ISSN (Print) 0473-453X
Discussion Paper No. 1203
ISSN (Online) 2435-0982

SOCIAL LEARNING AND STRATEGIC PRICING WITH RATING SYSTEMS

Chia-Hui Chen
Kong-Pin Chen
Junichiro Ishida

March 2023

The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

Social Learning and Strategic Pricing with Rating Systems

Chia-Hui Chen* Kong-Pin Chen† Junichiro Ishida‡

March 7, 2023

Abstract

Despite widespread use in online transactions, rating systems only provide summary statistics of buyers' diverse opinions at best. To investigate the consequences of this coarse form of information aggregation, we consider a dynamic lemons market in which buyers share their evaluations anonymously through a rating system. When the buyers have diverse preferences, the value of a good rating depends endogenously on the seller's pricing strategy, which in turn creates complicated dynamic interactions and results in stochastic price fluctuations. Occasional flash sales induced by the rating system yield a non-trivial welfare effect that stands in sharp contrast to standard adverse selection models: all buyers are weakly better off with information asymmetry than without. Incentivizing buyers to leave ratings may backfire by exacerbating the seller's strategic pricing incentives.

Keywords: online platform; rating system; anonymity; preference diversity; price fluctuation; dynamic adverse selection

JEL Classification: D82; D83; L11

*Institute of Economic Research, Kyoto University, Japan. E-mail: chchen@kier.kyoto-u.ac.jp.
†Institute of Economics, Academia Sinica, Taiwan. E-mail: kongpin@gate.sinica.edu.tw.
‡Institute of Social and Economic Research, Osaka University, Japan. E-mail: jishida@iser.osaka-u.ac.jp.

1 Introduction

The development of platform transactions has provided a channel for buyers to learn the quality of a seller that is very different from traditional modes of learning such as word-of-mouth communication and observational learning. Since transactions are centralized on a platform, a standardized reputation mechanism, which we broadly refer to as the "rating system," can be designed to streamline the transmission of information. Under a rating system, buyers can post their "ratings" for the products or sellers based on their experiences. Potential buyers can observe the ratings, or a summary of them, left by previous buyers before they make purchasing decisions. This trading environment constitutes a special form of social learning, which enables potential buyers to collect information on a much wider scale than they could through word-of-mouth communication and also to gain access to more detailed information—past buyers' evaluations rather than their actions—than through observational learning.

Rating systems are, however, subject to substantial communication frictions. There are several possible reasons for this,[1] but the one that stands out is the fact that under a rating system, buyers share their evaluations anonymously without knowing the identities, much less the preferences, of the raters. This inherently anonymous nature of rating systems creates a complication when the buyers have diverse preferences, as they may be satisfied for different reasons, and hence what a "good rating" implies could differ considerably across them. Particularly relevant in this regard is the difference in the way buyers evaluate the tradeoff between price and quality, where some are more sensitive to quality while others are more sensitive to price. This structure suggests a possibility that a seller—even a low-quality one—can induce good ratings from those who are more concerned about price by deliberately underpricing and temporarily boost her reputation. Under this structure, the pricing strategy of a seller plays a dual role: on one hand, it plays the conventional role of rent extraction; on the other hand, it also influences the buyer's rating directly and hence the beliefs of subsequent buyers and future revenues indirectly. The seller's pricing decisions must strategically balance the tradeoff between current and future revenues as dictated by the rating system.

In this paper, we develop a dynamic adverse selection model to examine this strategic incentive and its consequences on trading dynamics and welfare. The trading environment we

[1] As we will later argue in Section 6.1, it is also important that platforms must in principle rely on the voluntary participation of buyers: although some platforms offer small rewards for providing a rating, the size of the rewards is typically quite negligible, and it is in general difficult for platforms to induce buyers to exchange detailed information in an incentive-compatible manner. There are actually two sides to this limitation. On one hand, it is highly time-consuming for any buyer to provide a detailed account of how and in what sense he is satisfied. On the other hand, even if some previous buyers are somehow willing to provide such detailed reports, current buyers often do not bother to go through a pile of reports to extract relevant information.

consider consists of a long-lived seller and a sequence of short-lived buyers who arrive randomly over time. The quality of the good supplied by the seller is either high or low and is her private information. A crucial assumption of our analysis is that the buyers are also heterogeneous and have diverse preferences in that they are either price shoppers (who only care about price) or value shoppers (who also appreciate good quality). Upon arrival, a buyer observes the history of ratings left by previous buyers and the current price offer and decides whether to purchase the good given his belief.[2] If he chooses to purchase the good, he learns its quality after consumption and leaves a good rating (with some positive probability) if he is "satisfied."

Within this setup, we argue that any rating system, which aggregates diverse opinions of previous buyers into some summary statistics, entails an unintended consequence of inducing strategic behaviors from the sellers. The driving force behind our argument is the combination of anonymity and preference diversity (buyer heterogeneity) that is inevitably present in platform transactions. In the presence of these two factors, the rating system can only provide a crude indicator of quality because there are no cohesive standards among buyers for when to give a good rating: specifically, while value shoppers are satisfied only by high quality, there are also price shoppers who are satisfied by low price; as such, good ratings provided by price shoppers do not necessarily indicate high quality. The buyers can observe ratings scattered over time and the current price but have no means to differentiate informative signals (good ratings from value shoppers) and uninformative ones (those from price shoppers) due to the anonymous nature of rating systems. This implies that the value of a good rating is endogenous in our model, depending on the seller's pricing strategy at the time of purchase, which marks a crucial point of departure from dynamic adverse selection models with exogenous news such as Daley and Green (2012) and Kaya and Kim (2018).

Throughout the analysis, we focus on a natural class of equilibria, which we call *monotone equilibria*, and derive properties that hold robustly within this class. This class of equilibria exhibits an intuitive structure where both types of seller charge a higher price to extract surplus when the belief is relatively high, while they hold a flash sale to attract a broader customer base and garner good ratings when it is low. The seller's pricing strategy thus repeatedly and stochastically fluctuates between a high and a low price, giving rise to trading dynamics induced by the rating system. When the seller charges a low price, price shoppers are satisfied and leave good ratings which convey no information about the seller's quality. This strategic pricing introduces noise into the buyers' inference process and consequently slows down social learning. This finding is closely in line with Fan et al. (2016) who find that sellers tend to

[2] We therefore consider a posted price platform. Although eBay has made auctions with rating systems popular in the C2C environment, nowadays most of the items listed in eBay (or any other transaction platforms) are in posted prices (Einav et al., 2018).

underprice to acquire good ratings when their reputation is low, which incidentally indicates that price shoppers indeed exist and play a major role in rating systems.

All monotone equilibria of our model exhibit two absorbing states. At the upper absorbing state, the high-quality seller is identified as such, and the belief stays constant thereafter. Moreover, by belief consistency, only the high-quality seller can reach the upper absorbing state on the equilibrium path, meaning that the rating system we consider displays no false negative. In contrast, at the lower absorbing state, value shoppers do not have enough confidence in the seller and never purchase. Once the belief reaches this point, therefore, any good rating could only come from price shoppers and its reputation value is entirely lost. Since the low-quality seller cannot reach the upper absorbing state, she will eventually be trapped into the lower absorbing state. The problem is that the high-quality seller also has a positive probability of reaching the lower absorbing state, implying that the rating system may exhibit a false positive. This result suggests that in addition to transaction delays, noisy social learning entails a social cost that remains strictly positive even as the discount rate vanishes. To measure the extent of this efficiency loss, we explicitly compute the asymptotic probability of identifying the high-quality seller as such. Remarkably, the asymptotic probability is constant across *all* monotone equilibria of our model and depends only on two factors: the seller's initial reputation and the sensitivity of the value shoppers to the good's quality.[3]

It is worth noting, however, that despite this efficiency loss, distributional consequences of noisy social learning are not uniformly negative across all parties involved. In fact, the seller's strategic pricing incentive yields a non-trivial welfare effect that stands in sharp contrast to standard adverse selection models: the value shoppers are strictly better off with information asymmetry than without. This result stems from a crucial feature of our model that the seller's information manipulation occurs through her pricing decisions. To see this, if the seller is known to be of high quality for sure, there is no point in offering a low price as the seller only loses her current revenue without gaining any reputation. In this case, the seller always charges a price that equals the expected quality, and the value shoppers can capture no surplus. The situation changes substantially once there is uncertainty about the seller's type: when the belief is close to the lower absorbing state, even the high-quality seller is forced to salvage her reputation, which allows some value shoppers to capture positive surplus. Interestingly, an important qualifier of this result is that there are some price shoppers around, because the frequency of receiving a good rating would not depend on the current price otherwise, i.e., preference diversity benefits the value shoppers.

We also identify the seller-optimal equilibrium and examine how its efficiency is affected by

[3] To be more precise, the asymptotic probability is constant across all pooling equilibria of our model which constitute a larger set of equilibria including some non-monotone ones.

changes in some underlying parameters of the model. We find that as the seller becomes less patient, the efficiency of social learning improves. This is because when the seller discounts the future payoffs more heavily, there is less incentive to offer a low price to garner good ratings. This result suggests a counterintuitive possibility that the high-quality seller is often made better off, while the value shoppers are worse off, as the market environment becomes more competitive with higher market exit rates. Also, when satisfied buyers leave ratings at a higher frequency, the seller's strategic incentive might intensify and slow down social learning. This means that although many platforms adopt various measures to encourage buyers to leave ratings, such an attempt might backfire and actually impede social learning.

A distinctive feature of transactions mediated by platforms is that the platforms typically have some control over the flow of information. In the baseline model, we consider a rating system with minimum information content, where a buyer can only leave a good rating when he is satisfied. To examine the impact of providing more information on trading dynamics, we extend this rating system in two different directions. In the first, we make the history of prices publicly observable, on top of the history of ratings. In the second, we allow the buyers to leave bad ratings when they are unsatisfied. We argue that the first extension may drastically improve the efficiency of social learning while the second has little qualitative impact.

2 Literature

Our analysis belongs to the vast literature on dynamic adverse selection and reputation building. Although the literature has developed in various directions, there are very few works which incorporate strategic pricing into this framework.[4] As we will detail below, the key departures of our analysis from the previous literature are summarized by the following two elements: (i) both types of seller are strategic and attempt to maximize their payoffs; and (ii) they strategically set prices to manipulate the flow of information to their advantage.

Starting with Milgrom and Roberts (1982), Kreps and Wilson (1982), and Fudenberg and Levine (1989), many reputation models build on the framework where one long-run player interacts with a sequence of myopic short-run players, as in our trading environment.[5] This strand of literature assumes that the long-run player is either a strategic type, who maximizes his own payoffs, or a commitment type, who simply follows a pre-specified plan of action.

[4] An obvious technical challenge is that if one allows the seller to set a price, it works as a signal and hence changes the nature of the problem. For this reason, adverse selection models in trading environments typically consider buyer-offer games, where the buyer makes an offer and the seller either accepts or rejects.

[5] See Liu (2011), Liu and Skrzypacz (2014), and Pei (2020, 2022) for more recent examples of this approach. A continuous-time version of this strand of literature is developed by Faingold and Sannikov (2011).

While earlier models of this strand of literature look at the situation where all the past actions are perfectly observable to the current players, subsequent works extend this setting and incorporate imperfect observability of past actions (Fudenberg and Levine, 1992; Cripps et al., 2004). Among them, Ekmekci (2011) introduces a "rating system" which aggregates all the observable variables into a message in an arbitrary way and shows that a finite rating system with information censoring allows the long-run player to perpetually sustain a reputation.[6] The focus of this literature is different from ours, as it is generally concerned with whether the strategic type can approximately achieve his best feasible payoff, often called the Stackelberg payoff in the literature, when there are some commitment types around.

In our model, the seller strategically sets prices to interfere with the buyers' inference process—a strategy that may be broadly referred to as "signal jamming" (Holmstrom, 1999; Fudenberg and Tirole, 1986).[7] Related in this respect are Ekmekci et al. (2022), who analyze a dynamic stopping problem with manipulable signals. They consider a situation where a principal decides whether to terminate the game when an opportunity arrives stochastically while continuously receiving a signal about an agent's type. The agent in their model is either strategic (called noninvestible in their model) or committed (investible), where the strategic type chooses how much effort to exert at each instance to manipulate the signal-generating process. As noted above, a crucial point of departure is that both types of seller in our model are strategic. Aside from this, it is also important to note that the seller's information manipulation in our model occurs through pricing choices and hence affects not only the learning process but also the distribution of surplus directly. As we will discuss in more detail in Section 5.2, these elements of our model generate different forms of dynamic interaction and different welfare implications.

Several works analyze dynamic adverse selection models coupled with additional information sources. Daley and Green (2012) consider a dynamic lemons market where the value of a long-run seller's asset is gradually and publicly revealed to the market via some news process and characterize trading dynamics in this environment. Kaya and Kim (2018) consider a similar setup but instead introduce a news process that provides a private signal to each potential buyer and show that the private nature of the news process gives rise to diverse patterns of trading dynamics. These models share an important feature with our model that buyers have access to an additional, albeit imperfect, information source. A key distinction is that the news

[6] This is in contrast to Cripps et al. (2004) who are the first to analyze the long-run consequences of reputation building and point out that it is impossible to sustain a reputation in the long run under imperfect monitoring.

[7] In our analysis, signal jamming works in a slightly different way. As in Holmstrom (1999) and Fudenberg and Tirole (1986), the buyers correctly anticipate this manipulation, but it can still affect the way they aggregate information because strategic pricing effectively introduces noise into the rating system.

process in their models is exogenous whereas the rating system in our model is endogenous.[8] Because of this feature, asymptotic learning fails in our model as even the high-quality seller is driven out of the market with some probability.

