Journal of Experimental and Computational Methods in Applied Sciences, pp. 1–15

Quantifying Feedback-Loop Bias in Two-Sided Marketplaces: The Role of Ratings, Reviews, and Reputation Mechanisms

Rafael Sousa [1], Diego Mendes [2]

1. *Universidade Federal do Pampa (UNIPAMPA), 275 Avenida Antônio Trilha, Centro, São Gabriel 97300-000, Brazil*
2. *Instituto Federal do Espírito Santo (IFES), 1200 Avenida Vitória, Jucutuquara, Vitória 29040-780, Brazil*

Abstract

Two-sided marketplaces commonly rely on ratings, reviews, and reputation scores to reduce information asymmetry between participants. These signals are rarely passive summaries; they are often inputs to ranking, search, recommendation, and eligibility rules that shape who is seen, who matches, and who transacts. When reputational signals affect exposure and exposure affects the future stream of signals, a feedback loop arises in which early stochasticity can be amplified into persistent disparities. This paper formalizes feedback-loop bias as the discrepancy between outcomes that would arise under exposure that is conditionally independent of signal noise and outcomes under platform policies that couple exposure to noisy reputational states. A continuous-time stochastic model links true latent quality, transaction intensity, rating generation, review text informativeness, and platform ranking functions. The analysis decomposes bias into components attributable to selection on unobservables, endogenous sampling of raters, and nonlinear score aggregation. Identification strategies are developed for observational and experimental data, including randomized perturbations to ranking, instrumental variation in exposure, and panel approaches with seller fixed effects and shrinkage priors. Numerical methods are proposed for solving the induced distributional dynamics, including diffusion approximations and finite element discretizations of the associated Fokker–Planck equations. The framework yields estimands for amplification, persistence, and welfare-relevant distortion, and it supports mechanism design objectives that trade off efficiency, exploration, and disparity under explicit constraints.

Introduction

Two-sided marketplaces mediate interactions between at least two groups whose participation decisions are mutually dependent, such as buyers and sellers, riders and drivers, hosts and guests, or clients and service providers [1]. A central operational challenge is information: participants typically cannot directly observe the latent quality of the counterparty before transacting, and the platform cannot perfectly verify quality at scale. Ratings, written reviews, and derived reputation scores are widely used as a partial remedy. These mechanisms can improve matching efficiency by aggregating dispersed private experiences into public signals that guide future decisions. At the same time, most modern platforms do not treat reputational signals as neutral summaries; they embed them into search rankings, recommendations, eligibility thresholds, pricing modifiers, and trust badges. This embedding is often justified by user experience and safety goals, but it changes the statistical properties of the signals themselves because the process generating future observations depends on past observed values.

The central object of this paper is a class of distortions that arise when reputational signals are both outputs of marketplace interactions and inputs to the rules that determine which interactions occur. When exposure to buyers increases with a seller's reputation, the seller receives more transactions and therefore more ratings, which in turn refine or inflate reputation. Conversely, low exposure reduces the opportunity to accumulate informative feedback, which can lock sellers into a low-information state where random early negatives dominate. This coupling generates a feedback loop. The term feedback-loop bias is used here to denote systematic discrepancies between the distribution of realized reputations, matches, and welfare outcomes in the coupled system and the distribution that would obtain under an alternative system that breaks, weakens, or controls the coupling between exposure and reputational noise.

The distortion is not merely that some sellers are exposed more than others; rather, it is that exposure becomes endogenously correlated with errors in measured reputation, with consequences for inference and for fairness-relevant allocation.

A common intuition is that if ratings are unbiased estimators of quality, then sorting by ratings should asymptotically sort by quality. This intuition can fail even when individual ratings are unbiased conditional on quality, because the set of ratings observed is itself selected by the exposure policy. Suppose the platform ranks sellers by an estimated reputation that is a nonlinear function of past ratings, while buyers are more likely to transact with highly ranked sellers. The sellers who receive more ratings have reputations that are estimated with lower variance, and if the ranking function is convex in the estimate, the platform may amplify variance differences into mean differences by Jensen-type effects [2]. In addition, buyers who select into certain sellers may systematically differ in their propensity to rate or in their rating scale, generating endogenous rater composition. Written reviews add another layer: review propensity and textual sentiment can depend on buyer expectations, which are influenced by displayed reputation, producing expectation-driven measurement shifts. These mechanisms imply that the reputational signal is not a simple noisy measurement channel but an equilibrium object jointly determined by platform policy and participant behavior.

This paper develops a technical framework that keeps the two-sided nature explicit. On the demand side, buyers arrive stochastically, observe a subset of sellers through platform exposure mechanisms, and choose whether to transact based on displayed reputations, prices, and idiosyncratic tastes. On the supply side, sellers choose participation, effort, and potentially strategic behavior such as review solicitation. The platform chooses how to aggregate ratings and reviews into reputations and how to map reputations into exposure and eligibility. A key modeling goal is to separate latent quality from observed reputation, while allowing reputation to influence both matching intensity and the distribution of raters. This separation enables a precise definition of bias as a property of the joint process rather than as a purely statistical artifact.

The analysis is motivated by practical questions that arise when platforms or regulators attempt to evaluate rating systems. If one observes that high-reputation sellers earn more revenue, that observation alone does not imply efficient sorting; it may reflect amplification of early luck. If one observes that certain demographic groups have lower reputations, one must distinguish true quality differences, measurement differences in ratings conditional on quality, and feedback-loop differences due to exposure policies that reduce learning for some groups. Similarly, when researchers attempt to estimate the causal effect of reputation on outcomes, naive regressions can be confounded because reputation and exposure coevolve. Identifying causal effects requires either exogenous variation in displayed reputation or carefully modeled selection dynamics.

The paper makes several contributions [3]. First, it proposes a formal definition of feedback-loop bias grounded in counterfactual exposure processes. The definition treats exposure as a policy-controlled stochastic intensity and characterizes bias as the difference between coupled and decoupled steady states or finite-horizon distributions. Second, it provides a decomposition of the bias into components attributable to endogenous sampling, nonlinear aggregation, and strategic response. Third, it provides identification strategies and estimators that can be implemented with platform logs, including methods that combine randomized ranking perturbations with hierarchical modeling of quality. Fourth, it develops numerical methods for computing the distributional dynamics of reputation under alternative policies, including diffusion approximations and finite element solutions to Fokker–Planck equations describing the evolution of reputational states. Fifth, it analyzes mechanism design objectives for reputation and ranking rules under constraints, making explicit the trade-offs among efficiency, exploration, and disparity.

The modeling stance is deliberately technical and neutral. The goal is not to argue that all reputation systems are harmful or beneficial, but to provide analytical tools for quantifying the conditions under which feedback loops materially distort outcomes and for comparing policy alternatives. In some parameter regimes, coupling exposure to reputation accelerates learning and improves matching; in others, it creates persistent misallocation and unequal opportunity. A framework that yields measurable estimands and computational procedures can support empirical evaluation [4].

A final motivation concerns measurement and communication. Platforms often summarize complex feedback into a single scalar such as a star rating, sometimes rounded to a half-star. This quantization interacts with ranking thresholds, producing discontinuities where a small change in score yields a large change in exposure [5]. Additionally, users often interpret rating differences nonlinearly; the perceived difference between 4.6 and 4.8 stars may exceed that between 4.0 and 4.2. Modeling these nonlinearities is important because they affect both buyer choice and the platform's own use of the score [6]. The paper therefore emphasizes link functions and nonlinear transformations that map latent quality to displayed reputation.

Model of Two-Sided Interaction and Reputation Generation

Consider a marketplace with a finite or large population of sellers indexed by $i \in \{1, \dots, N\}$. Each seller has a latent, time-invariant quality parameter $q_i \in \mathbb{R}$ that summarizes the expected utility contribution to a buyer net of price in a normalized scale. Quality may represent reliability, service

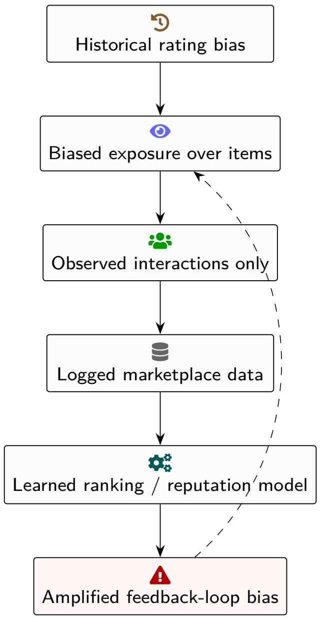

Figure 1: Stylized causal feedback diagram for ratings- and reputation-driven marketplaces. Initial historical biases in ratings affect exposure, restricting which buyer–seller pairs interact and are observed. The resulting logged data feed learning algorithms that update ranking and reputation models, which then adjust exposure yet again. Quantifying the amplification from historical bias to steady-state bias is central to measuring feedback-loop effects.

