Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Reviews and Self-Selection Bias with Operational Implications

Ningyuan Chen, Anran Li, Kalyan Talluri

To cite this article:
Ningyuan Chen, Anran Li, Kalyan Talluri (2021) Reviews and Self-Selection Bias with Operational Implications. Management Science 67(12):7472-7492. https://doi.org/10.1287/mnsc.2020.3892

http://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Vol. 67, No. 12, December 2021, pp. 7472–7492
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Reviews and Self-Selection Bias with Operational Implications

Ningyuan Chen,[a,b] **Anran Li,**[c] **Kalyan Talluri**[d]

[a] Department of Management, University of Toronto Mississauga, Mississauga, Ontario L5L 1C6, Canada; [b] Rotman School of Management, University of Toronto, Toronto, Ontario M5S 3E6, Canada; [c] Department of Management, The London School of Economics and Political Science, London WC2A 2AE, United Kingdom; [d] Imperial College Business School, Imperial College London, London SW7 2BU, United Kingdom

Contact: ningyuan.chen@utoronto.ca, https://orcid.org/0000-0002-3948-1011 (NC); a.li26@lse.ac.uk, https://orcid.org/0000-0001-7001-2240 (AL); kalyan.talluri@imperial.ac.uk, https://orcid.org/0000-0003-1608-6403 (KT)

Received: June 27, 2019
Revised: October 28, 2019; March 4, 2020
Accepted: April 20, 2020
Published Online in Articles in Advance: May 24, 2021

https://doi.org/10.1287/mnsc.2020.3892

Abstract. Reviews for products and services written by previous consumers have become an influential input to the purchase decision of customers. Many service businesses monitor the reviews closely for feedback as well as detecting service flaws, and they have become part of the performance review for service managers with rewards tied to improvement in the aggregate rating. Many empirical papers have documented a bias in the aggregate ratings, arising because of customers' inherent self-selection in their choices and bounded rationality in evaluating previous reviews. Although there is a vast empirical literature analyzing reviews, theoretical models that try to isolate and explain the bias in ratings are relatively few. Assuming consumers simply substitute the average rating that they see as a proxy for quality, we give a precise characterization of the self-selection bias on ratings of an assortment of products when consumers confound ex ante innate preferences for a product or service with ex post experience and service quality and do not separate the two. We develop a parsimonious choice model for consumer purchase decisions and show that the mechanism leads to an upward bias, which is more pronounced for niche products. Based on our theoretical characterization, we study the effect on pricing and assortment decisions of the firm when potential customers purchase based on the biased ratings. Our results give insights into how quality, prices, and customer feedback are intricately tied together for service firms.

History: Accepted by David Simchi-Levi, operations management.
Funding: The research of N. Chen is partly supported by the Natural Sciences and Engineering Research Council [Discovery Grant RGPIN-2020-04038].

Keywords: social learning • product reviews • consumer choice • assortment optimization and pricing • process variability

1. Introduction

It is widely accepted that reviews of a service or product left by past users are highly influential in determining the purchase decision for many customers. Product reviews and ratings on Amazon.com play a large role in how customers choose among competing products, reviews on Yelp guide dining and service choices, ratings on Tripadvisor for travel and stay choices, and so on. Indeed, as suggested by a recent survey, 93% customers rely on online ratings for purchase decisions (Podium 2017).

Online ratings are generally perceived as reflecting the quality of the product. As a consequence, in many service industries, such as the restaurant and hotel industries, managers closely track and follow reviews, and indeed, managers' performance is also evaluated based on the ratings. However, it is not necessarily the case that ratings reflect true qualities.[1] Consider the following situations that illustrate the biases inherent in reviews:

Example 1. A horror movie on IMDb has a high average rating posted by a few thousand viewers. Is this an accurate estimator for the quality of the movie? Because the viewers and voters of the movie tend to be horror movie fans and by no means a representative sample of the population, the ratings are confounded with tastes and do not reflect quality, say as perceived by a random customer from the general population. The bias may not matter to a fan of horror movies, but an executive considering hiring the director would certainly want to take this into account.

Example 2. Customers who choose to stay at a beachfront hotel may have an inherent preference for the sun, the sand, and the sea-front experience. After staying there, they may not be able to separate the great experience they had on the beach with the service quality of the hotel. Does the rating reflect the quality of the hotel's service? This would be of interest for a hotel chain comparing the performance of managers at its different properties.

This paper concentrates on a particular type of self-selection bias called the *acquisition bias*,[2] which is well documented empirically (see Li and Hitt 2008, Hu et al. 2017, and references therein). Customers who have a favorable predisposition toward the product (ex ante) are more likely to purchase a product and, hence, also to post ratings. The rating, therefore, may be a positively skewed indicator of the true quality. Such bias arises primarily because of consumers' bounded rationality as they do not generally take the trouble of elaborate calculations required to separate their innate tastes and preferences from their experiences or discount such effects in previous reviews. Indeed, when the review evaluation process becomes costly, consumers may simply take the average rating as indicative of the product's quality. Detecting the presence and effects of such bounded rationality is the cornerstone of research in behavioral economics (Kahneman 2003) and the empirical marketing literature (Dellarocas and Wood 2008, Li and Hitt 2008, Hu et al. 2017). The acquisition bias is also related to the phenomenon of the winner's curse (Thaler 1988) or postdecision surprise (Smith and Winkler 2006).