Finally, while our analysis centers around the seller's strategic incentive to influence the flow of information, there are some works that examine the role of rating systems with more direct focus on the process of buyer learning and its asymptotic properties (Crapis et al., 2017; Ifrach et al., 2019; Acemoglu et al., 2022). These works consider a rating system similar to ours, where a buyer leaves a good rating if his utility is above some threshold with preference heterogeneity among buyers, but abstract away from information asymmetry and the seller's strategic behaviors. Most notably, Acemoglu et al. (2022) construct a symmetric-information model of Bayesian learning from online reviews and investigate the speed of learning under different rating systems with an exogenously fixed price.[9] Their analysis identifies a selection effect which arises from the fact that the types of user who purchase the good depend on the information available at the time of purchase. Although cast in a different framework, this consideration is also important in our analysis because different pricing strategies attract different types of buyer, which in turn affects the value of a good rating.

3 Model

3.1 Setup

Consider a dynamic lemons market with a long-lived seller and a sequence of short-lived buyers who arrive randomly over time. Time is continuous and extends from 0 to infinity.

Seller. The seller, who has an unlimited supply of a good, lists it on a platform for sales by posted price. The seller is of fixed quality, which is either high ($i = H$) or low ($i = L$). The seller's type is her private information, and the common initial belief that the seller is of high quality is q_0. At each t, the seller posts a price p_t, where the cost of production is normalized

[8] Aside from the exogenous news process, another crucial difference is that the seller in their setting only has a single indivisible good, so that the seller of an asset can signal her type by passing on transaction opportunities. Signaling thus takes place in a different form. There are now several works which incorporate signaling via exit decisions in dynamic adverse-selection models (Bar-Isaac, 2003; Gul and Pesendorfer, 2012; Chen and Ishida, 2018).

[9] Acemoglu et al. (2022) later extend their analysis and briefly consider the case with endogenous pricing. Ifrach et al. (2019) also consider a similar symmetric-information environment, both with and without endogenous pricing. In either case, however, strategic interactions between the seller and buyers are not their main focus. Crapis et al. (2017) analyze optimal monopoly pricing in an asymmetric-information environment with non-Bayesian information updating, but restrict their attention to the case where the monopolist can adjust the price at most once.

to 0 for both types. The seller maximizes the discounted sum of profits with discount rate $r > 0$.

Buyers. The buyers have heterogeneous preferences, and hence what a good rating embodies is potentially different across them. To capture this aspect in the simplest manner, we assume that each arriving buyer is either a price shopper or a value shopper. Let $u^P(i,p)$ ($u^V(i,p)$) be the utility of a price (value) shopper when he purchases from the type i seller at price p. The price shoppers do not care about quality, and we normalize their utility from consumption to 0, i.e., for both $i = L, H$,

$$u^P(i,p) = -p.$$

In contrast, the value shoppers appreciate quality, and their utility is given by

$$u^V(i,p) = \begin{cases} 1-p & \text{if } i = H, \\ -k-p & \text{if } i = L, \end{cases}$$

where $k > 0$ is a parameter that measures the degree to which they value quality.[10] Each type of buyer arrives according to a Poisson process: for an interval of time $[t, t + \mathrm{d}t)$, a price shopper arrives with probability $m\mathrm{d}t$, and a value shopper with probability $\mathrm{d}t$. Each arriving buyer is myopic and exits the market immediately, regardless of whether he makes a purchase or not. In what follows, we refer to a buyer who arrives at time t as buyer t for clarity.

Rating system. Upon observing the price p_t, buyer t, if any, decides whether to make a purchase given his belief about the seller. The true quality is discovered after consumption, and the buyer leaves a good rating with probability $\lambda \in (0,1]$ at time t if his utility is weakly positive.

Remark: In our model, the buyers can only provide good ratings when they are satisfied, but have no choice but to remain silent when they are not. This specification is motivated by the fact that ratings are predominantly positive in online platforms (Nosko and Tadelis, 2015; Tadelis, 2016; Zervas et al., 2021).[11] For most part, we focus on this simplest setting

[10] It is worth noting that works such as Crapis et al. (2017), Ifrach et al. (2019), and Acemoglu et al. (2022) also incorporate buyer heterogeneity, but the nature of buyer heterogeneity in these works qualitatively differs from that in the current setting. In their works, the heterogeneity term is additive, so that buyers differ in terms of the threshold to give a good rating. This is different from our setting where the heterogeneity term is multiplicative, so that buyers differ in terms of the sensitivity to quality.

[11] Fradkin et al. (2021) argue that this constitutes the largest source of inefficiency in the rating system. There are many conceivable reasons for why buyers are reluctant to leave bad ratings, e.g., for fear of retaliation, reciprocation or psychological discomfort of giving negative feedbacks. In this paper, we do not explore why

for clarity, which allows us to illuminate how the seller's pricing incentives are affected by the rating system in a transparent manner. We will later discuss the case where there are more rating categories in Section 6.2 and show that our main insight still holds even when the rating system becomes "finer."

3.2 Information and strategies

Buyer t observes all the past ratings left by previous buyers and the current price offer p_t. Observe that the ratings are publicly observable while the current price offer is only privately observable to the current buyer. Let h_t be the history of ratings (including the time at which each rating is left) up to time t. The public belief at time t is defined as the probability that the seller is of high quality conditional on h_t and is denoted by q_t. In addition, buyer t observes p_t and further updates his belief. We denote the private belief of buyer t by $\hat{q}_t$, which is the probability conditional on both h_t and p_t.

We look for a Markov perfect Bayesian equilibrium in pure strategies (hereafter, simply an equilibrium). Let $\pi^i : [0,1] \to \mathbb{R}_+$ be the strategy of the type i seller, which is a mapping from q_t to a non-negative price p_t.

Definition 1. *The two types of seller adopt a pooling strategy at q if $\pi^L(q) = \pi^H(q)$; otherwise, they adopt a separating strategy. An equilibrium is a pooling equilibrium if $\pi^L(q) = \pi^H(q)$ for all q and a separating equilibrium if $\pi^L(q) \neq \pi^H(q)$ for some q.*

If the two types adopt a pooling strategy at some q_t, the current price offer reveals no information and $\hat{q}_t = q_t$. It is important to note, however, that the implication of "pooling" in our model is qualitatively different from its conventional one and should thus be interpreted with caution: even if the seller reveals her type by a separating price offer, this information will not be transmitted to the future buyers, meaning that separation at this stage does not contribute to building up her reputation (see Section 3.3 for more on this point).

Since the problem faced by myopic buyers is rather straightforward, instead of defining their strategies formally, we simply take them as if they are a non-strategic information-generating process. First, a price shopper purchases the good and is satisfied if and only if the price is 0, independently of his belief $\hat{q}_t$. Second, given $\hat{q}_t$, a value shopper purchases the good if p_t is no greater than $p(\hat{q}_t) \equiv \hat{q}_t - k(1 - \hat{q}_t)$, and is satisfied if the seller is of high quality. As such, at any time t, a price shopper leaves a good rating with probability λ if $p_t = 0$, and a value

this is the case, but rather take this nature of ratings as given and examine its consequences. In fact, since positive feedbacks play a critical role for their success, platforms even try to impose costs for consumers to leave negative ratings (Zervas et al., 2021). There are also reports on how Airbnb was accused of deleting negative reviews (see, e.g., https://qz.com/1333242/airbnb-reviews/)

shopper does so if $p(\hat{q}_t) \geq p_t$ and the seller is of high quality, also with probability λ. Let $\gamma^i(p_t; \hat{q}_t)$, or $\gamma^i(p_t)$ for short, denote the rate at which the type i seller receives a good rating. If $\hat{q}_t \geq \underline{q} \equiv \frac{k}{1+k}$, we have

$$\gamma^L(p_t; \hat{q}_t) = \begin{cases} \lambda m & \text{for } p_t = 0, \\ 0 & \text{for } p_t \in (0, p(\hat{q}_t)], \end{cases} \qquad \gamma^H(p_t; \hat{q}_t) = \begin{cases} \lambda(1+m) & \text{for } p_t = 0, \\ \lambda & \text{for } p_t \in (0, p(\hat{q}_t)]. \end{cases}$$

If $\hat{q}_t < \underline{q}$, value shoppers would never purchase the good, so that $\gamma^H(0; \hat{q}_t) = \lambda m$ and $\gamma^H(p_t; \hat{q}_t) = 0$ for any $p_t > 0$.

3.3 The static signaling problem and equilibrium selection

Although our analytical focus is on the seller's strategic response to the rating system, our model also entails a static signaling problem at each instant, given that an informed party—the seller—makes an observable pricing choice. The signaling problem in our model is somewhat irregular, as it differs from a typical setting characterized by the single-crossing property. The most distinctive feature of our signaling environment is that it is static in the sense that the current price is observable only to the current buyer, and any information conveyed by the price is not transmitted to the future buyers. As such, separation at this stage does not contribute to improving her future reputation, and hence the high-quality seller has no strong incentive to separate from the low-quality one by pricing differently.

As typical in signaling models, our model admits a plethora of signaling equilibria, but standard refinement arguments cannot be applied directly to obtain a sharp prediction due to its structure. Although it is possible to enumerate different signaling equilibria and analyze trading dynamics for each possible case, such an exercise would be rather tedious and provide little economic insight because all equilibria more or less exhibit similar dynamic properties (see Sections 5.1 and 5.2 for more detail). In the subsequent analysis, therefore, we focus on a class of equilibria which we deem as the most reasonable and admits a simple equilibrium structure. Specifically, we restrict our attention to the following class of equilibria:

1. (Pooling equilibria) Both types of seller adopt a pooling strategy in equilibrium, i.e., $\pi^L(q) = \pi^H(q)$ for all q;

2. (Full surplus extraction) Whenever the seller offers a strictly positive price, she always offers the price that equals the maximum willingness-to-pay of a value shopper, both on and off the equilibrium path.

Below, we comment on some technical issues specific to our signaling environment and provide arguments for our equilibrium selection.

Pooling equilibria. Although one might be tempted to focus on separating equilibria, a close inspection reveals that the structure of our signaling environment is not conducive to separation. First, the static nature of the problem suggests that the seller has no incentive to build reputation by sacrificing current gains: separation is beneficial only to the extent that it induces the current buyer to purchase the good, and any reputation gains can be capitalized only by charging a higher price to the current buyer. This essentially rules out all separating strategies in which the low-quality seller offers a higher price than the high-quality seller. Given this, the high-quality seller must set a higher price to separate from the low-quality seller, but separation in this direction is clearly difficult to achieve as the low-quality seller can easily mimic. In Section 5.1, we formalize this intuition and show that pooling equilibrium is the only possible equilibrium if we focus on a natural and intuitive class of equilibria, which we call *monotone equilibria*, in which the value functions are monotonically increasing in the belief.

Full surplus extraction. Even if we restrict our attention to pooling equilibria, the multiplicity of equilibria still remains. To fix ideas, suppose that the seller chooses to offer a strictly positive price at some belief q_t, and the maximum willingness-to-pay of a value shopper is $p(q_t)$ in some pooling equilibrium (where $\hat{q}_t = q_t$). The value shopper is then willing to accept an offer as long as $p(q_t) \geq p_t$. This means that we may construct an arbitrary equilibrium in which $\pi^L(q_t) = \pi^H(q_t) = \eta p(q_t)$ for some $\eta < 1$.[12] We argue, however, that in our posted-price environment where the seller retains all the bargaining power, it is unlikely that an equilibrium in which the seller is forced to offer a suboptimal price survives in the long run.

4 Analysis

4.1 Preliminaries

The notion of full surplus extraction implies that the seller's choice at each instant is essentially binary, either to choose a low price $p_t = 0$ or a high price $p_t = p(q_t) > 0$. For clarity, we thus represent the seller's strategy by a simpler and more intuitive form, defined over this

[12] For instance, this can be done by assigning $\hat{q}_t = 0$ to all off-path deviations. Standard refinement arguments, such as the Intuitive Criterion and D1, have no bite in selecting among different pooling equilibria because both types benefit equally from deviating to any $p_t \in (\eta p(q_t), p(q_t)]$. To see this, consider an equilibrium in which $\pi(q_t) = \pi^L(q_t) = \pi^H(q_t) < p(q_t)$ for some q_t. Suppose that we observe a deviation to some $p_t \in (\pi(q_t), p(q_t)]$. For both types, this deviation is profitable if and only if the belief assigned to it is such that the current buyer accepts p_t, i.e., $\hat{q}_t$ is high enough to satisfy $p(\hat{q}_t) \geq p_t$.

binary choice with no direct reference to p_t. Let $\sigma^i : [0,1] \to \{0,1\}$ denote the (pure) strategy of the type i seller, where she offers the low price if $\sigma^i(q_t) = 0$ and the high price if $\sigma^i(q_t) = 1$, with the interpretation that the type i seller is "having a sale" at q_t if $\sigma^i(q_t) = 0$. In what follows, we refer to $\sigma^i(q_t) = 0$ and $\sigma^i(q_t) = 1$ as the low-price and the high-price policy, respectively. Define $\sigma \equiv (\sigma^L, \sigma^H)$ as the seller's strategy profile and $\sigma(q) \equiv (\sigma^L(q), \sigma^H(q)) \in \{0,1\}^2$ as the profile at belief q. We also denote by $s_t \in \{0,1\}$ the seller's actual pricing choice at time t (as observed by buyer t) where $s_t = \sigma^L(q_t) = \sigma^H(q_t)$ on the equilibrium path (in a pooling equilibrium). Let $\hat{q}_t^s$ be the private belief after observing $s_t = s \in \{0,1\}$.