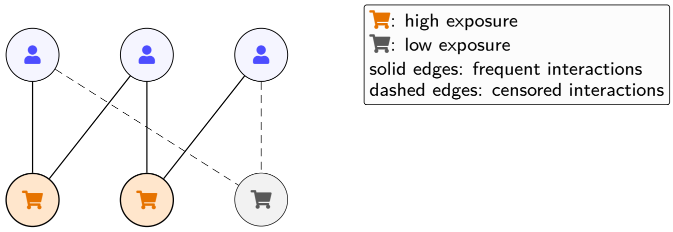

Figure 2: Bipartite view of a two-sided marketplace with biased exposure. A small subset of sellers attracts most of the traffic and feedback, whereas low-exposure sellers accumulate very few ratings and reviews. Because learning and evaluation rely primarily on frequently interacted edges, the resulting models may underestimate quality or opportunity for under-exposed sellers, contributing to feedback-loop bias.

level, or match-specific value. Buyers arrive over time; the arrival process can be modeled as a nonhomogeneous Poisson process with intensity $\Lambda(t)$, or as a discrete-time sequence of visits. For technical clarity, adopt continuous time $t \geq 0$ and let buyer arrivals be a Poisson process with intensity $\lambda_B(t)$. Upon arrival, a buyer is exposed to a subset of sellers through the platform's exposure policy. Let $E_i(t)$ denote the exposure rate of seller i at time t, interpreted as the probability that a randomly arriving buyer views seller i conditional on arriving, or as a normalized share of impressions.

Exposure depends on reputational state and possibly other covariates. Let $R_i(t)$ denote the platform's internal reputation state, which may include the history of ratings, reviews, cancellations, and other signals. The displayed reputation to buyers is $S_i(t) = g(R_i(t))$, where g may include smoothing, rounding, or clipping. The platform's exposure policy is represented as a mapping $E_i(t) = \pi_i(S(t), X(t))$, where $S(t)$ collects displayed reputations and $X(t)$ collects exogenous covariates such as location, category, and inventory constraints. In ranking-based systems, $E_i(t)$ is induced by the probability distribution over rank positions and click-through behavior. A convenient reduced form treats $E_i(t)$ as proportional to a softmax of a score:

$$E_i(t) = \frac{\exp\{\eta\, h(S_i(t), X_i(t))\}}{\sum_{k=1}^{N} \exp\{\eta\, h(S_k(t), X_k(t))\}}, \qquad (1)$$

where $\eta \geq 0$ controls concentration and h is a scoring function. As $\eta \to 0$, exposure becomes uniform; as η grows, exposure concentrates on top-scored sellers.

A buyer who is exposed to seller i chooses whether to transact [7]. Let $U_i(t)$ be the buyer's utility from transacting with i, modeled as

$$U_i(t) = \alpha_0 + \alpha_q q_i + \alpha_s S_i(t) - p_i(t) + \xi_i(t), \qquad (2)$$

where $p_i(t)$ is price and $\xi_i(t)$ is an idiosyncratic taste shock, often assumed to be i.i.d. Type I extreme value, yielding a logit choice structure. Conditional on being exposed to a choice set, the probability of choosing seller i is increasing in q_i and $S_i(t)$. The transaction intensity for seller i can then be written as

$$\lambda_i(t) = \lambda_B(t)\, E_i(t)\, \sigma_i(t), \qquad (3)$$

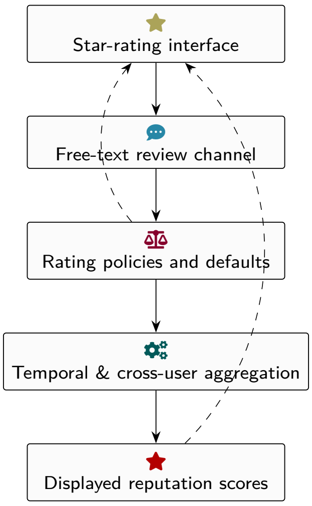

Figure 3: Decomposition of ratings, reviews, and reputation mechanisms. Interface design shapes which users rate and how they use discrete scales, review channels capture rich but selective feedback, policies and defaults control the mapping from raw inputs to stored signals, and aggregation schemes determine how past feedback is summarized. The final displayed reputation scores influence subsequent user decisions, closing the loop between mechanism design and long-run marketplace outcomes.

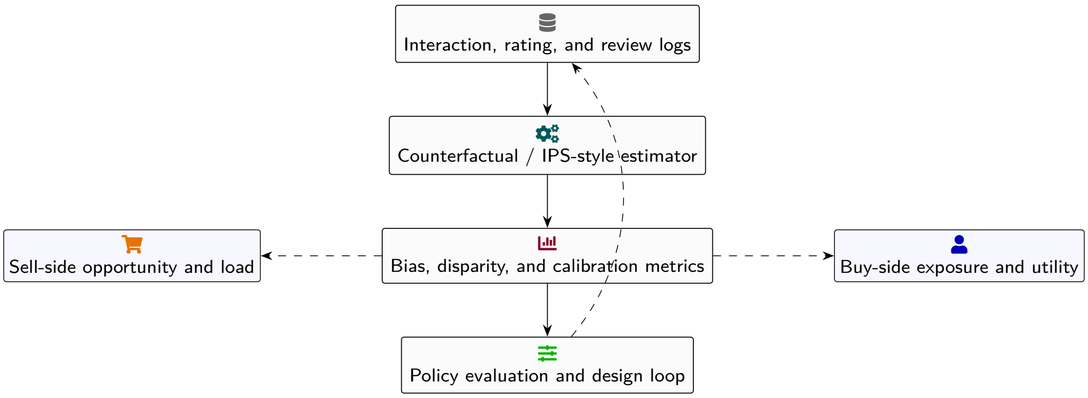

Figure 4: Estimation pipeline for quantifying feedback-loop bias. Logged interactions, ratings, and reviews are processed by counterfactual estimators that correct for exposure and selection effects. The resulting bias and disparity metrics are computed separately for buyer- and seller-side groups, and they inform policy updates to ranking and reputation mechanisms. The policy loop feeds back into the logs, enabling iterative analysis of long-run bias dynamics.

where $\sigma_i(t)$ is the conditional probability that an exposed buyer transacts with i, which may itself depend on $S_i(t)$ and prices in the choice set.

After a transaction, the buyer may leave a rating and possibly a text review. Let Z_{ij} denote an indicator that the jth transaction with seller i generates a rating, and let Y_{ij} be the numeric rating, often bounded and discrete, such as $\{1, 2, 3, 4, 5\}$. A standard starting point is a latent continuous satisfaction variable Y^*_{ij} with

$$Y^*_{ij} = q_i + \varepsilon_{ij}, \tag{4}$$

where ε_{ij} is mean-zero noise capturing idiosyncratic match quality and measurement error. Observed stars arise by thresholding: $Y_{ij} = m$ if $\tau_{m-1} < Y^*_{ij} \leq \tau_m$ for thresholds $\{\tau_m\}$. This ordered-response view is useful because it preserves a continuous latent scale while allowing discrete observed ratings. Review propensity can be modeled as

$$\mathbb{P}(Z_{ij} = 1 \mid q_i, S_i(t_{ij}), \varepsilon_{ij}) = \rho(\gamma_0 + \gamma_q q_i + \gamma_s S_i(t_{ij}) + \gamma_\varepsilon \varepsilon_{ij}), \tag{5}$$

where ρ is a logistic link. This allows selection into rating: dissatisfied buyers may be more likely to rate, or satisfied buyers may be more likely, depending on incentives and norms. Crucially, $S_i(t)$ can enter the propensity, capturing expectation effects: if displayed reputation raises expectations, identical experiences may be rated differently or may produce different propensities to leave feedback.

The platform aggregates observed ratings and other signals into the internal state $R_i(t)$. A widely used aggregation is a smoothed average, possibly with a Bayesian prior:

$$\hat{\mu}_i(t) = \frac{\kappa\mu_0 + \sum_{j:t_{ij} \leq t} Y_{ij}}{\kappa + n_i(t)}, \tag{6}$$

where $n_i(t)$ is the number of observed ratings by time t, μ_0 is a prior mean, and κ controls shrinkage. The internal state may include uncertainty, for instance a posterior variance $\hat{v}_i(t)$ under a Gaussian approximation, or a Beta-Binomial structure if ratings are binarized. Many ranking systems effectively operate on a conservative estimate such

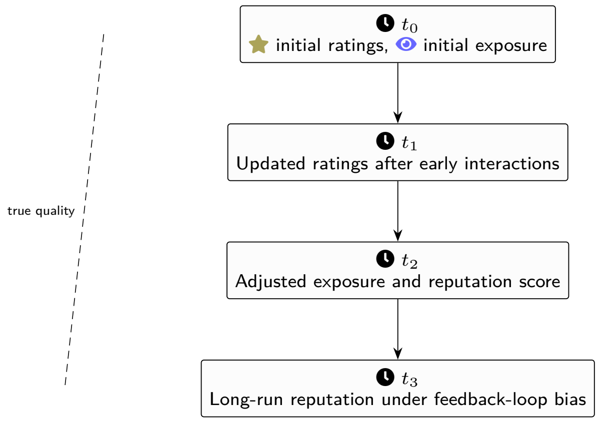

Figure 5: Temporal evolution of a single seller's reputation under biased feedback loops. Even when true underlying quality is stable, noisy and selective early ratings can alter exposure, which in turn shapes subsequent feedback and reputation updates. Measuring the divergence between the true-quality trajectory and the realized reputation trajectory provides one way to quantify feedback-loop bias over time.

as a lower confidence bound:

$$R_i(t) = \hat{\mu}_i(t) - z_\delta \sqrt{\hat{v}_i(t)}, \tag{7}$$

where z_δ is a quantile controlling risk aversion. This already introduces nonlinearity that can generate Jensen effects when combined with variable sample sizes across sellers.

