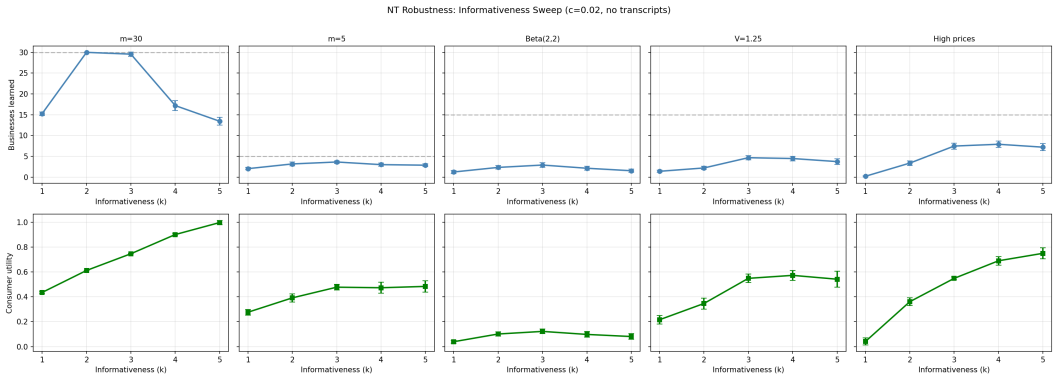


Figure 14: Robustness of informativeness sweep results without transcripts ($c = 0.02$, $T = 5{,}000$). Surviving-firm counts (top) first increase and then decrease in all five variants, most sharply for $m = 30$. Terminal welfare estimates (bottom) need not be monotone in k.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract summarizes the model and results, and the introduction elaborates on the motivation and connection to the application domain. All claims made in the abstract are supported by results in the paper.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: As a primarily theory paper, our model makes a number of simplifying assumptions and abstractions. We include a Model Interpretation discussion in the model section that describes limiting model choices, and a discussion of possible future extensions (many of which are relaxations of these assumptions) at the end of the paper.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: Full proofs of all results are included in the supplementary appendix.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [Yes]

Justification: A full description of the experiments is provided in Appendix E.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes]

   Justification: All code used in numerical simulations is included with the supplementary material.

   Guidelines:

   - The answer [N/A] means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: The experiments in the paper are numerical simulations that did not require explicit training or hyperparameter optimization. The parameters of the simulation, as specified in the main body and in the appendix, were chosen to illustrate variation in outcomes.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: Error bars and mean confidence intervals are provided on all relevant plots.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.

- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: All experiments were run on a laptop with an Intel Core i7-11370H processor (4 cores, 3.3 GHz) and 32 GB RAM. This information is included in the experiment description.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: We do not anticipate human or social harms as a result of the research conducted in this paper.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

Justification: In the introduction we motivate our research through the lens of potential social benefits and harms of agentic marketplaces. Our results speak to both positive and negative social impacts, and our comparison vis-a-vis transparency of search outcomes speaks to potential mitigation of possible negative outcomes.

Guidelines:

- The answer [N/A] means that there is no societal impact of the work performed.
- If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

Answer: [N/A]

Justification: No data or models are being released as part of this submission.

Guidelines:

- The answer [N/A] means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [N/A]

Justification: The paper does not use existing assets.

Guidelines:

- The answer [N/A] means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.

- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: The paper does not release new assets.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: This paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: This paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

- The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A]

    Justification: LLMs were used for editing and formatting purposes, and as a programming aid during the preparation of the numerical simulation code and result plots. LLMs were also used to review the details of completed proofs, which informed further edits. LLMs were not involved as any important, original, or non-standard component to the core research methodology.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.