Since the price cannot be negative, a value shopper purchases the good only if his belief is at least as high as $\underline{q}$. For $q_t \in [0, \underline{q})$, there is a continuation equilibrium in which both types of seller adopt the low-price policy.[13] In this region, only price shoppers would purchase, and hence any good ratings could only come from them. Since good ratings no longer reflect quality, there is no way for the seller to improve her reputation once q_t dips below $\underline{q}$. The seller's continuation payoff is then 0 for both types, and the game effectively ends at that point. We thus only define the seller's strategy over $Q \equiv [\underline{q}, 1]$ for simplicity, and say that the seller "exits the market" once q_t reaches $\underline{q}$. Note also that if the initial prior q_0 is below $\underline{q}$, no transactions take place on the equilibrium path. To rule out this trivial case, we make the following assumption.

Assumption 1. $q_0 > \underline{q}$.

Given our binary-choice framework and focus on pooling equilibria, the seller's equilibrium strategy can be expressed by two mutually exclusive belief sets (Q^l, Q^h), where $\sigma(q) = (0,0)$ if $q \in Q^l$ and $\sigma(q) = (1,1)$ if $q \in Q^h$, with the restriction that $Q^l \cup Q^h = Q$ and $Q^l \cap Q^h = \varnothing$. For clarity, we refer to Q^l and Q^h as the low-price and the high-price region, respectively. Note that Q^l and Q^h may or may not be connected. When both Q^l and Q^h are connected, we say that the strategy is a threshold strategy. Given this, we also say that an equilibrium is a *threshold equilibrium* if both types of seller adopt a threshold strategy.

4.2 Belief updating

With abuse of notation, we now redefine $\gamma^i(\cdot; \hat{q}_t)$, or $\gamma^i(\cdot)$ for short, in terms of the binary choice where for $s_t = s \in \{0,1\}$ and $\hat{q}_t^s \geq \underline{q}$,

$$\gamma^L(s; \hat{q}_t^s) = \lambda m(1-s), \;\; \gamma^H(s; \hat{q}_t^s) = \lambda + \lambda m(1-s).$$

[13] This constitutes an equilibrium if the belief after deviating to the high-price policy is low enough so that no buyers would purchase. Under Assumption 2 which we will postulate below, this is the unique continuation equilibrium within the class of equilibria under consideration.

When $\hat{q}_t^s < \underline{q}$, there is no difference between the low-price and the high-price policy, and $\gamma^i(s;\hat{q}_t^s) = \lambda m(1-s)$ for $i = L, H$ and $s = 0, 1$. Given the seller's strategy σ and the current belief q_t, the belief at the next instant is updated to $q_{t+\mathrm{d}t} = q_t^-(\sigma(q_t))$ if the seller receives no rating and jumps up to $q_{t+\mathrm{d}t} = q_t^+(\sigma(q_t))$ if she receives a good rating, where

$$q_t^-(\sigma(q_t)) = q_t - \left(\gamma^H(\sigma^H(q_t)) - \gamma^L(\sigma^L(q_t))\right) q_t(1-q_t)\mathrm{d}t, \tag{1}$$

$$q_t^+(\sigma(q_t)) = \frac{q_t\gamma^H(\sigma^H(q_t))}{q_t\gamma^H(\sigma^H(q_t)) + (1-q_t)\gamma^L(\sigma^L(q_t))}. \tag{2}$$

Observe that a good rating does not ensure high quality (because it might be left by a satisfied price shopper), and its absence does not indicate low quality (either because no buyer arrives or because a buyer arrives and is satisfied but fails to leave a rating). In what follows, we often suppress $\sigma(\cdot)$ and simply write q_t^- and q_t^+ whenever it is not confusing.

It is easy to see that if $\gamma^H(\sigma^H(q_t)) > \gamma^L(\sigma^L(q_t))$, the belief decreases continuously as long as the seller receives no good rating. For $q_t \geq \underline{q}$, this condition always holds in any pooling equilibrium because $\gamma^H(s) - \gamma^L(s) = \lambda$ for both $s = 0, 1$, which implies that the belief follows

$$\dot{q}_t = -\lambda q_t(1-q_t),$$

for all $q_t \in Q$ in the absence of good ratings. Once the seller receives a good rating, q_t jumps up to q_t^+. In particular, if $\sigma^L(q_t) = 1$, then $\gamma^L(1) = 0$ and q_t jumps up to 1 after receiving a good rating. In contrast, once q_t dips below the threshold $\underline{q}$, we have $\gamma^L(0) = \gamma^H(0) = \lambda m$, and the belief stays constant indefinitely as noted above. Both of these properties reflect an essential aspect of our model where the reputation value of a good rating is endogenous and depends on the seller's strategy profile at the time of purchase.

4.3 Seller's optimal strategy

Let $V^i(\cdot)$ denote the value function of the type i seller. For $q_t \in Q$, the value function is defined as

$$V^i(q_t) = \max_{s\in\{0,1\}} \left\{ p(\hat{q}_t^s)s\mathrm{d}t + e^{-r\mathrm{d}t}E[V^i(q_t + \mathrm{d}q_t) \mid \sigma, s] \right\}, \tag{3}$$

where

$$E[V^i(q_t + \mathrm{d}q_t) \mid \sigma, s] = \gamma^i(s)\mathrm{d}tV^i(q_t^+) + (1 - \gamma^i(s)\mathrm{d}t)V^i(q_t^-).$$

This formulation leads to an intuitive incentive compatibility constraint for each type. From (3), $\sigma^i(q_t) = 1$ if and only if

$$e^{-r\mathrm{d}t}E[V^i(q_t + \mathrm{d}q_t) \mid \sigma, s = 0] \leq p(\hat{q}_t^1)\mathrm{d}t + e^{-r\mathrm{d}t}E[V^i(q_t + \mathrm{d}q_t) \mid \sigma, s = 1].$$

In the limit, the condition above is reduced to

$$(\gamma^i(0) - \gamma^i(1))\Delta V^i(q_t) \leq p(\hat{q}_t^1), \tag{4}$$

where $\Delta V^i(q_t) \equiv V^i(q_t^+) - V^i(q_t)$ is the reputation value of a good rating. The left-hand side of (4) measures the future benefit brought by the low-price policy, and the right-hand side measures the loss of current revenue. Condition (4) provides an important characterization of the tradeoff faced by the seller in framing her pricing policy.

4.4 Equilibrium

There are two key observations which help us identify the equilibrium. First, since only the high-quality seller provides strictly positive benefits and is able to charge a strictly positive price (to value shoppers), both types of seller have an incentive to present themselves as the high-quality one. Second, since future buyers cannot distinguish good ratings from value shoppers (informative signals) and those from price shoppers (uninformative signals), this can be achieved by offering the low price and inducing good ratings from price shoppers. As is made clear in (4), this gives rise to the tradeoff between the low-price and the high-price policy. To balance this tradeoff, the seller may occasionally have a sale for $q_t > \underline{q}$, i.e., even when the willingness-to-pay of a value shopper is strictly positive. Note that the incentive to have a sale is intensified as the belief gets closer to the lower bound, because the seller is forced to exit the market once q_t gets below $\underline{q}$. These considerations suggest that there should exist some equilibria in which both types of seller adopt the high-price policy if and only if q_t is greater than some threshold $\overline{q} > \underline{q}$.

Consider a threshold strategy characterized by a pair of thresholds $(\underline{q}, \overline{q})$ such that

$$\sigma^H(q_t) = \sigma^L(q_t) = \begin{cases} 1 & \text{if } q_t \geq \overline{q}, \\ 0 & \text{if } \overline{q} > q_t \geq \underline{q}. \end{cases} \tag{5}$$

We will henceforth call $\underline{q}$ the lower threshold and $\overline{q}$ the upper threshold. Since the lower threshold is exogenously fixed at $\underline{q} = \frac{k}{1+k}$, each threshold equilibrium is uniquely characterized by the upper threshold $\overline{q}$. Under this class of strategies, we denote by $V^i(\cdot; \overline{q})$ the value function of the type i seller when the upper threshold is set at $\overline{q}$. In the case of threshold strategies, the belief sets (Q^l, Q^h) reduce to $Q^l = [\underline{q}, \overline{q})$ and $Q^h = [\overline{q}, 1]$.

The following proposition states that we can always construct an equilibrium by a threshold strategy described in (5). The class of threshold equilibria identified here is a natural one to fucus on because it contains all the key insights of our model in a simple and intuitive form.

Proposition 1. *For any given (λ, m, r), let $\overline{q}^*$ be the solution to*

$$\lambda m \left(\frac{1}{r} - V^L(\overline{q}^*; \overline{q}^*) \right) = p(\overline{q}^*), \tag{6}$$

if it exists for $\overline{q}^* < 1$; let $\overline{q}^* = 1$ if not. Then, for every $\overline{q} \in [\overline{q}^*, 1]$, there is a threshold equilibrium with upper threshold $\overline{q}$.

Proof. See Section 4.5 and Appendix A. □

Proposition 1 states that we can construct a continuum of threshold equilibria with any $\overline{q} \in [\overline{q}^*, 1]$. The threshold equilibria identified in Proposition 1 have a simple and intuitive structure: both types of seller adopt a bang-bang-type strategy by charging a positive price $p(q_t)$ if q_t is higher than $\overline{q}$, and having a sale if it is lower. What is especially important is the fact that both types of seller must adopt the low-price policy at the lower end of the belief space: when $q_t \in (\underline{q}, \overline{q})$, although the willingness-to-pay of a value shopper is still strictly positive, the seller nonetheless charges the low price to boost her reputation. Since the belief enters this region with positive probability for any prior $q_0 > \underline{q}$, occasional flash sales must occur with positive probability. The occurrence of flash sales is indeed a robust prediction of our model, as we will discuss later in Section 5.2.

The case when $\overline{q}^* = 1$ merits some remark. It occurs when either λ or m is large, or when r is small. When this happens, no tradeoff characterized by (4) exists, and both types of seller simply charge the low price for all $q_t \in Q$ in the unique equilibrium. As this case is rather trivial, we make the following assumption to focus our attention to more relevant cases.

Assumption 2. *There is an interior solution to (6) so that $\overline{q}^* < 1$.*

4.5 Sketch of the proof

Here, we provide a brief sketch of our equilibrium construction to develop some intuition and illustrate how the equilibrium looks; technical details are relegated to Appendix A. For the subsequent analysis, it is convenient to define $T(q, q')$ for any $q > q' \geq \underline{q}$ such that

$$q' = \frac{qe^{-\lambda T(q,q')}}{qe^{-\lambda T(q,q')} + 1 - q}. \tag{7}$$

In words, $T(q, q')$ is the length of time it takes for the belief to drop from q to q' in a pooling equilibrium when the seller receives no good rating.[14] Solving (7) for $T(q, q')$ gives

$$T(q, q') = \frac{1}{\lambda} \ln \frac{q(1 - q')}{q'(1 - q)}.$$

Also, for any $q \in Q^l$, define $\bar{t}$ as the length of time it takes for the belief to drop back to q when the seller receives a good rating now and no rating afterwards. Since the belief jumps up to $\frac{q(1+m)}{q(1+m)+(1-q)m}$ after receiving a good rating, $\bar{t}$ is the solution to

$$q = \frac{q(1 + m)e^{-\lambda \bar{t}}}{q(1 + m)e^{-\lambda \bar{t}} + (1 - q)m},$$

from which we obtain

$$\bar{t} = \frac{1}{\lambda} \ln \frac{1 + m}{m}.$$

Note that $\bar{t}$ is independent of the initial belief, implying that if the seller receives a good rating at time t and no rating afterwards, the belief goes back to q_t at time $t + \bar{t}$ regardless of the value of q_t (as long as $q_t \in Q^l$). This fact reflects a convenient property of Poisson process.

To illustrate how the equilibrium looks like, it is instructive to categorize the lower-price region Q^l into different sets, based on how many ratings the seller needs to enter the high-price region Q^h. From (1) and (2), we know that $q_t^- < q_t$ under a pooling strategy, and that there is a jump to q_t^+ when the seller receives a good rating. Given any threshold strategy described in (5) and any $q_t \in Q^l$, we define $Q_n \equiv [\underline{q}_n, \underline{q}_{n-1}), n = 1, 2, ...,$ recursively as follows. Let $Q_1 = [\underline{q}_1, \underline{q}_0)$, with $\underline{q}_0 = \bar{q}$ and $\underline{q}_1 = \max\{\underline{q}, \frac{\bar{q}m}{\bar{q}m+(1-\bar{q})(1+m)}\}$, where $\frac{\bar{q}m}{\bar{q}m+(1-\bar{q})(1+m)}$ is the value of q_t such that $q_t^+ = \bar{q}$. This means that for any $q_t \in Q_1$, the belief enters Q^h if the seller receives a good rating within a period of length $T(q_t, \underline{q}_1)$. For $n > 1$, if $\underline{q}_{n-1} > \underline{q}$, we recursively define $\underline{q}_n(q_t) = \max\{\underline{q}, \frac{\underline{q}_{n-1}m}{\underline{q}_{n-1}m+(1-\underline{q}_{n-1})(1+m)}\}$. Similarly, for any $q_t \in Q_n$, it takes n good ratings within a period of length $T(q_t, \underline{q}_n)$ for the belief to enter Q^h. We continue this process until we find some $\bar{n}$ such that $\max\{\underline{q}, \frac{\underline{q}_{\bar{n}-1}m}{\underline{q}_{\bar{n}-1}m+(1-\underline{q}_{\bar{n}-1})(1+m)}\} = \underline{q}$. For any $n \neq 1$, $Q_n \subseteq Q^l$ and both types of

[14] Note that $T(q, q')$ can be determined independently of the strategy σ in any pooling equilibrium because $\gamma^H(s) - \gamma^L(s) = \lambda$ for both $s = 0, 1$.