To model the coupled dynamics compactly, treat the arrival of ratings as a point process [8]. Let $N_i(t)$ be the counting process for ratings received by seller i up to time t. Conditional on the filtration generated by histories, $N_i(t)$ has stochastic intensity $\lambda_i(t)\,\bar{\rho}_i(t)$, where $\bar{\rho}_i(t)$ is the expected rating propensity among transacting buyers at time t. A diffusion approximation for the evolution of an average rating estimate can be derived when ratings accumulate frequently. Let $r_i(t)$ denote a continuous approximation to the internal reputation state. For example, for an exponential moving average with forgetting rate $\alpha > 0$,

$$dr_i(t) = \alpha\big(\tilde{y}_i(t) - r_i(t)\big)\,dt, \tag{8}$$

where $\tilde{y}_i(t)$ is the instantaneous average of incoming ratings, which is itself random and depends on the realized transactions. Under high-frequency arrivals, $\tilde{y}_i(t)$ can be approximated as q_i plus noise with variance inversely proportional to the instantaneous rating arrival intensity. This yields an SDE of the form

$$dr_i(t) = \alpha\big(q_i - r_i(t)\big)\,dt + \frac{\alpha\sigma_\varepsilon}{\sqrt{\lambda_i(t)\bar{\rho}_i(t)}}\,dW_i(t), \tag{9}$$

where $W_i(t)$ is a standard Brownian motion and σ_ε captures rating noise on the latent scale. The key feature is that the diffusion coefficient depends on $\lambda_i(t)$, which depends on exposure, which depends on $r_i(t)$ through $S_i(t)$. Hence the noise intensity depends on the state, producing state-dependent learning rates and heteroskedastic measurement error. This endogenous heteroskedasticity is one pathway by which bias arises, because sellers in low-exposure states experience slower variance reduction and remain vulnerable to noise-driven downward drifts.

Text reviews can be incorporated by treating sentiment scores or embeddings as additional signals. Let T_{ij} be a scalar sentiment extracted from review text, modeled as $T_{ij} = \beta_T q_i + \nu_{ij}$ with its own noise. The platform may combine stars and text via a weighted score, or via a latent factor model. A general representation is to let the internal state follow a Bayesian update for a latent quality estimate $\theta_i(t)$ based on observed stars and text. Even when the update is optimal given the observation model, the selection of observations is endogenous, so the posterior is not necessarily centered at the true quality in the coupled equilibrium.

Definition and Decomposition of Feedback-Loop Bias

Feedback-loop bias is defined relative to a counterfactual in which the mapping from reputational state to exposure is altered while preserving the rating-generation mechanism conditional on realized transactions. Let $\mathcal{P}$ denote the observed platform policy, consisting of an aggregation rule for reputations and an exposure rule π. Let $\mathcal{P}_0$ denote a benchmark policy that weakens coupling, such as exposure that depends only on exogenous covariates X or on a debiased estimate that is conditionally independent of rating noise. The precise choice of $\mathcal{P}_0$ depends on the evaluative question, but a common benchmark is one that uses the same average exposure across sellers as $\mathcal{P}$ while randomizing its allocation with respect to reputational noise. Formally, define a coupling parameter η in the exposure function and consider the path of policies indexed by η, with $\eta = 0$ corresponding to uniform exposure within a category and $\eta = \eta^\star$ corresponding to the deployed policy [9].

Let $O(t)$ denote an outcome of interest at time t, such as the cross-sectional distribution of reputations $\{S_i(t)\}$, the correlation between exposure and latent quality, the fraction of transactions allocated to top-quality sellers, or welfare. Let $\mathbb{E}_\eta[O(t)]$ denote expectation under the policy with coupling η. The feedback-loop bias for outcome O over horizon t is

$$\mathcal{B}_O(t;\eta^\star,\eta_0) = \mathbb{E}_{\eta^\star}[O(t)] - \mathbb{E}_{\eta_0}[O(t)]. \tag{10}$$

When focusing on steady states, define $\mathcal{B}_O(\infty;\eta^\star,\eta_0)$ as the difference in stationary expectations if the process is ergodic. In practice, finite-horizon bias is often more relevant because platforms change policies and sellers churn.

To make the definition operational, choose outcomes that capture amplification of noise. One useful outcome is the exposure-weighted mean quality:

$$\bar{q}_E(t) = \sum_{i=1}^{N} E_i(t)\, q_i, \tag{11}$$

which measures allocative efficiency in the sense of matching buyers to higher quality. Another is the exposure-weighted mean reputational error:

$$\bar{\epsilon}_E(t) = \sum_{i=1}^{N} E_i(t)\, \big(S_i(t) - q_i\big), \tag{12}$$

where q_i and $S_i(t)$ must be on commensurate scales; if stars are discrete, a link function can map them to a latent utility scale. Under an idealized unbiased measurement channel with exposure independent of noise, $\mathbb{E}[\bar{\epsilon}_E(t)]$ would be close to zero. Under coupling, $\mathbb{E}[\bar{\epsilon}_E(t)]$ can deviate from zero because exposure is larger when $S_i(t)$ is high, and high $S_i(t)$ can occur due to positive noise realizations that also drive exposure, creating selection on noise.

A decomposition clarifies multiple channels. Let $S_i(t) = s(q_i, \mathcal{H}_i(t))$ where $\mathcal{H}_i(t)$ is the history of observed ratings and reviews. Write $S_i(t) = q_i + u_i(t)$ where $u_i(t)$ is the reputational error. Although $\mathbb{E}[u_i(t) \mid q_i]$ may be zero in a hypothetical scenario with exogenous sampling, under endogenous sampling it can become nonzero. Consider the covariance term

$$\mathbb{E}[\bar{\epsilon}_E(t)] = \mathbb{E}\left[\sum_i E_i(t) u_i(t)\right] = \sum_i \mathbb{E}[E_i(t) u_i(t)]. \tag{13}$$

Even if $\mathbb{E}[u_i(t)] = 0$ marginally, the term $\mathbb{E}[E_i(t)u_i(t)]$ can be positive if $E_i(t)$ is increasing in $S_i(t)$ and hence in $u_i(t)$. For smooth exposure policies, a local approximation yields

$$\mathbb{E}[E_i u_i] \approx \mathbb{E}\left[\frac{\partial E_i}{\partial S_i}\right] \mathbb{E}[u_i^2] + \frac{1}{2}\mathbb{E}\left[\frac{\partial^2 E_i}{\partial S_i^2}\right] \mathbb{E}[u_i^3] + \cdots, \tag{14}$$

highlighting that variance and skewness of reputational errors interact with the curvature of exposure mapping. This indicates that even symmetric noise can produce positive bias when exposure is locally convex in the reputational signal. The effect is magnified when $\mathbb{E}[u_i^2]$ is large, which is typical for low-sample sellers. Hence, policies that aggressively exploit noisy reputations can induce a form of winner's-curse selection where those with upward noise are preferentially sampled and reinforced.

Endogenous rater composition introduces another component [10]. Suppose rating noise decomposes as $\varepsilon_{ij} = a_{b(ij)} + e_{ij}$ where a_b is a buyer-specific rating tendency and e_{ij} is idiosyncratic. If exposure induces a change in which buyers transact with which sellers, then the distribution of $a_{b(ij)}$ conditional on seller i depends on $S_i(t)$ and therefore on $u_i(t)$. This creates a term where reputational error is correlated with future rater tendencies, potentially producing drift. A simplified representation uses the conditional mean of noise given displayed reputation:

$$\mathbb{E}[\varepsilon_{ij} \mid q_i, S_i(t_{ij})] = \delta(S_i(t_{ij})), \tag{15}$$

where $\delta(\cdot)$ captures expectation effects and selection. If δ is decreasing, high displayed reputations raise expectations and can generate harsher ratings for the same experience, which can dampen runaway amplification. If δ is increasing, social proof can generate leniency, which can accelerate amplification. In either case, coupling changes the mapping from q_i to the distribution of observed ratings, so the reputational signal is no longer conditionally unbiased.