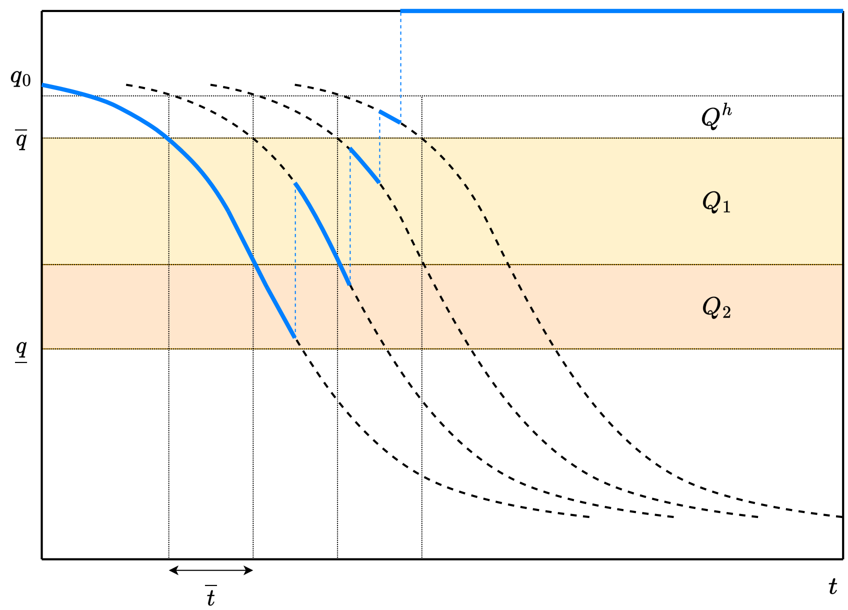

Figure 1: Construction of Q_n and sample belief path. The figure depicts a sample belief path of the high-quality seller. Each jump represents the arrival of a good rating. Once a good rating arrives in Q^h, the belief jumps up to 1 and stays there indefinitely.

seller adopt the low-price policy. Figure 2 plots the case with $\overline{n} = 2$, which provides a succinct view of how the belief evolves in equilibrium.

We now show by construction that any threshold strategy with $\overline{q} \geq \overline{q}^*$ constitutes an equilibrium. In the proof (in Appendix A), we prove that the following properties hold under this class of strategies.

Lemma 1. *$V^i(q; q)$ is increasing in q for all $q \in [\underline{q}, 1)$ and $i = H, L$.*

Lemma 2. *Given any $\overline{q} \in [\overline{q}^*, 1)$, for all $q \in [\underline{q}, 1)$, (i) $V^H(q; \overline{q}) - V^L(q; \overline{q}) > 0$, and (ii) $V^i(q; \overline{q})$ is increasing in q for $i = H, L$.*

Lemma 2 is easy to understand. It simply says that, for any given belief, the continuation payoff of the high-quality seller is always greater than that of the low-quality seller, and also that the continuation payoff is strictly increasing in the belief for both types. Lemma 1 is less obvious, but it says that the benefit of high reputation always dominates the potential loss of an increase in $\overline{q}$.

Given these results, we can verify that as long as $\overline{q} \geq \overline{q}^*$, any pair $(\underline{q}, \overline{q})$ can constitute the thresholds of an equilibrium. First, pick any $\overline{q} \geq \overline{q}^*$. For $q_t \in Q^l$, we need to show that it is the best response for both types of seller to adopt the low-price strategy. Suppose that the off-path belief after deviating to the high-price policy is low enough so that $\hat{q}_t^1 < \underline{q}$.[15] Then, from (4), the condition for adopting the low-price policy is given by

$$\lambda m \Delta V^L(q_t) > \max\{p(\hat{q}_t^1), 0\} = 0 \text{ and } \lambda \Delta V^H(q_t) > \max\{p(\hat{q}_t^1), 0\} = 0.$$

These conditions hold ince $\Delta V^i(q) > 0$ for all $q \in Q$ by Lemma 2. For $q_t \in Q^h$, to prevent downward deviation, it must be that

$$(\gamma^i(0) - \gamma^i(1)) \left(\frac{1}{r} - V^i(q_t; \overline{q})\right) \leq p(\overline{q}).$$

To show that this holds for the low-quality seller, note that

$$\lambda m \left(\frac{1}{r} - V^L(q_t; \overline{q})\right) < \lambda m \left(\frac{1}{r} - V^L(\overline{q}; \overline{q})\right) \leq \lambda m \left(\frac{1}{r} - V^L(\overline{q}^*; \overline{q}^*)\right) = p(\overline{q}^*) \leq p(\overline{q}) \leq p(q_t),$$

where the first inequality is by Lemma 2, and the second inequality is by Lemma 1. For the high-quality seller, since $\gamma^H(0; \hat{q}_t^0) - \gamma^H(1; \hat{q}_t^1) \leq \lambda m$,[16] we have

$$\lambda m \left(\frac{1}{r} - V^H(q_t; \overline{q})\right) < \lambda m \left(\frac{1}{r} - V^L(q_t; \overline{q})\right) \leq p(q_t),$$

which holds by Lemma 2. Again, it is optimal for both types to adopt the high-price policy when $q_t \in Q^h$.

5 Discussion

5.1 Monotone equilibria

Our model admits a plethora of equilibria, including some unrealistic ones, which requires us to adopt some restrictions to reduce the set of equilibria. One such restriction is our exclusive focus on pooling equilibria, which at a glance does not appear to have any solid foundation.

[15] As noted above, it is difficult to restrict off-path beliefs by the standard refinement arguments in our model. The current choice of off-path beliefs reflects an intuitive restriction that there are "no reputation gains from (upward) deviation." If the off-path belief $\hat{q}_t^1$ is very high for some reason, we cannot construct any monotone equilibrium in pure strategies. Since any equilibrium involving mixed strategies is substantially more complicated, we leave it for future research. We must emphasize, however, that the prospect of such an exercise is unclear since any monotone equilibrium must possess similar dynamic properties even if the seller randomizes in some intermediate range of the belief.

[16] The value of $\gamma^H(0; \hat{q}_t^0) - \gamma^H(1; \hat{q}_t^1)$ depends on the off-path belief assigned to a downward deviation: if $\hat{q}_t^0 \geq \underline{q}$, $\gamma^H(0; \hat{q}_t^0) - \gamma^H(1; \hat{q}_t^1) = \lambda m$; if not, $\gamma^H(0; \hat{q}_t^0) - \gamma^H(1; \hat{q}_t^1) = \lambda(m-1)$.

Before we discuss equilibrium properties of our model, we first deal with this issue and argue that our focus on pooling equilibria is hardly restrictive, as it can be derived from a more structural restriction. To this end, we consider an intuitive class of equilibria, which we call *monotone equilibria*, defined as follows.

Definition 2. *An equilibrium is a monotone equilibrium if the value function $V^i(\cdot)$ is monotonically increasing for both $i = L, H$.*

The notion of monotone equilibrium is a natural restriction to impose, as it captures an essential premise of all reputation models, i.e., higher reputation is always better for the seller. The following statement provides a structural foundation of pooling equilibria in our setting.

Proposition 2. *Any pure-strategy monotone equilibrium must be a pooling equilibrium.*

Proof. See Appendix A. □

It is out of the scope of the current analysis to explore whether there exists any pure-strategy separating equilibrium. Proposition 2 suggests, however, that even if such an equilibrium exists, it must possess an unappealing property of reputation non-monotonicity, i.e., the seller's continuation payoff decreases when her reputation improves at some point. It should be noted, though, that the set of monotone equilibria is still potentially very large, as Proposition 2 does not imply that all pure-strategy monotone equilibria have been identified in Proposition 1: for instance, we can construct an equilibrium in which Q^l and Q^h are disconnected. In what follows, we say that an equilibrium property holds robustly if it holds in all pure-strategy monotone equilibria of our model.

5.2 Rating systems and occasional flash sales

One of the main contentions of our analysis is that any rating system entails an unintended consequence of inducing strategic behaviors from sellers, which in turn introduces noise into the buyers' inference process and impedes social learning. In our model, the seller fluctuates stochastically between the two policies, and hence occasionally engages in a sale to boost her reputation. From the viewpoint of social learning, the existence of the low-price region creates an efficiency loss. To see this, suppose $\overline{q} = \underline{q}$, so that $\sigma(q) = (1, 1)$ for all $q \in Q$. Under this strategy profile, a good rating is always a sure sign of high quality, and the low-quality seller exits the market at the earliest timing $T(q_0, \underline{q})$. As we have seen above, however, this cannot happen in equilibrium, since the seller always has an incentive to switch to the low-price policy when the belief gets close enough to the lower threshold. Although we make this point under a set of restrictions, slow social learning due to the occurrence of occasional flash sales is a prediction that holds robustly in our setting.

Proposition 3. *In all pure-strategy monotone equilibria, $\sigma(q) = (0,0)$ for all $q < \overline{q}^*$. Moreover, in all pure-strategy monotone equilibria in which both types of seller offer the high price with positive probability, $\sigma(q) = (1,1)$ for q sufficiently close to 1.*

Proof. Let $V^i(\cdot)$ denote the value function of a monotone equilibrium. As we will show in Proposition 5, the threshold equilibrium with $\overline{q} = \overline{q}^*$ maximizes the low-quality seller's expected payoff, i.e., $V^L(q; \overline{q}^*) \geq V^L(q)$ for all $q \in Q$. This means that for all $q \in Q$, we have

$$\lambda m \left(\frac{1}{r} - V^L(q) \right) \geq \lambda m \left(\frac{1}{r} - V^{L*}(q) \right) > p(q),$$

which implies that the incentive compatibility constraint cannot be satisfied in any monotone equilibrium.

To show that the seller must adopt the high-price policy when the belief is sufficiently high, fix some $\varepsilon > 0$ such that $\sigma(q) = (0,0)$ for all $q \in (1-\varepsilon, 1)$. Then, in the limit, we have

$$\lim_{q \to 1} V^i(q) \leq \lim_{q \to 1} e^{-rT(q, 1-\varepsilon)} V^i(1-\varepsilon) = 0,$$

where the second term is the best-case payoff where the seller receives no good rating until q_t reaches $1 - \varepsilon$. This means that if there is such an ε, $V^i(\cdot)$ must decrease at some point, contradicting monotonicity. Therefore, the seller must adopt the high-price policy when the belief is sufficiently high in any monotone equilibrium in which she offers the high price with any positive probability. □

It is important to note that Proposition 3 is not a trivial restatement of Proposition 1: it says that the seller adopts the low-price policy when q_t is below $\overline{q}^*$ and the high-price policy when q_t is close to 1 in *any* pure-strategy monotone equilibrium in which the seller offers the high price with positive probability,[17] including those that do not belong to the class identified in Proposition 1. What is especially important is the fact that there is a threshold $\overline{q}^*$ below which the seller cannot adopt the high-price policy. To see this, suppose that there is a threshold equilibrium with $\overline{q} < \overline{q}^*$. In this equilibrium, the value of a good rating is extremely high when $q_t \in [\overline{q}, \overline{q}^*)$ because q_t jumps up to 1 after one good rating, which allows the seller to charge $p_t = 1$ indefinitely. More precisely, the value of a good rating in this range is given

[17] We can trivially construct an equilibrium in which the seller never charges a positive price. This equilibrium is perhaps of little interest because the seller's expected payoff is invariably 0, making it highly unlikely for such a market to survive.

by

$$\Delta V^i(q_t) = V^i(1) - V^i(q_t),$$

which is strictly decreasing in q_t in any monotone equilibrium by Lemma 2. By definition, therefore, we have

$$\Delta V^L(q_t) > \Delta V^L(\overline{q}^*) = p(\overline{q}^*) > p(q_t),$$

for any $q_t \in [\underline{q}, \overline{q}^*)$. In a nutshell, the incentive to deviate and (secretly) set a low price to garner a good rating is too strong at this point.

Observe that our model exhibits two absorbing states: $q_t = 1$ and $q_t = \underline{q}$. Once q_t reaches either one of them, no additional information will be revealed thereafter, and the belief stays constant over time. On the equilibrium path, the low-quality seller can never receive a good rating in Q^h: once q_t jumps above $\overline{q}$, it will gradually fall back. This means that the rating system never identifies a low-quality seller as a high-quality one, i.e, it never displays a false negative. On the other hand, there is a strictly positive probability that q_t reaches $\underline{q}$ for both types. Since the low-quality seller can never reach the upper absorbing state, it is easy to see that the low-quality seller will eventually be revealed as such as t tends to infinity. The problem is that there is also a strictly positive probability that the high-quality seller reaches the lower absorbing state and is forced to exit the market.

The fact that the rating system may display a false positive is the reason why even the high-quality seller needs to engage in a sale when the belief is closer to the lower threshold. The occurrence of flash sales by the high-quality seller generates an important welfare implication that stands in sharp contrast to standard adverse selection models. In standard adverse selection models, either static or dynamic, buyers usually cannot benefit from the presence of information asymmetry. This is not the case in our setting because information asymmetry induces the high-quality seller to hold a sale occasionally, which generates potential surplus for value shoppers. Since the expected payoff of a price shopper is always 0 by assumption, this result implies that all buyers are weakly better off with information asymmetry than without. A formal statement goes as follows.

Proposition 4. *The expected payoff of a value shopper is always 0 if* $q_0 = 1$ *(no asymmetric information), whereas it is strictly positive in all pure-strategy monotone equilibria if* $q_0 \in (\underline{q}, 1)$.

Proof. The proof is straightforward. Suppose that the seller is known to be of high quality, so that $q_t = 1$ for any history of the game. Then, there is clearly no incentive for the high-quality seller to offer the low price, and the expected payoff of a value shopper is always 0 in any equilibrium. If $q_0 < 1$, on the other hand, Proposition 3 states that the high-quality seller must offer $\sigma^H(q) = 0$ for $q \leq \overline{q}^*$. Since any arriving value shopper earns a payoff of $p(q_t)$ in this region and also there is a strictly positive probability that the belief enters this region, the expected payoff must be strictly positive in any monotone equilibrium with information asymmetry. □

Two elements of our model are essential for Propositions 3 and 4. First, it is important that both types of seller we consider are strategic. This is a departure from canonical reputation models which assume at least one commitment type and analyze how strategic types mimic the commitment type. If we applied this standard approach to our setting, we would take the high-quality seller as the commitment type, assuming that she always offers the high price, and analyze how the low-quality seller mimics the high-quality type. While such an approach makes the analysis more tractable and clarifies the distortion generated by the low-quality seller, it cannot describe how the high-quality seller would respond. Our assumption that both types are strategic allows us to capture dynamic interactions between the two types, which we believe are crucial for the problem we consider.

Second, it is also important that the rating system is an endogenous information-generating process in that the value of a good rating depends on the seller's strategy profile at the time of purchase. The contrast is clear with Daley and Green (2012) who consider an exogenous news process in a dynamic lemons market. In their model, there is a belief lower bound which serves as a reflecting barrier rather than an absorbing state. At the lower bound, low-quality sellers exit gradually (by accepting a low offer) while high-quality sellers never exit. There is no false positive in their setting because news arrives exogenously, and high-quality sellers can afford to wait for the arrival of good news even if the belief is approaching the lower bound. In our setting, on the other hand, once the belief reaches the lower threshold, the seller can only sell to price shoppers, and therefore has no chance to improve her reputation thereafter.[18]

In our model, the value of a good rating is endogenous because the rating system aggregates information anonymously from buyers with diverse preferences. The importance of anonymity is evident, as our entire argument rests on the fact that buyers cannot differentiate informative signals and uninformative ones. To see the role of preference diversity, note that if all buyers are price shoppers, both types of seller obviously cannot charge any positive price. If all buyers are value shoppers, on the other hand, there is no incentive to have a sale because the likelihood of receiving a good rating is unaffected in this case; as such, the seller is willing to have a sale only when there are buyers who are satisfied by low price. This suggests that the combination of anonymity and preference diversity is the crucial driving force of the trading

[18] Indeed, sellers with low ratings choose to leave the market (Fan et al., 2016).

dynamics (Proposition 3) as well as of the welfare implication (Proposition 4).

5.3 Seller optimality

As noted above, Proposition 1 does not exhaust all possible forms of pooling equilibrium. Moreover, even if we restrict attention to the class of equilibria identified in Proposition 1, there is still a continuum of equilibria. Those different equilibria obviously entail different efficiency properties, with some more favorable for the seller and others more favorable for the buyers. In this subsection, we propose a selection criterion in order to identify the most desirable equilibrium from the viewpoint of the high-quality seller.

Definition 3. *An equilibrium is the seller-optimal equilibrium if it maximizes the high-quality seller's expected payoff among all pure-strategy monotone equilibria.*

The notion of seller optimality is an important criterion for platforms when the seller's participation decision is endogenous: if the high-quality seller's expected payoff from some platform is not high enough, she may choose not to join the platform in the first place; in the worst case, the initial prior q_0 might dip below $\underline{q}$, which forces the platform to shut down entirely. The following result provides a characterization based on this notion.

Proposition 5. *The threshold equilibrium with $\overline{q} = \overline{q}^*$ is the seller-optimal equilibrium. Moreover, it also maximizes the low-quality seller's expected payoff.*

Proof. See Appendix A. □

Proposition 5 demonstrates the importance of Proposition 1 as it suggests that, if there is an equilibrium that does not belong to the class in Proposition 1, then we can identify one in that class that is more favorable to the seller. This result stems from the fact that the seller has an excessively strong incentive to deviate unilaterally to the low-price policy: as noted in Section 5.2, even if the high-price policy benefits the seller, she would be tempted to secretly undercut the price and garner a good rating in the high-price region. As a consequence, the set of beliefs in which the high-price policy benefits the seller is always larger and strictly contains the set of beliefs in which the incentive compatibility constraints are satisfied for both types. This observation implies that the seller-optimal equilibrium is the one in which the seller offers the high price as much as possible, i.e., the threshold equilibrium with $\overline{q} = \overline{q}^*$.

5.4 Comparative statics

We now conduct some comparative statics to derive more specific implications of our model. Since there is a continuum of equilibria, our comparative statics will focus on the seller-optimal equilibrium characterized in Proposition 5, i.e., we focus on how the value of $\overline{q}^*$ changes in response to changes in underlying parameters. Note that although we look at the seller-optimal equilibrium, the value of $\overline{q}^*$ itself has relevance beyond this particular equilibrium because this indicates the lower bound of all possible upper thresholds.

Proposition 6. *(i) There exists some $\underline{r} > 0$ such that $\overline{q}^* = 1$ for all $r < \underline{r}$ and $\overline{q}^* \to \underline{q}$ as $r \to \infty$, where $\overline{q}^*$ is strictly decreasing in r for $r \in [\underline{r}, \infty)$. (ii) There exists some $\overline{\lambda} > 0$ such that $\overline{q}^* = 1$ for all $\lambda > \overline{\lambda}$ and $\overline{q}^* \to \underline{q}$ as $\lambda \to 0$. (iii) There exists some $\overline{m} > 0$ such that $\overline{q}^* = 1$ for all $m > \overline{m}$ and $\overline{q}^* \to \underline{q}$ as $m \to 0$.*

Proof. We prove each part of the proposition in turn.

(i) Given some $q_t < 1$, the expected payoff at each $\tau > t$ is bounded away from 1. Therefore, as r increases, $\frac{1}{r} - V_L(\overline{q}^*)$ must decrease. To satisfy (6), $\overline{q}^*$ must decrease, which simultaneously increases the left-hand side of (6) and decreases the right-hand side. Note also that the right-hand side of (6) is bounded by 1, while the left-hand side goes to infinity as r goes to 0. Therefore, when r is close enough to 0, (6) does not have a solution, so that $\overline{q}^* = 1$ by definition. Finally, as $r \to \infty$, $\frac{1}{r} - V_L(\overline{q}^*)$ converges to 0, which implies that $\overline{q}^*$ must also converge to $\underline{q}$.

(ii) As λ approaches 0, since $\frac{1}{r} - V^L(\overline{q}^*)$ is bounded from above, the right-hand side of (6) must also go to 0. The solution $\overline{q}^*$ thus converges to the lower bound $\underline{q}$ as λ goes to 0. Also, when λ becomes sufficiently large, (6) does not have a solution, so that $\overline{q}^* = 1$ by definition.

(iii) The proof is essentially the same as in case (ii) and is omitted. □

The first result concerns the effect of the discount rate r. It is intuitively clear that as r increases (more discounting), the seller's preferences move towards the immediate revenue and away from the future gains. Correspondingly, there is less incentive to offer the low price, and the low-price region reduces. At the other end, when r is small enough, (6) does not admit a solution, and $\overline{q}^* = 1$. This presents an intriguing situation given that the expected payoff is 0 for both types in this equilibrium. Nevertheless, when $\overline{q}^* = 1$, it is impossible to construct a pooling equilibrium with $\overline{q} < 1$ because the incentive to (secretly) set the low price and garner a good rating is too strong, not just for the low-quality seller but also for the high-quality seller.

The discount rate r can be interpreted in different ways, ranging from the seller's patience to the (exogenous) probability of exiting the market; in principle, it can include all factors that make the seller more or less forward-looking. When the seller is more forward-looking, the frequency of sales increases, which generally hurts the seller but benefits the value shoppers as we have seen above. This suggests a counterintuitive possibility that the seller is made better off, while the buyers are worse off, in a more competitive market environment where the seller is driven out of the market with higher frequency.

A comparative statistics result that might be of particular interest to platform designers is the effect of λ, which measures the frequency of providing a rating after a transaction. Intuition suggests that an increase in λ would provide more information to buyers and expedite social learning; indeed, platforms usually adopt various measures to encourage buyers to leave ratings. As it turns out, though, this type of intervention can backfire in our setting: with a higher λ, the seller expects to receive more ratings in a given interval of time, which generates a similar effect to a decrease in r and raises $\overline{q}^*$. Table 1 shows this general tendency where an increase in λ induces an increase in $\overline{q}^*$.[19] This finding leads to a conclusion, also emphasized in other works such as Daley and Green (2012), Kaya and Kim (2018), Ekmekci et al. (2022), and Acemoglu et al. (2022), that more information does not necessarily improve learning in strategic environments.

λ	0.2	0.4	0.6	0.8
$\overline{q}^*$	0.53	0.70	0.84	0.90

Table 1: Effect of λ on $\overline{q}^*$

Finally, we also discuss the effect of m, which can be interpreted as measuring the extent of buyer heterogeneity. Proposition 6 suggests that there is an equilibrium in which the upper threshold $\overline{q}$ is arbitrarily close to $\underline{q}$ as the number of price shoppers goes to 0. To see this, consider the case of homogeneous buyers where $m = 0$ and all arriving buyers are value shoppers. Then, $\gamma^L(s) = 0$ and $\gamma^H(s) = \lambda$ for both $s = 0, 1$, and there is no benefit from having a sale. In equilibrium, therefore, $\sigma(q) = (1, 1)$ for all $q \in Q$, which obviously benefits the seller but hurts the buyers. This is a formal representation of the argument made in Section 5.2 that an important qualifier of Proposition 4 is $m > 0$, i.e., there is some preference diversity among buyers.

5.5 Buyer preferences and asymptotic learning

As noted above, one distinctive feature of our model is that the value of a good rating is endogenously determined: one good signal is enough to identify the high-quality seller when the belief is above $\overline{q}$ while it has no reputation value once it gets below $\underline{q}$. Because of this feature, even the high-quality seller may be forced to exit the market, i.e., social learning not

[19] The values in the table are computed by using the value functions derived in Appendix B. The parameter values used in the exercise are $r = 0.2$, $m = 0.3$, and $\underline{q} = \frac{1}{3}$.

only takes time but may also fail in the long run. Since only the high-quality seller can generate surplus, the likelihood of this event obviously yields a crucial welfare implication.

To measure the extent of this efficiency loss, let $P_t^i \equiv \mathbb{P}[q_t > \underline{q} \mid i]$ be the probability that the type i seller stays in the market at time t in a given equilibrium, and let $P^i \equiv \lim_{t\to\infty} P_t^i$. Since the low-quality seller never reaches the upper absorbing state, we clearly have $P^L = 0$, i.e., the low-quality seller is eventually driven out of the market. The asymptotic efficiency of learning is thus determined by P^H which we call the asymptotic probability of learning. Note that there are infinitely many belief paths that can be realized under each equilibrium strategy profile. The following result shows, however, that the asymptotic probability is independent of the seller's strategy and depends only on two parameters, q_0 and k, in any pooling equilibrium.

Proposition 7. *In all pooling equilibria, the asymptotic probability of leaning is*

$$P^H = 1 - \frac{\underline{q}(1-q_0)}{q_0(1-\underline{q})} = 1 - k\frac{1-q_0}{q_0}.$$

Proof. Although the low-quality seller exits the market at various points in time, the distribution of exiting times actually does not matter for the asymptotic probability. Fix some equilibrium σ^*, and let $f^i(t)$ be the probability of the type i seller exiting at time t under σ^*. Define a discrete set T of t such that $f^i(t) > 0$. Under this formulation, the asymptotic probability can be written as $P^i = 1 - \sum_{t\in T} f^i(t)$. Since the low-quality seller eventually exits at some point, we know $\sum_{t\in T} f^L(t) = 1$.

Observe that in any pooling equilibrium, both types of seller are driven out of the market if and only if q_t reaches $\underline{q}$. Therefore, for each $t \in T$, if the seller is forced to exit, we have

$$q_t = \underline{q} \;\Leftrightarrow\; \frac{q_0 f^H(t)}{q_0 f^H(t) + (1-q_0) f^L(t)} = \underline{q},$$

which can be written as

$$q_0(1-\underline{q}) f^H(t) = \underline{q}(1-q_0) f^L(t).$$

Summing them up for all $t \in T$ then gives

$$q_0(1-\underline{q})(1-P^H) = \underline{q}(1-q_0)(1-P^L) = \underline{q}(1-q_0).$$

Solving this for P^H yields the result. □

What is remarkable about this result is that the asymptotic probability of learning is constant across *all* pooling equilibria of this model. Observe that the asymptotic probability depends linearly on and is strictly decreasing in k. In other words, aside from the seller's initial reputation q_0, the asymptotic probability is determined by how selective the value shoppers are. If the value shoppers care little about quality so that k goes to 0, the rating system can almost surely identify the high-quality seller, despite its obvious limitations on information transmission. If they are very selective, on the other hand, learning opportunities are limited, and social learning fails asymptotically.

6 Extensions

In our baseline model, the buyers are only given coarse information, i.e., ratings which only indicate that some buyers are satisfied. What then happens if we allow the buyers to have access to more information? In this section, we discuss two possibilities—one in which the buyers can observe past prices (instead of only the current one) and the other in which the buyers are given the option to leave bad ratings—to address this issue and illuminate the impact of providing more information to buyers.