Nonlinear aggregation contributes further. Many platforms discretize, cap, or threshold reputations. Let $S_i(t) = \mathrm{clip}(\mathrm{round}(g(\hat{\mu}_i(t))))$ for some function g. The rounding operator introduces quantization error whose sign depends on the fractional part of $g(\hat{\mu}_i(t))$. When exposure depends on $S_i(t)$, the system effectively endogenizes the quantization error because sellers are differentially sampled depending on which side of a rounding boundary they fall. Threshold-based eligibility, such as removing sellers below 4.2 stars, introduces absorbing states in which low scores terminate future sampling, creating selection bias in the observed distribution of ratings among remaining sellers. This can make cross-sectional averages appear high even when underlying quality is moderate, and it can complicate inference because low-quality sellers are underrepresented.

To describe amplification and persistence, define an amplification factor based on the sensitivity of long-run exposure to early noise. Let $u_i(t_0)$ be the reputational error at an early time t_0 after a small number of ratings. Define

$$A(t;t_0) = \frac{\mathrm{Cov}\big(E_i(t), u_i(t_0)\big)}{\mathrm{Var}\big(u_i(t_0)\big)}, \tag{16}$$

with covariance and variance taken across sellers and stochastic realizations [11]. If $A(t;t_0)$ remains positive for large t, early noise has persistent exposure effects, indicating hysteresis. This concept links to dynamic treatment effects where early reputation acts as a treatment that changes future data collection.

Frequency-domain diagnostics can also quantify feedback. Let $x_i(t)$ be a detrended time series of reputational changes for seller i, such as $x_i(t) = dS_i(t)/dt$ or discrete

Symbol	Description	Domain
u	User index	$\mathcal{U}$
i	Item / provider index	$\mathcal{I}$
r_{ui}	Observed rating for pair (u, i)	$\mathbb{R}$
$\hat{r}_{ui}$	Predicted rating	$\mathbb{R}$
e_{ui}	Exposure indicator	$\{0, 1\}$
p_{ui}	Propensity / exposure probability	$[0, 1]$
t	Discrete time step	$\{1, \ldots, T\}$

Table 1: Key notation used to model exposure, feedback, and ratings in the two-sided marketplace.

Marketplace	# Users	# Items / Providers	# Interactions
Ride-sharing	120,842	18,503	2,741,906
Food delivery	84,190	9,237	1,154,220
Freelance services	41,322	6,801	597,488
All platforms	246,354	34,541	4,493,614

Table 2: Dataset statistics across three representative two-sided marketplaces.

differences. In systems with algorithmic updates and periodic demand cycles, coupling can produce oscillatory behavior where exposure and ratings reinforce on certain time scales. The power spectral density $P_i(\omega)$ of $x_i(t)$ can be estimated, and the aggregate spectrum can be compared across policies to see whether coupling amplifies specific frequencies. While frequency analysis does not by itself establish bias, it can detect algorithmically induced resonance where periodic rank updates interact with demand seasonality.

A signal-to-noise characterization provides an interpretable scalar. Define a latent-scale signal variance across sellers as $\mathrm{Var}(q_i)$ and a noise variance of reputational estimates as $\mathbb{E}[\hat{v}_i(t)]$. A logarithmic signal-to-noise ratio can be reported as

$$\mathrm{SNR}_{\mathrm{dB}}(t) = 10 \log_{10} \left(\frac{\mathrm{Var}(q_i)}{\mathbb{E}[\hat{v}_i(t)]} \right), \tag{17}$$

interpreted analogously to decibels. Coupling changes $\hat{v}_i(t)$ endogenously by changing transaction intensities. A policy may increase average $\mathrm{SNR}_{\mathrm{dB}}$ by concentrating exposure, yet simultaneously increase feedback-loop bias by making exposure depend more strongly on errors for low-sample sellers. Reporting both a bias measure and an SNR-like measure helps distinguish learning acceleration from allocative distortion.

Identification and Statistical Estimation under Endogenous Sampling

Estimating feedback-loop bias requires separating latent quality from observed reputation and quantifying counterfactual outcomes under alternative exposure coupling. This is challenging because quality is unobserved and because the data generating process depends on platform policy. A practical approach is to specify a joint model for transactions, ratings, and exposure, estimate its parameters, and then simulate counterfactual policies. Another approach is to identify certain bias components nonparametrically using randomized perturbations or quasi-experimental variation [12]. The present section develops estimators that combine these ideas.

Assume access to platform logs containing impression counts, click and transaction events, prices, displayed reputations, and rating outcomes with timestamps. Let I_{it} denote impressions for seller i in a small time interval around t, and let C_{it} and T_{it} denote clicks and transactions. In a ranking system, impressions are the first stage through which exposure is realized. A reduced-form model can treat impressions as Poisson with intensity proportional to $E_i(t)$ times buyer arrivals. Transactions are then generated conditional on impressions with a conversion probability that depends on q_i, $S_i(t)$, and price. Ratings occur conditional on transactions and follow an ordered-response model.

A latent-quality model can be specified as a hierarchical state-space model. Let q_i be drawn from a population distribution, such as $q_i \sim \mathcal{N}(\mu_q, \sigma_q^2)$ within category. Observed ratings provide noisy measurements. Conditional on q_i and buyer-level shocks, the likelihood of stars can be written using thresholds. The platform's displayed score $S_i(t)$ is a deterministic function of past observed ratings and perhaps other signals. If the analyst knows the aggregation rule, $S_i(t)$ is computable from history; if not, it can be parameterized and estimated. The exposure

Bias type	**Mechanism**	**Typical effect on data**
Position bias	Higher slots receive more attention	Overestimates quality of top-ranked items
Popularity bias	Popular items shown more often	Feedback concentrates on a few items
Rating inflation	Lenient users give systematically high scores	Compresses rating scale near the top
Strategic rating	Agents manipulate ratings for advantage	Breaks link between ratings and true quality
Survival bias	Low-rated items are removed from display	Under-represents poorly performing items

Table 3: Conceptual taxonomy of feedback-loop biases arising from ranking, visibility, and strategic behavior.

Setting	**Exposure policy**	**Feedback signal**	**Primary purpose**
Offline i.i.d.	Uniform over candidate set	Historical ratings	Benchmarking prediction quality
Logged bandit	Production ranking logs	Clicks + ratings	Off-policy evaluation
Simulated feedback loop	Learned policy with re-ranking	Synthetic ratings	Long-horizon bias accumulation
Cold-start scenario	Popularity-based seeding	Early reviews	Effect of bias on new items
Unbiased oracle	Counterfactual random exposure	Full rating matrix	Upper bound on achievable gains

Table 4: Experimental configurations used to stress-test feedback-loop behavior under different exposure mechanisms.

process can be modeled as

$$\log E_i(t) = \eta h(S_i(t), X_i(t)) - \log\left(\sum_k \exp\{\eta h(S_k(t), X_k(t))\}\right) + \zeta_i(t), \tag{18}$$

where $\zeta_i(t)$ captures stochastic exploration, ad placements, or unmodeled factors. The dependence of $E_i(t)$ on $S_i(t)$ is central; identification of η requires variation in $S_i(t)$ that is not fully determined by q_i and past exposure, which is generally not available without experiments.

Randomized interventions provide clean identification. Many platforms conduct A/B tests that randomize ranking weights, display formats, or exploration rates [13]. Suppose that at times or for subsets of users, the platform perturbs the exposure mapping by adding random noise to the ranking score or by randomizing the order of a subset of sellers. Let Z_{it} be a randomized instrument that affects $E_i(t)$ but is independent of q_i and rating noise conditional on covariates. Then one can identify the causal effect of exposure on future reputation updates by comparing the evolution of $S_i(t)$ across instrument levels. A simple dynamic IV estimand targets

$$\beta_{\text{dyn}}(\Delta) = \frac{\mathbb{E}[S_i(t+\Delta) - S_i(t) \mid Z_{it}=1] - \mathbb{E}[S_i(t+\Delta) - S_i(t) \mid Z_{it}=0]}{\mathbb{E}[E_i(t) \mid Z_{it}=1] - \mathbb{E}[E_i(t) \mid Z_{it}=0]}, \tag{19}$$

which measures how incremental exposure translates into incremental reputation over horizon Δ. This captures feedback strength. However, the numerator depends on rating composition and buyer behavior, so interpreting β_{dyn} requires assumptions about stable rating mechanisms across instrument arms.

When randomized experiments are unavailable, quasi-experimental variation may come from discontinuities in the platform's policy. For example, if a trust badge appears when $S_i(t)$ crosses a threshold $\bar{s}$, then sellers near the threshold experience a discrete change in exposure. A regression discontinuity design can estimate the local effect of badge-induced exposure on subsequent ratings. Let $S_i(t^-)$ be the running variable just before badge assignment. Under continuity assumptions, the difference in outcomes around $\bar{s}$ identifies a local average treatment effect. Because the running variable itself is noisy and endogenously sampled, one must account for manipulation and measurement error, but the design can still be informative if threshold crossing is largely stochastic for near-threshold sellers.