6.1 Observable price history

Suppose that in addition to the history of ratings, the buyers can also observe the history of prices, i.e., each buyer t can observe the whole history of prices $\{p_\tau\}_{\tau=0}^{t}$. Alternatively, since the price when there is no rating is irrelevant, we may simply assume that the price at the time of purchase is appended to each rating; as we show below, these two cases are actually equivalent. In our baseline model, given the current belief q_t, there are two possible belief realizations for the next instant, q_t^- and q_t^+, depending on whether a good rating is left or not. When the price information is appended to a rating, there are three possible belief realizations, depending not only on whether a good rating is left, but also on whether the price is high or low. We denote these new possibilities by q_t^{s+} which indicates the updated belief when the seller receives a good rating with $s_t = s$.

With this augmented rating system, we can construct an equilibrium in which the low-price policy is never adopted. The reason for this is that the current price is now observable to the future buyers as well, which suppresses the temptation to (secretly) undercut the price to obtain a good rating. In this equilibrium, price shoppers never purchase, and hence the learning process is as efficient as when there are no price shoppers (although this is detrimental to the value shoppers as we have discussed above).

Proposition 8. *Consider a rating system in which the price at the time of purchase is appended* *to each rating. Under this rating system, there exists an equilibrium in which $\sigma(q) = (1, 1)$ for all $q \in Q$.*

Proof. The proof is straightforward. Given that both types of seller always adopt the high-price policy, $s_t = 0$ is off the equilibrium path. For clarity, suppose we assign $\hat{q}_t^0 = 0$ to all deviations. We then have $q_t^{0+} = 0$ even if the seller receives a good rating.[20] This entirely eliminates any deviation incentive, because this information is now transmitted to the future buyers (in contrast to the case where the price can only be observed by the current buyer), and the seller is forced to exit the market immediately after receiving a good rating. □

Although not many platforms provide this information, there are actually several price-tracking tools available online such as CamelCamelCamel, Honey, and Keepa. Proposition 8 indicates that these price trackers not only help buyers find the best timing to make a purchase but also have the potential to improve social learning. More generally, the efficiency of social learning can be improved substantially by providing the "right information" to buyers.

We must note, however, that it could be a tall order to implement such an outcome in practice, because the augmented rating system requires far more cognitive effort from buyers. The argument circles back to our original premise that the rating system is subject to various communication frictions, which makes it difficult to convey detailed information to all buyers. In fact, the augmented rating system requires buyers to figure out how ratings and prices are correlated over time, whereas many buyers in reality do not bother to go into such details. It is also not entirely clear whether there exist any summary statistics that are intuitively and immediately understandable for all buyers. The simple rating system we consider has a definite advantage in this sense since at any t such that $q_t < 1$, a buyer simply needs to count how many good ratings the seller has received.[21]

6.2 Bad ratings

The previous subsection shows that a rating system which provides two-dimensional information can work quite well. Here, we enrich the rating system in a different way, by making it "finer" with more rating categories. Suppose that, in addition to leaving a good rating with

[20] Of course, this is a rather extreme, though simple, way to construct an equilibrium. A more conservative approach would be to suppose that the buyers believe that the two types deviate randomly with equal probability such that $\hat{q}_t^0 = q_t$. In this case, the two types receive a good rating with the same probability and hence $\hat{q}_t^0 = q_t$. Since the value of a good rating is 0, no type has an incentive to offer the low price.

[21] The fact that $q_t < 1$ implies that the seller has not received a good rating in the high-price region. Given this, the current belief depends only on the number of good ratings due to the memoryless property of exponential distribution.

probability λ^g when a buyer is satisfied, he can also leave a bad rating with probability λ^b if he is unsatisfied. We say that a buyer is unsatisfied when his utility from consumption is negative. Given this classification, a value shopper is unsatisfied when he purchases from the low-quality seller, regardless of the price, implying that only the low-quality seller can receive a bad rating. Therefore, after a bad rating is left, q_t drops down to 0, and the seller is forced to exit the market. On the other hand, if no rating is left, the belief is updated to

$$q_t^- = q_t - (\gamma^H(\sigma^H(q_t)) - \gamma^L(\sigma^L(q_t)) - \lambda^b)q_t(1-q_t)\mathrm{d}t,$$

where for $\hat{q}_t^s \geq \underline{q}$,

$$\gamma^L(s) = \lambda^g m(1-s),\ \gamma^H(s) = \lambda^g + \lambda^g m(1-s).$$

Notice a slight difference from (1), where we now have λ^b in the second term. If $\lambda^b < \lambda^g$, the belief would go down when there is no rating, but at a slower rate than in our baseline model where bad ratings are not allowed, i.e., $\lambda^b = 0$. In this case, the analysis remains essentially the same. The only notable difference is that the low-quality seller may receive a bad rating, which drives her out of the market immediately. Since this event occurs whenever the good is purchased by a value shopper, it is as if the low-quality seller is driven out of the market at a constant rate for some exogenous reasons. The possibility of bad ratings then yields an effect that is equivalent to an increase in the discount rate r, as long as she offers a price that is acceptable for value shoppers. Other than this, the tradeoff involved in setting a low price still works in the same way as before, and an equilibrium characterized by some upper threshold $\overline{q}$ still exists. Furthermore, Proposition 6 implies that the presence of bad ratings benefits the high-quality seller while it hurts the value shoppers.

If $\lambda^b \geq \lambda^g$, on the other hand, the belief increases as long as the seller receives no rating, i.e., "no news is good news." Unless the seller receives a bad rating, in which case q_t drops to 0, q_t will always stay above $\underline{q}$ for any $q_0 \geq \underline{q}$. This completely changes the incentive structure of the seller, and may produce equilibria that are very different from what we characterize in Proposition 1. However, though theoretically possible, it is highly unlikely that buyers leaves bad ratings with higher frequency than good ones in online platforms, given that reviews are predominantly positive (Nosko and Tadelis, 2015; Tadelis, 2016; Zervas et al., 2021); even at some intuitive level, it is hard to conceive of a situation where a seller establishes her reputation by having received no ratings. In all likelihood, λ^b may be positive but should lie somewhere between 0 and λ^g, so that "no news is bad news" is the appropriate specification in online platforms. It is nonetheless of some interest, from the purely theoretical point of view, to investigate this "no news is good news" case because this difference in the specification of learning often leads to markedly different dynamics.[22]

Finally, one may conjecture that the efficiency of social learning can be improved by extending the rating scale in the other direction, e.g., creating an additional category of "very good." Suppose that a buyer can leave a very good rating when he is "very satisfied," i.e., when his utility is above some positive threshold. For instance, in our setting, we may allow a value shopper to leave a very good rating when he purchases from the high-quality seller at the low price. While this rating system allows the high-quality seller to separate from the low-quality seller, and discourages the low-quality seller from having a sale, it does not fundamentally solve the problem at hand. In our model, we assume that the utility of a price shopper is always 0 simply for clarity, but the same problem persists if we introduce a third type, say a bargain shopper, who derives a very high utility regardless of the quality when he can purchase at the low price. Ultimately, this problem boils down to the fact that we cannot compare utilities across individuals. Different buyers not only have different preferences, with some emphasizing quality and others emphasizing price, but also have different evaluation standards—different rules to map a realized utility to a rating—and hence it is fundamentally infeasible for platforms to device a clear rule of what should constitute a good rating, no matter how elaborate the rating system becomes.

7 Conclusion

Under a rating system, buyers share their evaluations anonymously without knowing the identities of the raters. When the buyers have diverse preferences, this creates a room for strategic sellers to exploit. This paper explores the consequences of this inherently coarse nature of rating systems and its efficiency implications on social learning. In our model, both types of seller switch back and forth between the two pricing policies and occasionally hold a flash sale to garner good ratings. This form of strategic pricing introduces noise into the buyers' inference process and impedes social learning but at the same time generates a non-trivial welfare effect that stands in sharp contrast to standard adverse selection models: information asymmetry, when combined with preference heterogeneity among buyers, strictly benefits buyers who value quality. Since information is generated endogenously, social learning fails in the long run, as the high-quality seller is forced to exit the market after a string of bad luck. Incentivizing buyers to leave ratings exacerbates the seller's strategic pricing incentive and can be detrimental to social learning.

[22] See, for instance, Board and Meyer-ter-Vehn (2013) who show that reputation dynamics can be markedly different between these two cases.

References

Acemoglu, Daron, Ali Makhdoumi, Azarakhsh Malekian, and Asu Ozdaglar, "Learning from Reviews: The Selection Effect and the Speed of Learning," *Econometrica*, 2022, *90* (6), 2857–2899.

Bar-Isaac, Heski, "Reputation and Survival: Learning in a Dynamic Signalling Model," *Review of Economic Studies*, 2003, *70* (2), 231–251.

Board, Simon and Moritz Meyer-ter-Vehn, "Reputation for Quality," *Econometrica*, 2013, *81* (6), 2381–2462.

Chen, Chia-Hui and Junichiro Ishida, "Hierarchical Experimentation," *Journal of Economic Theory*, 2018, *177*, 365–404.

Crapis, Davide, Bar Ifrach, Costis Maglaras, and Marco Scarsini, "Monopoly Pricing in the Presence of Social Learning," *Management Science*, 2017, *63* (11), 3586–3608.

Cripps, Martin W., George J. Mailath, and Larry Samuelson, "Imperfect Monitoring and Imperfect Reputations," *Econometrica*, 2004, *72* (2), 407–432.

Daley, Brendan and Brett Green, "Waiting for News in the Market for Lemmons," *Econometrica*, 2012, *80* (4), 1433–1504.

Einav, Liran, Chiara Farronato, Jonathan Levin, and Neel Sundaresan, "Auctions versus Posted Prices in Online Markets," *Journal of Political Economy*, 2018, *126* (1), 178–215.

Ekmekci, Mehmet, "Sustainable Reputations with Rating Systems," *Journal of Economic Theory*, 2011, *146* (2), 479–503.

_ , Leandro Gorno, Lucas Maestri, Jian Sun, and Dong Wei, "Learning from Manipulable Signals," *American Economic Review*, 2022, (forthcoming).

Faingold, Eduardo and Yuliy Sannikov, "Reputation in Continuous-Time Games," *Econometrica*, 2011, *79* (3), 773–876.

Fan, Ying, Jiandong Ju, and Mo Xiao, "Reputation Premium and Reputation Management: Evidence from the Largest e-Commerce Platform in China," *International Journal of Industrial Organization*, 2016, *46*, 63–76.

Fradkin, Andrey, Elena Grewal, and David Holtz, "Reciprocity and Unveiling in Two-Sided Reputation Systems: Evidence from an Experiment on Airbnb," *Marketing Science*, 2021, *40* (6), 1013–1029.

Fudenberg, Drew and David K. Levine, "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica*, 1989, *57* (4), 759–778.

_ **and** _ , "Maintaining a Reputation when Strategies are Imperfectly Observed," *Review of Economic Studies*, 1992, *59* (3), 561–579.

_ **and Jean Tirole**, "A "Signal-Jamming" Theory of Predation," *RAND Journal of Economics*, 1986, *17* (3), 366–376.

Gul, Faruk and Wolfgang Pesendorfer, "The War of Information," *Review of Economic Studies*, 2012, *79* (2), 707–734.

Holmstrom, Bengt, "Managerial Incentive Problems: A Dynamic Perspective," *Review of Economic Studies*, 1999, *66* (1), 169–182.

Ifrach, Bar, Costis Maglaras, Marco Scarsini, and Anna Zseleva, "Bayesian Social Learning from Consumer Reviews," *Operations Research*, 2019, *67* (5), 1209–1221.

Kaya, Ayca and Kyungmin Kim, "Trading Dynamics with Private Buyer Signals in the Market for Lemons," *Review of Economic Studies*, 2018, *85* (4), 2138–2352.

Kreps, David M. and Robert Wilson, "Reputation and Imperfect Information," *Journal of Economic Theory*, 1982, *27* (2), 253–279.

Liu, Qingmin, "Information Acquisition and Reputation Dynamics," *Review of Economic Studies*, 2011, *78* (4), 1400–1425.

_ **and Andrzej Skrzypacz**, "Limited Records and Reputation Bubbles," *Journal of Economic Theory*, 2014, *151*, 2–29.

Milgrom, Paul and John Roberts, "Predation, Reputation, and Entry Deterrence," *Journal of Economic Theory*, 1982, *27* (2), 280–312.

Nosko, Chris and Steven Tadelis, "The Limits of Reputation in Platform Markets: An Empirical Analysis and Field Experiment," 2015. NBER Working Paper No. 20830.

Pei, Harry, "Reputation Effects under Interdependent Values," *Econometrica*, 2020, *88* (5), 2175–2202.

—, "Reputation Building under Observational Learning," *Review of Economic Studies, forthcoming*, 2022.

Royden, Halsey and Patrick Fitzpatrick, *Real Analysis*, Pearson Education Asia Limited, 2010.

Tadelis, Steven, "Reputation and Feedback System in Online Platform Markets," *Annual Review of Economics*, 2016, *8* (2), 321–340.

Zervas, Georgios, Davide Proserpio, and John W. Byers, "A First Look at Online Reputation on Airbnb, Where Every Stay is Above Average," *Marketing Letters*, 2021, *32* (1), 1–16.

Appendix A: Proofs

Proof of Proposition 1: In the main text, we prove Proposition 1 based on Lemmas 1 and 2. Here, we provide the proof of each statement.

Proof of Lemma 1. Consider two hypothetical situations: in situation 1, the thresholds are given by $(\underline{q}', \overline{q}')$ with initial prior $\overline{q}'$; in situation 2, they are given by $(\underline{q}'', \overline{q}'')$ with initial prior $\overline{q}''$. Assume $\frac{(1-\underline{q}')\overline{q}'}{\underline{q}'(1-\overline{q}')} = \frac{(1-\underline{q}'')\overline{q}''}{\underline{q}''(1-\overline{q}'')}$, so that $T(\underline{q}', \overline{q}') = T(\underline{q}'', \overline{q}'') \equiv T$. Since we here consider different lower thresholds, with abuse of notation, we write $V^i(\cdot; \underline{q}, \overline{q})$ to explicitly denote its depends on the lower threshold as well.

We first show $V^i(\overline{q}'; \underline{q}', \overline{q}') < V^i(\overline{q}''; \underline{q}'', \overline{q}'')$. To this end, note that in the absence of any good rating, the seller exits the market at time T in both scenarios by assumption. Moreover, since both types of seller adopt the low-price policy when the belief is below the upper threshold, the probability of receiving a good rating, and hence entering the high-price region, is the same between the two situations. Since $\overline{q}'' > \overline{q}'$, the seller can charge a higher price in situation 2 than in situation 1. Also, in the high-price region, a single good rating allows the seller to earn a continuation payoff of $\frac{1}{r}$. Therefore, $V^i(\overline{q}''; \underline{q}'', \overline{q}'') > V^i(\overline{q}'; \underline{q}', \overline{q}')$. Given this, we lower $\underline{q}''$ to $\underline{q}'$ so as to compare the two situations with the same lower threshold. As this change only makes the seller's exit less likely, we in general have $V^i(\overline{q}''; \underline{q}', \overline{q}'') > V^i(\overline{q}''; \underline{q}'', \overline{q}'')$, which in turn implies $V^i(\overline{q}''; \underline{q}', \overline{q}'') > V^i(\overline{q}'; \underline{q}', \overline{q}')$. As this holds for any $\overline{q}'' > \overline{q}'$ and $\underline{q}' = \underline{q}'' = \underline{q}$, $V^i(q; \underline{q}, q)$ is increasing in q. $\square$

Proof of Lemma 2. We first prove $V^H(q) > V^L(q)$ for all $q \in Q$. For any belief higher than $\underline{q}$, the high-quality seller always has a higher probability of receiving a good rating under the pooling strategy considered here. This implies that, starting with any q_t and at any time $t' > t$, the belief distribution of the high-quality seller first-order stochastically dominates that of the low-quality seller. Moreover, since the instantaneous payoff weakly increases with q_t (and strictly increases for $q_t \in Q^h$), the expected payoff of the high-quality seller is always higher than that of the low-quality seller at any time $t' > t$, which proves the claim.

We next show that $V^i(\cdot)$ is increasing. For this purpose, consider two initial beliefs $q_0 = q'$ and $q_0 = q''$, with $q'' > q'$, and compare $V^i(q')$ and $V^i(q'')$. Let $\{\tilde{q}_\tau(q)\}_{\tau=0}^{\infty}$ denote the belief path that is realized when it starts from q and the seller receives no good rating afterwards. First consider the case $q'' > q' > \overline{q}$, so that the seller charges the high price. With no good rating, both $\tilde{q}_t(q')$ and $\tilde{q}_t(q'')$ gradually decrease. However, as long as $\tilde{q}_t(q'') > \tilde{q}_t(q') \geq \overline{q}$, the seller charges a positive price under both cases, and receives a good rating with equal probability. Moreover, the belief will jump to 1 with a single good rating, in which case the seller receives a continuation payoff of $\frac{1}{r}$. This means that before time $T(q', \overline{q})$, the expected payoff under p'' is higher than that under q'. If the seller has received no good rating, we have $\tilde{q}_t(q'') > \overline{q} > \tilde{q}_t(q')$ for $t > T(q', \overline{q})$. The comparison in this case is harder because although the seller can charge a higher price under q'', the probability of receiving a good rating is lower. To show that the expected payoff under q'' is higher in this contingency, we need to show that $V^i(q'') > V^i(q')$ when $q'' > \overline{p} > q'$, i.e., $V^i(q'') > V^i(q')$ when $q'' > q' > \overline{p}$ if $V^i(q'') > V^i(q')$ when $q'' > \overline{p} > q'$. We will deal with this case in the final step of the proof.

Next consider the case where $\overline{q} > q'' > q'$. The argument in this case is similar to the previous one. Since the seller charges the low price under both cases, the probability of receiving a good rating is the same. However, given that $q'' > q'$, it requires fewer good ratings to jump back to the high-price region under q'' than under q'. Moreover, the probability of hitting the lower threshold is higher under q' than under q''. Again, this implies that $V^i(q'') > V^i(q')$ when $\overline{p} > q'' > q'$ if $V^i(q'') > V^i(q')$ when $q'' > \overline{p} > q'$.

The argument above implies that the proof boils down to establishing $V^i(q'') > V^i(q')$ when $q'' > \overline{p} > q'$. Note that when the seller receives a good rating before time $\min\{T(q'', \overline{q}), T(q', \underline{q})\}$, the belief under q'' reaches 1 while the belief under q' jumps to a value strictly smaller than 1. Therefore, the expected payoff under q'' is higher than that under q'. If the seller has received no rating until $T(q'', \overline{q})$, the seller charges a higher price under q'' but has a lower probability of receiving a good rating. Also, since $\tilde{q}_t(q'') > \overline{q} \geq \overline{q}^*$, we know that $\lambda m(\frac{1}{r} - V^L(\overline{q}, \overline{q})) < p(\tilde{q}_t(q''))$. Given that $V^H(\overline{q}; \overline{q}) > V^L(\overline{q}; \overline{q})$, this implies $\lambda m(\frac{1}{r} - V^H(\overline{q}, \overline{q})) < p(\tilde{q}_t(q''))$ as well. Therefore, by (4), the high-price policy yields a higher continuation payoff for both types of seller than the low-price policy before time $T(q'', \overline{q})$. Note that, if the seller were to adopt the low-price policy before reaching $\overline{q}$, her continuation payoff would be exactly the same as the continuation payoff under q'. Since the seller can do even better by adopting the high-price policy, $V^i(q'') > V^i(q')$

when $q'' > \overline{p} > q'$, which completes the proof. □

Proof of Proposition 2. The following proof exploits the value functions derived in Appendix B. It is clear that there cannot be a separating equilibrium in which the low-quality seller offers the high price and the high-quality seller offers the low price, i.e., $\sigma(q_t) = (1, 0)$ for some q_t. Under this profile of strategies, the belief jumps up to 1 after one good rating. The low-quality seller surely finds it profitable to deviate since $\lambda m(\frac{1}{r} - V^L(q_t; \overline{q})) > p(\hat{q}_t^1) = 0$. The remainder of the proof thus focuses on ruling out a separating strategy in which the low-quality seller offers the low price and the high-quality seller offers the high price.

Consider $\sigma(q_t) = (0, 1)$ for some q_t. We then have $\hat{q}_t^1 = 1$ and $q_t^+ < 1$. The proof is still straightforward in this case if $m \geq 1$ because $q_t^+ \leq q_t$ by (2), which implies $\Delta V^i(q_t) < 0$. Since $\lambda m \Delta V^i(q_t) \leq 0 < p(\hat{q}_t) = 1$, a deviation to the high-price policy is profitable for the low-quality seller. In what follows, we thus assume $m < 1$.

We now claim that if $\sigma(q_t) = (0, 1)$ for some q_t, then there must be some interval of q_t such that $\sigma(q_t) = (1, 1)$. Suppose otherwise. Then, it is either $\sigma(q_t) = (0, 1)$ or $\sigma(q_t) = (0, 0)$ for all q_t, in which case the low-quality seller's expected payoff is always 0. She would then deviate to the high-price policy to receive a positive payoff every time the high-quality seller offers the high price. This means that there must be some q_t such that

$$\lambda m \left(\frac{1}{r} - V^L(q_t) \right) \leq p(q_t).$$

Given that both $V^L(\cdot)$ and $p(\cdot)$ are increasing, we can define

$$\overline{z} \equiv \inf \left\{ q_t \mid \lambda m \left(\frac{1}{r} - V^L(q_t) \right) \leq p(q_t) \right\}.$$

Then, for all $q_t \in (\overline{z}, 1)$,

$$\lambda m \Delta V^L(q_t) \leq \lambda m \left(\frac{1}{r} - V^L(q_t) \right) \leq p(q_t) < 1. \tag{8}$$

This implies that, by (4), we cannot have $\sigma(q_t) = (0, 1)$ for any $q_t \in (\overline{z}, 1)$.

For notational simplicity, define

$$\dot{V}_t^i \equiv \frac{\mathrm{d}V^i}{\mathrm{d}q} \dot{\hat{q}}_t,$$

as the time derivative of the value function given that there is no rating. Following (13) and

(14) derived in Appendix B, for q_t such that $\sigma(q_t) = (1,1)$, it must be

$$\dot{V}_t^L = rV^L(q_t) - p(q_t) > -1.$$

Also, for q_t such that $\sigma(q_t) = (0,0)$, it must be that

$$\dot{V}_t^L = rV^L(q_t) - \lambda m \Delta V^L(q_t) > -1$$

by (8). This argument establishes that $\dot{V}_t^L > -1$ for all $q_t \in [\overline{z}, 1)$.

For $q_t < \overline{z}$, since $\lambda m(\frac{1}{r} - V^L(q_t)) > p(q_t)$, we cannot have $\sigma(q_t) = (1,1)$. Define

$$\underline{z} \equiv \sup \{q_t \mid \sigma(q_t) = (0,1)\}.$$

Note that $\underline{z} < \overline{z}$ by their definitions and by (8). Moreover, $\sigma(q_t) = (0,1)$ requires that $\lambda m \Delta V^L(q_t) > 1$. For $q_t < \overline{z}$, only $\sigma(q_t) = (0,0)$ or $\sigma(q_t) = (0,1)$ is possible. On the other hand, for $q_t > \underline{z}$, we cannot have $\sigma(q_t) = (0,1)$ by definition. Therefore, for $q_t \in (\underline{z}, \overline{z})$, it must be that $\sigma(q_t) = (0,0)$. If $\sigma(q_t) = (0,0)$ for all $q_t \leq \underline{z}$, the proof is complete. If not, by the definition of $\underline{z}$, we can find a small number $\varepsilon > 0$ such that $\sigma(q_t) = (0,1)$ and

$$\dot{V}_t^L = rV^L(q_t) - \lambda m \Delta V^L(q_t) > -\frac{\lambda m}{r}, \tag{9}$$

for all $q_t \in (\underline{z} - \varepsilon, \underline{z})$, given that $V^L(q_t^+) < \frac{1}{r}$. The equality in (9) comes from the fact that for $q_t \in (\underline{z} - \varepsilon, \underline{z})$, since $\sigma^L(q_t) = 0$, the continuation payoff can be written as

$$V^L(q_t) = \lambda m V^L(q_t^+)\mathrm{d}t + (1 - r\mathrm{d}t)(1 - \lambda m \mathrm{d}t)V^L(q_t^-),$$

which can be reduced to

$$\lambda m \Delta V^L(q_t) = rV^L(q_t^-) - \dot{V}_t^L.$$

Also, for $q_t \in (\underline{z} - \varepsilon, \underline{z})$, it will take at most $\frac{1}{\lambda(1-m)} \ln \frac{1}{m}$ for the belief to drop from q_t^+ to q_t.[23]

[23] Note that q_t^+ can be either smaller or larger than $\overline{z}$. If it is smaller, then $\sigma(q_t^+) = (0,1)$. If it is larger, then $\sigma(q_t^+) = (0,0)$. The high-quality seller has a higher probability of receiving a good rating in the latter case, since price shoppers can also leave good ratings. This means that when no rating is left after time t, the belief will decrease faster when $q_t^+ > \overline{z}$ than when $q_t^+ < \overline{z}$. In this case, the length of time it takes for the belief to drop back to q_t given by $\tilde{t}$ such that $\frac{\lambda q_t e^{-\lambda \tilde{t}}}{\lambda q_t e^{-\lambda \tilde{t}} + \lambda m(1-q_t)e^{-\lambda m \tilde{t}}} = \tilde{t}$, which yields $\tilde{t} = \frac{1}{\lambda(1-m)} \ln \frac{1}{m}$.

If $\lambda m \leq r$, then $\dot{V}_t^L > -1$ for all $q_t \in (\underline{z} - \varepsilon, \underline{z})$ by (9), and therefore

$$\lambda m \Delta V^L(q_t) = -\lambda m \int_{t-\tilde{t}(q_t)}^{t} \dot{V}_\tau^L \mathrm{d}\tau \leq \lambda \frac{\lambda m}{\lambda(1-m)} \ln \frac{1}{m} = \frac{m}{1-m} \ln \frac{1}{m} < 1,$$

given $m < 1$, which in turn implies that $\lambda m \Delta V^L(q_t) < 1$. Therefore, a deviation to the high-price policy is profitable for the low-quality seller. This implies that for $q_t \in [\underline{z} - \varepsilon, \underline{z}]$, it must be $\sigma(q_t) = (0, 0)$ as well. We can then repeat this process to show that $\sigma(q_t) = (0, 0)$ for all $q_t \leq \underline{z} - \varepsilon$, which completes the proof.