Panel methods can control for seller fixed effects as proxies for quality. Consider a model for ratings conditional on transacting buyers:

$$Y^*_{ij} = \alpha_i + \beta S_i(t_{ij}) + \phi^\top W_{ij} + \varepsilon_{ij}, \tag{20}$$

where α_i is a seller fixed effect and W_{ij} are controls. Here β captures expectation effects: how displayed reputation influences reported satisfaction holding seller constant. Estimating β is difficult because $S_i(t)$ is a function of past Y, but lagged values and instruments can be used. A generalized method of moments approach can exploit orthogonality conditions between future shocks and past instruments. For instance, if one can find variables that shift displayed reputation without directly affecting satisfaction, such as random rounding changes or interface experiments, they can serve as instruments [14].

Selection into rating can be addressed by jointly modeling rating propensity. Let Z_{ij} be rating occurrence and specify

$$\mathbb{P}(Z_{ij}=1) = \rho(\omega_0 + \omega_s S_i(t_{ij}) + \omega_x^\top X_{ij}), \tag{21}$$

and condition the rating model on $Z_{ij} = 1$ using a selection correction. In a parametric setting, one can use a bivariate probit or Heckman-type correction on the latent scale, though the discreteness of stars suggests ordered models with selection. Bayesian hierarchical approaches can integrate over missing ratings, treating unobserved ratings as latent and using the propensity model to infer their distribution. This helps separate true satisfaction from observed feedback, which is essential when exposure changes the propensity to rate.

To connect estimation to feedback-loop bias, one needs an estimate of latent quality q_i that is not itself contaminated by the same feedback one intends to measure. A pragmatic approach is to estimate q_i using

Method	Ride-sharing	Food delivery	Freelance
MF-Naive	0.892	0.926	0.978
IPW-MF	0.847	0.883	0.932
DR-MF	0.832	0.871	0.918
Causal Embeddings	0.824	0.862	0.909
Oracle (unbiased)	0.801	0.843	0.891

Table 5: RMSE (↓) of rating prediction models across platforms; lower values indicate better fit to unbiased counterfactual ratings.

Method	ECE ↓	Brier score ↓	AUC ↑
MF-Naive	0.084	0.192	0.731
IPW-MF	0.057	0.176	0.754
DR-MF	0.049	0.169	0.768
Causal Embeddings	0.043	0.162	0.781

Table 6: Calibration and ranking quality of rating-to-outcome models on held-out data.

a model that incorporates the selection mechanism and uses all available signals, including outcomes that are less affected by displayed reputation, such as refunds, repeat purchase rates, or objective delivery metrics. If such auxiliary outcomes are available, they can anchor quality estimates. Let D_{ij} be an objective performance metric with model $D_{ij} = \theta q_i + \upsilon_{ij}$. Even if υ_{ij} is noisy, it may be less sensitive to expectation effects, providing an additional measurement channel. Multi-signal factor models can combine stars, text sentiment, and objective metrics to estimate q_i with reduced bias.

Once parameters are estimated, counterfactual simulation computes $\mathbb{E}_\eta[O(t)]$ under alternative coupling η. The simulation must preserve the causal structure: exposure affects transactions, which affects which ratings are observed. For each seller, simulate buyer arrivals and match outcomes using estimated demand parameters, generate ratings and reviews using the estimated measurement model, update reputations using the platform aggregation rule, and feed reputations into exposure via the chosen η. Repeating over many stochastic paths yields Monte Carlo estimates of outcomes and thus of feedback-loop bias.

Inference for bias estimands can be done via bootstrap over sellers and time blocks to capture dependence. When simulation is nested inside estimation, a common approach is to use a parametric bootstrap: sample parameters from their estimated sampling distribution, run counterfactual simulations, and compute the distribution of bias. Hypothesis tests can target whether coupling increases allocative efficiency while also increasing disparity. For example, test the null that $\mathcal{B}_{\bar{q}_E}(t;\eta^\star,0) = 0$ using a bootstrap confidence interval. Tests for amplification persistence can target whether $A(t;t_0)$ decays to zero [15]. A permutation test can be constructed by reassigning early rating shocks across sellers while preserving their qualities, then recomputing exposure trajectories; this isolates the role of early noise.

Descriptive statistics remain important as diagnostics. One can report the distribution of $n_i(t)$ across sellers, the relationship between $S_i(t)$ and $n_i(t)$, and the fraction of exposure allocated to low-sample sellers. A hallmark of strong feedback is a heavy-tailed exposure distribution where a small fraction of sellers receives a large share of impressions. If, for example, the top 5% of sellers by displayed reputation receives 60% of impressions, then variance reduction and learning are concentrated, and low-sample sellers remain noisy. Such summary measures can be compared across policy variants.

Distributional Dynamics and Numerical Methods

Analytical characterization of the coupled reputation-exposure system often requires studying the evolution of the distribution of reputational states across sellers. When N is large and sellers are exchangeable within categories, mean-field approximations can be useful. Consider a representative seller with latent quality q drawn from $f_Q(q)$ and a reputation state $r(t)$ evolving according to an SDE with state-dependent diffusion:

$$dr(t) = a(r(t), q)\,dt + b(r(t), q)\,dW(t), \tag{22}$$

where $a(r, q) = \alpha(q - r)$ under a simple smoothing update and $b(r, q) = \alpha\sigma_\varepsilon/\sqrt{\lambda(r)\bar{\rho}(r)}$ under the diffusion approximation, with $\lambda(r)$ induced by exposure mapping from r to impressions and transactions. The cross-sectional density $p(r, t \mid q)$ of reputations conditional on quality satisfies a Fokker–Planck equation:

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial r}\big(a(r,q)p\big) + \frac{1}{2}\frac{\partial^2}{\partial r^2}\big(b(r,q)^2 p\big). \tag{23}$$

Policy	Avg. rating	Gini exposure ↓	New-item impressions (%) ↑
Logged production	4.62	0.71	6.3
Unbiased random	4.21	0.29	24.8
Debiased ranking	4.48	0.44	15.6
Popularity-only	4.67	0.83	2.1

Table 7: Counterfactual simulation of alternative ranking policies and their effect on exposure inequality and exploration.

Variant	User NDCG@10 ↑	Item Gini ↓	Coverage@10 (%) ↑
Full model	0.842	0.48	37.2
w/o user covariates	0.819	0.51	33.9
w/o item covariates	0.812	0.54	31.7
w/o temporal component	0.803	0.57	29.5
w/o propensity correction	0.781	0.63	24.1

Table 8: Ablation analysis of model components on accuracy, exposure inequality, and catalog coverage.

The unconditional density is $p(r,t) = \int p(r,t \mid q) f_Q(q)\, dq$. Outcomes such as the expected exposure-weighted mean quality can be written as integrals over q and r, for example

$$\mathbb{E}[\bar{q}_E(t)] \approx \frac{\int q\, \lambda(r)\, p(r,t \mid q)\, f_Q(q)\, dr\, dq}{\int \lambda(r)\, p(r,t)\, dr}. \quad (24)$$

This integral highlights that coupling enters through $\lambda(r)$, which weights the joint distribution. When $\lambda(r)$ is steep, the integral becomes dominated by high-r regions whose mass depends on the diffusion and drift, thereby connecting feedback-loop bias to PDE properties.

Closed-form solutions are rare because $b(r,q)$ depends on $\lambda(r)$, which can be exponential in r under softmax-like ranking [16]. Numerical methods are therefore central. A finite difference discretization on a bounded domain for r can solve the PDE, but finite element methods are often more stable under variable diffusion coefficients and complex boundary conditions. Let the reputational state domain be truncated to $r \in [r_{\min}, r_{\max}]$ with reflecting or absorbing boundaries representing clipping and deplatforming. Define a mesh with nodes $\{r_k\}$ and basis functions $\{\varphi_k(r)\}$. Seek an approximate solution $p_h(r,t) = \sum_k c_k(t)\varphi_k(r)$. Multiplying the PDE by a test function and integrating yields a weak form leading to a system of ODEs:

$$M\dot{c}(t) = K(c(t)), \quad (25)$$

where M is a mass matrix and K encodes drift and diffusion operators, potentially nonlinear if b depends on p through mean-field coupling. Time stepping can be performed with implicit schemes such as backward Euler or Crank–Nicolson to preserve stability when diffusion is small in high-exposure regions and large in low-exposure regions. Positivity preservation is important because p is a density; stabilization methods or constrained solvers can enforce $c_k(t) \geq 0$.

A complementary approach uses direct simulation of the SDE for many sellers, which is Monte Carlo rather than PDE-based. Simulation is straightforward but can be computationally intensive when exploring policy parameters. Variance reduction techniques can help. One can use common random numbers across policy variants to estimate bias differences with lower variance, because $\mathcal{B}_O$ is a difference of expectations. Specifically, simulate the same Brownian paths and demand shocks under two policies and compute the difference in outcomes pathwise, then average. This is particularly useful when $\eta^\star$ and η_0 are close.

Parametric analysis of coupling can be conducted by treating η as a continuous parameter in the exposure mapping and differentiating outcomes with respect to η. Under regularity conditions, one can compute sensitivity via likelihood ratio or pathwise derivatives. For example, if exposure shares are softmax, then [17]

$$\frac{\partial E_i}{\partial \eta} = E_i \left(h(S_i, X_i) - \sum_k E_k h(S_k, X_k) \right), \quad (26)$$

which shows that increasing η increases exposure for above-average scored sellers and decreases it for below-average. Sensitivity of bias measures to η can be computed by differentiating integrals or by finite differences in simulation. This enables local optimization where the platform chooses η to balance objectives.

The coupled system can be framed as a control problem. Let π_θ denote a family of policies parameterized by θ, which may include coupling strength, exploration rate, and confidence-bound penalties. Define an objective

$$J(\theta) = \mathbb{E}_\theta[W(t)] - \lambda_{\text{bias}}\, \mathbb{E}_\theta[B(t)] - \lambda_{\text{disp}}\, \mathbb{E}_\theta[D(t)], \quad (27)$$

where $W(t)$ is welfare, $B(t)$ is a bias measure such as $\bar{\epsilon}_E(t)^2$, and $D(t)$ is a disparity measure such as the variance of exposure across protected groups conditional on quality. The weights λ_{bias} and λ_{disp} encode policy priorities. Gradient-based optimization can be applied if J is differentiable, using simulation-based gradients. When J is nonconvex due to threshold effects, derivative-free methods or stochastic approximation may be used. However, even with smooth policies, nonconvexity can arise because the steady-state distribution of reputations changes discontinuously when the system transitions between exploration-dominated and exploitation-dominated regimes.

Analogies to propagation phenomena can yield additional intuition and numerical tools. Consider a representation in which reputational influence spreads over a graph or manifold of seller similarity, such as category adjacency or geographic proximity. Let $u(x,t)$ be a reputation field over a continuous space x capturing seller attributes, and let exposure-induced learning be modeled as a diffusion with damping:

$$\frac{\partial u}{\partial t} = \kappa \Delta u - \gamma u + s(x,t), \tag{28}$$

where $s(x,t)$ is a source term from ratings and Δ is a Laplacian. This resembles heat or sound propagation with attenuation. While this field model is not literal for individual seller scores, it is useful when ranking algorithms use similarity-based smoothing, such as transferring priors across similar sellers. In such cases, local shocks can influence nearby sellers, and coupling can create spatially correlated biases [18]. Numerical solvers for diffusion and Helmholtz-type equations can then be repurposed for reputation smoothing analysis.

Frequency spectrum analysis can be integrated into the numerical study by simulating reputation and exposure time series under different update cadences. Suppose the platform updates rankings at discrete intervals Δt and buyers arrive with daily seasonality $\lambda_B(t) = \lambda_0(1 + \theta_s \sin(2\pi t/T))$ with period T. Coupling can create harmonics in exposure time series. By computing discrete Fourier transforms of simulated exposure shares, one can quantify how much variance occurs at the fundamental frequency and its multiples. If policy changes increase energy at certain frequencies, this may indicate unstable oscillations where reputations overshoot and correct, which can be undesirable for predictability. Stability can also be studied by linearizing the mean-field dynamics around a fixed point and examining eigenvalues; if the real part is positive, small perturbations grow, indicating amplification instability.

Empirical Quantification of Amplification, Persistence, and Welfare Distortion

To operationalize the framework in real data, one must select measurable estimands that reflect feedback-loop bias while remaining identifiable. This section proposes a set of estimands centered on amplification, persistence, and welfare distortion, and outlines a practical estimation workflow grounded in panel logs.

Amplification is the degree to which early random shocks to reputation translate into long-run exposure differences. Suppose that at onboarding, new sellers have limited feedback. Let t_0 be an early time such as the moment after the first m ratings. Define an early reputation statistic $S_i(t_0)$ and decompose it as $S_i(t_0) = \mathbb{E}[S_i(t_0) \mid q_i] + u_i(t_0)$, where $u_i(t_0)$ is the residual. Amplification at horizon $t_1 > t_0$ can be measured by regressing exposure at t_1 on early residuals controlling for quality proxies:

$$E_i(t_1) = \alpha + \psi u_i(t_0) + \varphi^\top \hat{q}_i + \epsilon_i. \tag{29}$$

Here $\hat{q}_i$ is an estimate or proxy for quality using objective metrics or long-run averages. The coefficient ψ estimates how much early noise predicts later exposure. Because $u_i(t_0)$ is constructed from observed ratings, it may still be correlated with unobserved quality components, so the regression should be interpreted cautiously unless $\hat{q}_i$ is credible and early variation is plausibly random, for instance when onboarding traffic is randomized.

Persistence concerns how long early shocks matter [19]. One can compute $\psi(\Delta)$ as a function of horizon $\Delta = t_1 - t_0$ and examine its decay. If $\psi(\Delta)$ remains materially positive for large Δ, the system exhibits path dependence. In practice, persistence may be mediated by seller churn and by platform re-ranking. Conditioning on survival can itself induce bias; sellers with low early reputations may exit, so persistence among survivors may understate the effect. A joint model of exposure and exit can address this. Let $H_i(t)$ be a hazard of exit with

$$\log H_i(t) = \chi_0 + \chi_s S_i(t) + \chi_e \log E_i(t) + \chi_q q_i. \tag{30}$$

Estimating this hazard helps distinguish whether early shocks reduce future exposure directly or indirectly by increasing exit.

Welfare distortion is measured by comparing realized matching and utility to a counterfactual with reduced coupling. Define buyer surplus from matching to seller i as $v(q_i) - p_i$, where v is an increasing function. Aggregate expected surplus under policy η over horizon t is

$$\mathrm{CS}_\eta(t) = \mathbb{E}_\eta\left[\int_0^t \sum_i \lambda_i(s)\, \mathbb{E}[v(q_i) - p_i(s) \mid \text{match } i]\, ds\right]. \tag{31}$$

Seller surplus may depend on transactions and costs, and platform revenue may depend on fees. Total welfare can be defined as a weighted sum. Feedback-loop bias in welfare is then $\mathrm{CS}_{\eta^\star}(t) - \mathrm{CS}_{\eta_0}(t)$ or similarly for total welfare.

Because q_i is latent, welfare must be approximated using estimated quality or observed outcomes. One approach is to map observed repeat purchase, complaint rates, or objective performance to a utility metric. For example, if complaint rate c_i is observed and is decreasing in quality, one can define a monotone mapping $v(q_i) \approx v_0 - \omega c_i$ for some ω calibrated from buyer retention [20]. This yields a welfare proxy. Another approach is to estimate demand parameters from choice data and compute expected utility using a discrete-choice model where q_i enters as a random coefficient. The accuracy of welfare comparisons then depends on the model fit.

A key empirical challenge is disentangling the direct effect of reputation on buyer choice from the effect mediated through exposure. A platform may show different sellers to different buyers; therefore observed transactions are conditional on exposure. If one estimates a choice model only on exposed sets, the estimated effect of reputation on choice may be biased if exposure selection correlates with unobserved components of utility. A two-stage model can help: first model exposure assignment, then model choice conditional on exposure with correction for selection. In logit frameworks, control-function approaches can incorporate the residual from the exposure equation into the choice equation. Alternatively, one can treat exposure as an instrumented variable in a structural model.

To quantify bias in reputational measurement itself, compare the observed distribution of reputations to an estimated distribution of latent quality. If the reputation estimator is a posterior mean under a model that correctly accounts for selection, then $S_i(t) - \hat{q}_i$ should have limited systematic dependence on exposure after conditioning on uncertainty. An empirical diagnostic is to regress reputational residuals on log exposure:

$$S_i(t) - \hat{q}_i = \beta_0 + \beta_1 \log(E_i(t) + \epsilon) + \nu_i. \quad (32)$$

A positive β_1 suggests that higher exposure is associated with upward reputational residuals, consistent with selection on noise or composition effects. To avoid mechanical correlation due to the aggregation formula, one can use lagged exposure or exposure shocks from experiments.

Hypothesis testing can be framed around whether coupling induces statistically detectable amplification beyond what would occur under independent sampling. One can simulate a null model in which ratings are generated for each seller at a rate independent of reputation but matched to the observed number of ratings, then compute the distribution of amplification measures under this null [21]. Comparing observed amplification to the null yields a p-value for whether the coupling is stronger than expected. This is a form of model-based randomization inference.