If $\lambda m > r$, we can only ensure $\dot{V}_t^L > -1$ for $q_t > \overline{z}$. In this case, we first show that there exists some $\underline{l} > 0$ such that $\dot{V}_t^L > -1$ for $q_t \in (\overline{z} - \underline{l}, \overline{z})$. To this end, consider $q_t = \overline{z} - l$ for some $l \in (0, \overline{z})$. Then,

$$\begin{aligned}
\lambda m \Delta V^L(q_t) &= -\lambda m \int_{t-T(\overline{z}, \overline{z}-l)}^{t} \dot{V}_\tau^L \mathrm{d}\tau \\
&= -\lambda m \left[\int_{t-T(\overline{z}, \overline{z}-l)}^{t-t_l} \dot{V}_\tau^L d\tau + \int_{t-t_l}^{t} \dot{V}_\tau^L \mathrm{d}\tau \right] \\
&\leq \lambda m \left[\left(\frac{1}{\lambda(1-m)} \ln \frac{1}{m} - t_l \right) + t_l \frac{\lambda m}{r} \right] \\
&= \frac{m}{1-m} \ln \frac{1}{m} + \lambda m t_l \left(\frac{\lambda m}{r} - 1 \right),
\end{aligned}$$

where the inequality comes from the facts that: (i) $\dot{V}_t^L > -1$ for $q_t > \overline{z}$; (ii) $\dot{V}_t^L > -\frac{\lambda m}{r}$ for $q_t \in (\overline{z} - l, \overline{z}]$ by (9); (iii) $T(q_t, \underline{q}) \leq \frac{1}{\lambda(1-m)} \ln \frac{1}{m}$. If

$$T(\overline{z}, \overline{z} - l) < \frac{1 - \frac{m}{1-m} \ln \frac{1}{m}}{\lambda m (\frac{\lambda m}{r} - 1)},$$

then $\lambda m \Delta V^L(q_t) < 1$. Therefore, if $l < \overline{l}(\overline{z})$, where $\overline{l}(q)$ is such that

$$\frac{1}{\lambda} \ln \frac{q[1 - (q - \overline{l}(q))]}{(q - \overline{l}(q))(1 - q)} = \frac{1 - \frac{m}{1-m} \ln \frac{1}{m}}{\lambda m (\frac{\lambda m}{r} - 1)},$$

then $\lambda m \Delta V^L(q_t) < 1$, which implies $\dot{V}_t^L > -1$. Since $\overline{l}(q)$ is continuous and strictly positive on $q \in [\underline{z}, \overline{z}]$, by the extreme value theorem (Royden and Fitzpatrick, 2010, p.26), there exists $\underline{l} \equiv \min_{q \in [\underline{z}, \overline{z}]} \overline{l}(q) > 0$, which proves the claim.

This argument suggests that

$$\dot{V}_t^L = rV^L(q_t) - \lambda m \Delta V^L(q_t) > -1,$$

for $q_t > \overline{z} - \underline{l}$. By applying the same argument, we can next show that given that $\dot{V}_t^L > -1$ for $q_t > \overline{z} - \underline{l}$, there again exists $\underline{l} > 0$ such that for $q_t \in (\overline{z} - 2\underline{l}, \overline{z} - \underline{l}]$, $\lambda m \Delta V^L(q_t) < 1$. By going through this procedure repeatedly, we can have $\lambda m \Delta V^L(q_t) < 1$ for some $q_t \in (\underline{z} - \varepsilon, \underline{z})$. However, this contradicts the previous fact that $\sigma(q_t) = (0, 1)$ for q_t slightly below $\underline{z}$. Therefore, we cannot have $\sigma(q_t) = (0, 1)$ for any $q_t \in Q$ in equilibrium, and the only possibility is $\sigma(q_t) = (0, 0)$ for all $q_t < \underline{z}$. This proves that no separating strategies can constitute an equilibrium, and hence that only pooling equilibrium is possible. □

Proof of Proposition 5. For the following proof, we denote by $V^i(\cdot; \sigma)$ the value function of the type i seller under strategy σ. Let σ^* be the strategy of the seller-optimal equilibrium and $V^{i*}(\cdot) \equiv V^i(\cdot; \sigma^*)$ for brevity.

To obtain the seller-optimal equilibrium, consider a social planner who determines (Q^l, Q^h) to maximize the type i seller's expected payoff among all pure-strategy pooling equilibria. Alternatively, the social planner determines whether to implement $\sigma(q_t) = (0, 0)$ or $\sigma(q_t) = (1, 1)$ at each q_t. Note that the social planner can implement $\sigma(q_t) = (0, 0)$ at any q_t while she can implement $\sigma(q_t) = (1, 1)$ only if q_t is sufficiently large. Specifically, the social planner can implement $\sigma(q_t) = (1, 1)$ only if

$$\lambda m \left(\frac{1}{r} - V^i(q_t; \sigma) \right) \leq p(q_t), \tag{10}$$

for both $i = L, H$. We impose (10) as a constraint because if this fails to hold, there is no way to induce both types to offer the high price in a pooling equilibrium. Note that in any monotone equilibrium with strategy σ, there is a unique $\overline{q}^i(\sigma)$ such that

$$\lambda m \left(\frac{1}{r} - V^i(\overline{q}^i(\sigma); \sigma) \right) = p(\overline{q}^i(\sigma)).$$

Under this formulation, the incentive compatibility constraint for $\sigma(q_t) = (1, 1)$ can be written as $q_t \geq \max\{\overline{q}^L(\sigma), \overline{q}^H(\sigma)\}$.

We first claim that the value function of the seller-optimal equilibrium must satisfy

$$V^{H*}(q_t) = \max_{s \in \{0,1\}} \{p(q_t) s \mathrm{d}t + e^{-r\mathrm{d}t} E[V^{H*}(q_t + \mathrm{d}q_t) \mid s]\}$$

for $q_t \geq \max\{\overline{q}^L(\sigma^*), \overline{q}^H(\sigma^*)\}$,[24] where

$$E[V^{i*}(q_t + \mathrm{d}q_t) \mid s] = \gamma^i(s)\mathrm{d}tV^i(q_t^+(s,s)) + (1 - \gamma^i(s)\mathrm{d}t)V^i(q_t^-(s,s)).$$

Alternatively, for $q_t \geq \max\{\overline{q}^L(\sigma^*), \overline{q}^H(\sigma^*)\}$, the social planner chooses the high-price policy if and only if

$$e^{-r\mathrm{d}t}E[V^{H*}(q_t + \mathrm{d}q_t) \mid s = 0] \leq p(q_t)\mathrm{d}t + e^{-r\mathrm{d}t}E[V^{H*}(q_t + \mathrm{d}q_t) \mid s = 1],$$

which is in the limit reduced to

$$\lambda(1+m)\left(V^{H*}(q_t^+(0,0)) - V^{H*}(q_t)\right) \leq p(q_t) + \lambda\left(\frac{1}{r} - V^{H*}(q_t)\right). \tag{11}$$

Now suppose on the contrary that there is a subset $Q_0 \subset [\max\{\overline{q}^L(\sigma^*), \overline{q}^H(\sigma^*)\}, 1]$ such that for $q_t \in Q_0$, (11) holds strictly but $\sigma^*(q_t) = (0,0)$.[25] We can then pick a $q' \in Q_0$ and construct an alternative strategy σ' such that $\sigma'(q') = (1,1)$ and $\sigma'(q_t) = \sigma^*(q_t)$ for all $q_t \neq q'$. Then, when (11) is satisfied, we have $V^i(q_t; \sigma') > V^*(q_t)$. This means that the incentive compatibility constraint is relaxed in that $\overline{q}^i(\sigma') < \overline{q}^i(\sigma^*)$. This is, however, a contradiction because σ' raises the expected payoff while satisfying the incentive compatibility constraint. This shows that (11) is a necessary condition for the seller optimality.

Given this, we now claim that the seller-optimal equilibrium must be a threshold equilibrium. Suppose on the contrary that there are q and $q' < q$ such that $\sigma^*(q) = (0,0)$ but $\sigma^*(q') = (1,1)$ in the seller-optimal equilibrium. Observe that $\sigma^*(q') = (1,1)$ implies $q' \geq \overline{q}^H$. Observe also that $q_t \geq \max\{\overline{q}^L(\sigma^*), \overline{q}^H(\sigma^*)\}$ implies (11) because $V^{H*}(q_t) \leq \frac{1}{r}$. These two facts in turn imply $\sigma^*(q) = (1,1)$ because $q > q' \geq \overline{q}^H$ and (11) must hold at q, a contradiction. Therefore, the seller-optimal equilibrium must be a threshold equilibrium where $\sigma^*(q_t) = (1,1)$ if and only if $q_t \geq \max\{\overline{q}^L, \overline{q}^H\}$. Then, by statement (i) of Lemma 2 and the definition of $\overline{q}^*$, we have $\overline{q}^L = \overline{q}^* > \overline{q}^H$, suggesting that the threshold equilibrium with $\overline{q} = \overline{q}^*$ achieves the seller optimality.

By the same argument, to maximize the low-quality seller's expected payoff, we must have

$$e^{-r\mathrm{d}t}E[V^{L*}(q_t + \mathrm{d}q_t) \mid s = 0] \leq p(q_t)\mathrm{d}t + e^{-r\mathrm{d}t}E[V^{L*}(q_t + \mathrm{d}q_t) \mid s = 1],$$

[24] Notice a difference from the seller's problem where each type chooses her own policy taking the other type's strategy as given. Here, a unilateral deviation by one type is not considered, and the maximum price the seller can charge depends on the public belief rather than the private belief as a consequence.

[25] When (11) holds with equality, the social planner is indifferent between the low-price and high-price policy. In this case, we assume that the social planner implements the high-price policy.

which becomes

$$\lambda m \left(V^{L*}(q_t^+(0,0)) - V^{L*}(q_t)\right) \leq p(q_t).$$

Again, since $V^{L*}(q_t^+(0,0)) < \frac{1}{r}$, this condition holds if $q_t \geq \overline{q}^L$. Given that the seller-optimal equilibrium must be a threshold equilibrium, this condition holds for all $q_t \geq \overline{q}^L = \overline{q}^*$. This shows that the threshold equilibrium with $\overline{q} = \overline{q}^*$ also maximizes the low-quality seller's expected payoff. □

Appendix B: The value functions

In this appendix, we show how to derive the value function for each type, which we use for the numerical example in Section 5.4 as well as for Proposition 2. Given the equilibrium strategy of our focus, this can be done piecewise for Q^h and each Q_n. In what follows, we denote by $\{\tilde{q}_\tau(q)\}_{\tau=0}^{\infty}$ the realized belief path, starting from q, when the seller receives no good rating as defined in the proof of Lemma 2.

First, if the high-quality seller receives a good rating for any $q_t \in Q^h$, she can reveal her true type and earns a continuation payoff of $\frac{1}{r}$. For an interval $[t, t+dt)$, this event occurs with probability λdt. The value function of the high-quality seller must hence satisfy

$$V^H(q_t) = \left(p(q_t) + \frac{\lambda}{r}\right) \mathrm{d}t + (1 - r\mathrm{d}t)(1 - \lambda \mathrm{d}t) V^H(q_t^-).$$

Taking the limit, we obtain the following first-order ordinary differential equation:

$$p(q_t) + \frac{\lambda}{r} = (r + \lambda) V^H(q_t) - \frac{dV^H}{dq} \dot{\tilde{q}}_t,$$

where $\dot{\tilde{q}}_t = -\lambda q_t (1 - q_t)$ is the rate of change of $\tilde{q}_t$. Solving this differential equation gives

$$e^{-(r+\lambda)T(q,q^*)} V^H(q) = V^H(q^*) - \int_0^{T(q,q^*)} e^{-(r+\lambda)\tau} \left(p(\tilde{q}_\tau(q^*)) + \frac{\lambda}{r}\right) \mathrm{d}\tau, \tag{12}$$

for any $q \in Q^h$, where $q \leq q^* \equiv \frac{\overline{q}(1+m)}{\overline{q}(1+m)+(1-\overline{q})m}$.[26] Similarly, the value function of the low-quality

[26] For the purpose of this proof, it suffice to construct a value function for $q \in [\underline{q}, q^*]$, because q_t never goes above q^* on the equilibrium path (unless the initial prior q_0 is greater than q^*). Obviously, we can derive $V^i(q)$ for $q \in (q^*, 1)$ by following the same procedure.

seller must satisfy

$$V^L(q_t) = p(q_t)dt + (1 - rdt)V^L(q_t^-),$$

which leads to

$$p(q_t) = rV^L(q_t) - \frac{\mathrm{d}V^L}{\mathrm{d}q}\dot{\tilde{q}}_t.$$

Again, solving the differential equation gives

$$e^{-rT(q,q^*)}V^L(q) = V^L(q^*) - \int_0^{T(q,q^*)} e^{-r\tau} p(\tilde{q}_\tau(q^*))d\tau, \tag{13}$$

for any $q \in Q^h$.

We now go down the belief space and consider each Q_n in turn. For $q_t \in Q_n$, the value function of the type i seller must satisfy

$$V^i(q_t) = \gamma^i(0)V^i(q_t^+)\mathrm{d}t + (1 - r\mathrm{d}t)(1 - \gamma^i(0)\mathrm{d}t)V^i(q_t^-),$$

which leads to

$$\gamma^i(0)\Delta V^i(q_t) = rV^i(q_t^-) - \frac{\mathrm{d}V^i}{\mathrm{d}q}\dot{\tilde{q}}_t,$$

where $\dot{\tilde{q}}_t = -\lambda m q_t(1 - q_t)$. Solving this differential equation, we obtain

$$e^{-(r+\gamma^i(0))T(q,q')}V^i(q) = V^i(q') - \int_0^{T(q,q')} e^{-(r+\gamma^i(0))\tau}\gamma^i(0)V^i(\tilde{q}_t(q'))\mathrm{d}\tau, \tag{14}$$

for $q', q \in Q_n$.

Given these, we can obtain the value function for the whole belief space. We first obtain $V^i(q)$ for $q \in Q^h$ by assigning a value $V^i(q^*)$ to (12) and (13). Then, by continuity of the value function, we can further derive $V^i(q)$ for $q \in Q_1$, $q \in Q_2$, and so forth, until we pin down $V^i(\underline{q})$. Note that the value of $V^i(\underline{q})$ continuously increases with the initial value of $V^i(q^*)$. The correct value of $V^i(q^*)$ should be the one that leads to $V^i(\underline{q}) = 0$.