Review text offers additional observables that can reveal expectation effects. If sentiment scores derived from text, T_{ij}, are less tightly linked to star ratings, one can test whether displayed reputation predicts sentiment conditional on objective outcomes. For example,

$$T_{ij} = \delta_0 + \delta_1 S_i(t_{ij}) + \delta_2 D_{ij} + \delta_3 p_i(t_{ij}) + \epsilon_{ij}. \quad (33)$$

A nonzero δ_1 conditional on D_{ij} suggests that text sentiment is influenced by displayed reputation, which would imply that reviews are not purely reflective of experience. Such expectation effects feed back because platforms often summarize sentiment into badges or highlights.

Logarithmic transforms can stabilize estimation because exposure and transaction counts are heavy-tailed. Using $\log(1+I_{it})$ and $\log(1+T_{it})$ can reduce heteroskedasticity. Similarly, mapping star ratings to a logit scale can linearize effects. If S_i is an average star rating in $[1, 5]$, one can normalize to $[0, 1]$ by $\tilde{S}_i = (S_i - 1)/4$ and use $\text{logit}(\tilde{S}_i)$ as a continuous index. This avoids boundary compression near 5 stars, where small changes in stars correspond to large changes in perceived quality for some buyers. Such transformations can be important when estimating how exposure responds to reputation.

In many marketplaces, the distribution of ratings is skewed toward high values, with a ceiling effect. This can be modeled with asymmetric noise or with mixture models where buyers have different rating thresholds. A mixture ordered-probit model can represent that some buyers are lenient and others strict. Estimating such mixtures helps separate seller quality from buyer composition, which is essential for bias decomposition. The resulting posterior uncertainty over q_i can be propagated into counterfactual simulations, yielding intervals for bias measures rather than point estimates.

Mechanism Design and Optimization of Reputation and Ranking under Bias Constraints

Mechanism design in this context concerns choosing how to elicit, aggregate, and display feedback, and how to use it in ranking and eligibility, to achieve objectives subject to constraints [22]. Feedback-loop bias matters because it changes the mapping from latent quality to outcomes, and because the platform's policy influences the statistical environment in which learning occurs. The problem is not merely to estimate quality but to do so while controlling the consequences of using estimates to allocate exposure.

A baseline formulation treats the platform as choosing an exposure function $\lambda(r)$, an aggregation rule producing r, and a display rule producing S, to maximize expected welfare. Let W be an objective combining buyer utility, seller revenue, and platform revenue. The platform faces a constraint that allocations should not be overly sensitive to noise. One can impose a constraint on amplification, such as $\text{Cov}(E_i(t_1), u_i(t_0)) \leq \bar{c}$ for selected horizons, or a constraint on exposure disparity conditional on quality, such as

$$\text{Var}\big(E_i(t) \mid q_i \in [q, q + dq]\big) \leq \bar{v}, \quad (34)$$

which limits how much exposure varies among similar-quality sellers. Implementing such constraints requires a model that predicts how policy changes affect the distribution of reputations and exposures.

A natural class of policies balances exploitation with exploration. In bandit terms, each seller is an arm with unknown mean quality. Showing a seller yields information through ratings and yields reward through buyer surplus. A purely greedy policy ranks by current posterior mean, which can amplify noise. An upper-confidence-bound or Thompson-sampling policy introduces exploration. For reputation systems, a conservative lower-confidence bound may be used for safety, but it can also penalize low-sample sellers and create lock-in. A symmetric confidence-bound policy can be considered:

$$\text{score}_i(t) = \hat{\mu}_i(t) + \beta\sqrt{\hat{v}_i(t)}, \tag{35}$$

where $\beta > 0$ encourages exploring uncertain sellers by increasing their score. This tends to allocate some exposure to low-sample sellers, increasing learning and potentially reducing long-run bias due to early noise [23]. However, exploration may reduce short-run buyer surplus if uncertain sellers have lower expected quality.

The design of the aggregation rule can mitigate bias by accounting for selection and by reporting uncertainty. Displaying only a point estimate S_i hides variance differences and encourages overreaction to small-sample fluctuations. Displaying a credible interval or a volume indicator can allow buyers to interpret scores appropriately. From a modeling perspective, one can define a displayed reputation as a shrinkage estimator:

$$S_i(t) = \hat{\mu}_i(t) = \frac{\kappa\mu_0 + \sum Y_{ij}}{\kappa + n_i(t)}, \tag{36}$$

which reduces variance for low n_i. Shrinkage reduces amplification of early noise because extreme early averages are pulled toward μ_0. However, if μ_0 differs across groups or categories and is estimated from historical data affected by prior bias, shrinkage can import historical distortions. A group-conditional prior can partially address this, but it raises fairness concerns if group labels are sensitive. An alternative is to use category-level priors based on objective metrics rather than historical ratings.

Review solicitation and incentives are part of the mechanism. If the platform offers prompts or reminders to rate, the selection function $\rho(\cdot)$ changes. A mechanism that increases rating rates uniformly can reduce variance and thereby reduce the curvature-based bias term associated with $\mathbb{E}[u_i^2]$. However, if prompts are targeted based on reputation or exposure, they can reinforce feedback loops. Therefore, a design principle is to avoid making rating propensity a steep function of displayed reputation. In the model, this means controlling γ_s in the propensity equation, potentially by interface design that standardizes prompts regardless of reputation [24].

Eligibility thresholds create discontinuities that can magnify bias. If sellers below a threshold are suspended, then a small early negative shock can have permanent effects. One can replace hard thresholds with smooth penalties. For example, instead of deplatforming below $\bar{s}$, reduce exposure smoothly using a logistic function:

$$E_i(t) \propto \frac{1}{1 + \exp\{-k(S_i(t) - \bar{s})\}}, \tag{37}$$

where k controls steepness. Smaller k reduces discontinuity and can reduce hysteresis. Safety concerns may still require thresholds for extreme cases, but for moderate ranges, smoothing can improve statistical properties.

Optimization can be formalized as selecting k, η, and exploration parameters to minimize a combined loss:

$$L(\theta) = \mathbb{E}_\theta\left[\int_0^t \ell\big(q_{m(s)}, S_{m(s)}(s)\big)\,ds\right] + \lambda\,\mathbb{E}_\theta\left[\bar{\epsilon}_E(t)^2\right], \tag{38}$$

where $m(s)$ denotes the seller matched at time s and ℓ penalizes low-quality matches, while the second term penalizes bias. Solving this requires simulation or PDE methods described earlier. Gradient estimates can be obtained via finite differences in θ or via stochastic gradients if the policy is differentiable and the simulation is reparameterized. In practice, platforms may restrict to a small grid of policy parameters and choose the best via offline evaluation and controlled experiments.

A fairness-relevant constraint can be expressed in terms of equality of opportunity for exposure conditional on quality. Let G_i denote group membership. A constraint might require that for any quality quantile, the expected exposure is similar across groups:

$$|\mathbb{E}[E_i(t) \mid q_i \in Q_\tau, G_i = 0] - \mathbb{E}[E_i(t) \mid q_i \in Q_\tau, G_i = 1]| \le \epsilon_\tau, \tag{39}$$

where Q_τ denotes a neighborhood of the τth quantile [25]. Because q_i is latent, enforcement would rely on estimated quality or proxies, and uncertainty must be acknowledged. A robust approach is to enforce constraints with respect to posterior distributions over q_i, yielding probabilistic constraints. This often increases computational complexity but aligns with the uncertainty inherent in reputations.

The welfare effects of bias constraints can be non-monotone. Reducing coupling may reduce allocative efficiency in the short run if reputations contain useful information. But reducing coupling can also improve efficiency in the long run by preventing lock-in and by improving learning for a broader set of sellers. Therefore, mechanism design should evaluate time horizons explicitly. A policy that increases buyer surplus by 2% over one week might decrease it over a year if it discourages entry or reduces learning. Dynamic evaluation is essential.

The framework also clarifies the role of transparency. If buyers understand that low-sample reputations are uncertain, they may discount them, reducing the platform's need to control coupling. However, if buyers over-trust star averages, then the platform's ranking policy has greater responsibility for preventing overreaction to noise. In model terms, buyer utility coefficient α_s measures how responsive demand is to displayed reputation. High α_s increases feedback strength because small changes in S generate large changes in $\lambda_i(t)$. Interface design that reduces α_s by contextualizing scores can therefore reduce feedback-loop bias without changing ranking algorithms [26]. Estimating α_s from choice data is thus directly relevant to mechanism design.

Finally, review text offers opportunities for richer signals but also for new biases. Text-based reputation features can be more informative than stars, yet they may be more sensitive to language style, cultural norms, and expectations, introducing systematic differences in sentiment expression. If platforms use text embeddings to rank sellers, they may inadvertently couple exposure to linguistic patterns correlated with group membership rather than quality. The bias framework extends by treating text features as additional noisy measurements with their own selection and expectation effects. Mechanism design can mitigate this by calibrating text-based scores against objective outcomes, or by limiting their use in exposure decisions.

Conclusion

This paper developed a technical framework for quantifying feedback-loop bias in two-sided marketplaces where ratings, reviews, and reputation mechanisms interact with ranking and exposure policies. The core idea is that reputational signals are not merely measurements of latent quality but endogenous objects shaped by the policy rules that use them. When exposure depends on reputation and reputation depends on the stream of interactions produced by exposure, stochastic early variation and endogenous rater composition can be amplified into persistent differences in visibility and outcomes. Feedback-loop bias was defined as the difference in expected outcomes between the deployed coupled policy and a benchmark policy that weakens coupling, with outcomes including exposure-weighted reputational error, allocative efficiency, and welfare proxies.

A decomposition highlighted several channels: selection on reputational noise through convex exposure mappings, endogenous rater composition and expectation effects that shift rating behavior as a function of displayed reputation, and nonlinear aggregation and thresholding that introduce discontinuities and absorbing states. Identification strategies were described for both experimental and observational settings, including randomized ranking perturbations, discontinuity-based designs around policy thresholds, hierarchical latent-quality models that incorporate rating propensity, and simulation-based counterfactual evaluation. The distributional dynamics induced by state-dependent learning were connected to Fokker–Planck equations, motivating numerical approaches such as finite element discretization and Monte Carlo simulation with variance reduction. Additional diagnostics such as frequency spectrum analysis and logarithmic signal-to-noise measures were incorporated to characterize stability and learning under alternative policies.

Mechanism design implications were framed as optimization under bias and disparity constraints, emphasizing exploration-exploitation trade-offs, uncertainty-aware display and aggregation, and the risks introduced by hard thresholds and by uncalibrated text-based reputation features. The overall result is a set of estimands and computational tools for evaluating when and how reputation systems may generate bias through feedback, and for comparing policy alternatives in terms of efficiency, learning, and distributional outcomes without presuming a universally optimal design [27].

Conflict of interest

Authors state no conflict of interest.

References

[1] A. Sisto and C. Kamath, *Ensemble feature selection in scientific data analysis*, Sep. 20, 2013. DOI: 10.2172/1097710

[2] M. Perc, M. Ozer, and J. Hojnik, "Social and juristic challenges of artificial intelligence," *Palgrave Communications*, vol. 5, no. 1, pp. 1–7, Jun. 25, 2019. DOI: 10.1057/s41599-019-0278-x

[3] G. Choy et al., "Current applications and future impact of machine learning in radiology.," *Radiology*, vol. 288, no. 2, pp. 318–328, Jun. 26, 2018. DOI: 10.1148/radiol.2018171820

[4] M. Bonart, A. Samokhina, G. Heisenberg, and P. Schaer, "An investigation of biases in web search engine query suggestions," *Online Information Review*, vol. 44, no. 2, pp. 365–381, Dec. 5, 2019. DOI: 10.1108/oir-11-2018-0341

[5] J. Velkova and A. Kaun, "Algorithmic resistance: Media practices and the politics of repair," *Information, Communication & Society*, vol. 24, no. 4, pp. 523–540, Aug. 26, 2019. DOI: 10.1080/1369118x.2019.1657162

[6] N. V. Patel, "Selection bias in two-sided e-commerce marketplaces: A framework for propensity score matching implementation," *Journal of Business Intelligence Systems and Computational Social Science Applications*, vol. 11, no. 5, pp. 1–20, 2021.

[7] K. Sakaguchi, R. L. Bras, C. Bhagavatula, and Y. Choi, *Winogrande: An adversarial winograd schema challenge at scale*, Jul. 24, 2019.

[8] M. Momeni and M. R. Saradjian, "Evaluating ndvi-based emissivities of modis bands 31 and 32 using emissivities derived by day/night lst algorithm," *Remote Sensing of Environment*, vol. 106, no. 2, pp. 190–198, Jan. 30, 2007. DOI: 10.1016/j.rse.2006.08.005

[9] R. Singh, H. Kumar, and R. Singla, "Sampling based approaches to handle imbalances in network traffic dataset for machine learning techniques," in *Computer Science & Information Technology (CS & IT)*, Academy & Industry Research Collaboration Center (AIRCC), Jun. 1, 2013, pp. 37–48. DOI: 10.5121/csit.2013.3704

[10] G. S. Nelson, "Bias in artificial intelligence.," *North Carolina medical journal*, vol. 80, no. 4, pp. 220–222, Jul. 5, 2019. DOI: 10.18043/ncm.80.4.220

[11] D. Guo, S. I. Ktena, F. Huszar, P. K. Myana, W. Shi, and A. Tejani, *Deep bayesian bandits: Exploring in online personalized recommendations*, Aug. 3, 2020.

[12] M. Moore, K. Ng, Y. Kim, K. Robinson, M. Thompson, and J. Reich, *Debriefing digital simulations: Exploring designs that support teacher educators in using teacher candidate performance data*, Feb. 12, 2019. DOI: 10.31235/osf.io/pwqb8

[13] T. Kehrenberg, Z. Chen, and N. Quadrianto, "Tuning fairness by balancing target labels," *Frontiers in artificial intelligence*, vol. 3, pp. 33–33, May 12, 2020. DOI: 10.3389/frai.2020.00033

[14] J. Chen, "Reinforcement machine learning robots with the reset policy and leptokurtic noise," *SSRN Electronic Journal*, Jan. 1, 2017. DOI: 10.2139/ssrn.3637255

[15] K. N. Johnson, *Automating the risk of bias*, Nov. 14, 2019.

[16] S. Strauß, "From big data to deep learning: A leap towards strong ai or 'intelligentia obscura'?" *Big Data and Cognitive Computing*, vol. 2, no. 3, pp. 16–, Jul. 17, 2018. DOI: 10.3390/bdcc2030016

[17] T. Kihara, R. Bendor, and D. Lomas, "Chi extended abstracts - designing an escape room in the city for public engagement with ai-enhanced surveillance," in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, ACM, May 2, 2019, pp. 1–6. DOI: 10.1145/3290607.3313003

[18] C. Blex and T. Yasseri, "Positive algorithmic bias cannot stop fragmentation in homophilic networks," *The Journal of Mathematical Sociology*, vol. 46, no. 1, pp. 1–18, Sep. 13, 2020. DOI: 10.1080/0022250x.2020.1818078

[19] Y. Wang, "Guest editorial: Special issue on artificial intelligence, machine learning, data science, and libraries," *International Journal of Librarianship*, vol. 5, no. 1, pp. 1–3, Jul. 23, 2020. DOI: 10.23974/ijol.2020.vol5.1.166

[20] H. Boon-Peng and K. Yusoff, "The utility of copy number variation (cnv) in studies of hypertension-related left ventricular hypertrophy (lvh): Rationale, potential and challenges," *Molecular cytogenetics*, vol. 6, no. 1, pp. 8–8, Mar. 1, 2013. DOI: 10.1186/1755-8166-6-8

[21] T. Sühr, S. Hilgard, and H. Lakkaraju, *Does fair ranking improve minority outcomes? understanding the interplay of human and algorithmic biases in online hiring*, Dec. 1, 2020.

[22] E. J. Gerritse, F. Hasibi, and A. P. de Vries, "Ictir - bias in conversational search: The double-edged sword of the personalized knowledge graph," in *Proceedings of the 2020 ACM SIGIR on International Conference on Theory of Information Retrieval*, ACM, Sep. 14, 2020, pp. 133–136. DOI: 10.1145/3409256.3409834

[23] M. Saers and D. Wu, "Lata - handling ties correctly and efficiently in viterbi training using the viterbi semiring," in Germany: Springer International Publishing, Mar. 8, 2018, pp. 284–295. DOI: 10.1007/978-3-319-77313-1_22

[24] N. A. Obuchowski et al., "Quantitative imaging biomarkers: A review of statistical methods for computer algorithm comparisons," *Statistical methods in medical research*, vol. 24, no. 1, pp. 68–106, Jun. 11, 2014. DOI: 10.1177/0962280214537390

[25] E. Magrani, "Governance of internet of things and ethics of artificial intelligence," *Revista Direitos Culturais*, vol. 13, no. 31, pp. 153–190, Dec. 10, 2018. DOI: 10.20912/rdc.v13i31.2816

[26] K.-C. Yang, O. Varol, C. A. Davis, E. Ferrara, A. Flammini, and F. Menczer, "Arming the public with artificial intelligence to counter social bots," *Human Behavior and Emerging Technologies*, vol. 1, no. 1, pp. 48–61, Feb. 6, 2019. DOI: 10.1002/hbe2.115

[27] T. Takahashi, T. Inagaki, S. Nariai, J. Kodama, M. Arakawa, and Y. Ema, "A calculation method of deposition profiles in chemical vapor deposition reactors using bio-inspired algorithms," *physica status solidi c*, vol. 12, no. 7, pp. 908–911, May 7, 2015. DOI: 10.1002/pssc.201510056