Despite the well-documented phenomenon, consumers seem to rely on these reviews precisely because they are considered objective and reliable. For example, Dellarocas et al. (2007) find that the average numerical rating of online consumer reviews can best predict the future movie revenues compared with any other measure. It is, therefore, important for firms to understand the nature of the bias and its implications for operational decisions.

The first question we attempt to address is how to relate the magnitude of the bias to product attributes. For example, the firm may plan to launch two products: one aimed at a niche market and one aimed at all market segments. An understanding of the connection between product popularity and self-selection biases may provide valuable insights into product-launching strategies.

Next, how is the market share affected by self-selection biases? Consider a retailing platform listing substitutable products. Given that the ratings of all the products are biased to some degree, the choice probability of arriving customers is also distorted by the biased ratings. Which assortment is optimal to offer then in the presence of such bias?

The main theoretical findings and contributions of our research are as follows: First, under a classic choice model, we quantify the acquisition bias and the impact on the rating of an arriving customer. In particular, we show that the expected bias, which is defined to be the difference between the expected rating of the customer and the true quality, is proportional to the logarithm of the reciprocal of the choice probability. An implication of this result is that the ratings of niche products tend to have a higher bias than those of popular products. It may provide theoretical explanations to the empirical observations in Li and Hitt (2008), in which they find that, as a product becomes more popular (i.e., the choice probability increases), the ratings usually exhibit a downward trend. This is consistent with their hypothesis: early adopters of products when they are less mainstream tend to give higher ratings. Taking one step further, our model predicts that acquisition bias serves as the "negative-reinforcing" mechanism; it offsets the intrinsic quality gap between products.

Second, we characterize the asymptotic outcome of social learning. After a large number of customer arrivals, we show that the average ratings and the choice probability (market shares) of all products converge to a limit. The bias stated earlier is systematic and cannot be averaged out. Interestingly, the limiting choice probability resembles the multinomial logit (MNL) model with known qualities except that the standard deviation of the random utilities is almost doubled. Because the random utilities represent the idiosyncratic preferences of customers, our results reveal a surprising and unintended consequence of social learning. Under social learning, the products look less dissimilar quality-wise, and as a result, it has the same effect *as if* the customers are more heterogenous than they actually are. This effect is related to the negative-reinforcing mechanism mentioned and leads to a wider range of choices made by customers. Therefore, the market share of popular products (high quality or low price) is cannibalized by niche products under social learning. The wide range of choices does not necessarily benefit customers as their surplus is lowered by the biased ratings, which induce suboptimal product choices.

Third, as a concrete application, we show how biases and social learning affect the optimal assortment and pricing decisions. In particular, compared with the known quality case, we show that the revenue-ordered assortment may not be optimal, and the optimal assortment cannot be found efficiently. In contrast, the optimal pricing has a structure similar to the case of known qualities: all products have the same markups. We also show that the acquisition bias and social learning always benefit the firm in terms of revenues if the firm optimizes either pricing or assortment.

Our model is applicable to product ratings in online retailing or service ratings, such as in the hospitality or entertainment industry. For both categories, reviews left by past users play a large role in a consumer's purchasing decision. Thus, in what follows, we freely switch between products and services in illustrating our modeling choices.

The relevant literature is reviewed in Section 2. We present the model and quantify the acquisition bias in Section 3. The firm's operational assortment and pricing decisions under the acquisition bias are studied in Section 4. Section 5 extends our model to the underreporting bias. Section 6 is devoted to numerical experiments to strengthen our theories.

2. Literature Review

Social learning theory describes how people learn from each other via observation or communication

(Smallwood and Conlisk 1979). Following the seminal papers of Banerjee (1992) and Bikhchandani et al. (1992), many subsequent papers adopt the following framework: the underlying state of the world is unobserved, and an arriving agent receives a private signal, observes the actions of past agents, updates beliefs in a Bayesian framework, and makes a decision. The main findings of these works are that herd behavior can happen, and the consequent learning failure is a Bayesian equilibrium in a game with incomplete information. However, Smith and Sørensen (2000) show that, if the private signal is unbounded, then learning is achieved asymptotically, that is, the state is eventually inferred correctly.

With the advent of the internet and users posting reviews of products and services, many researchers have studied the issue of bias in reviews and note social influence and the presence of both selection and underreporting bias. Ying et al. (2006), in a two-stage selection and prediction model, find evidence of selection bias in a movie ratings database. Koh et al. (2010) investigate a question very similar to ours: whether ratings reflect true quality (defined in a concrete way in their paper). However, their emphasis and empirical reasoning is based on linking ratings and cultural differences. Engler et al. (2015) argue that customers rate satisfaction rather than quality given the selection bias. Muchnik et al. (2013), using randomized experiments, investigate another source of bias that they call *social bias*, by which previous reviews influence a customer's ratings. In an early well-known paper, Salganik et al. (2006) conduct an experiment on how users rate songs with and without social influence (ratings of previous consumers). In a finding partially echoing some of our results, they find that the quality of niche products is exaggerated in reviews.

Our paper studies learning from reviews and is, thus, also related to prominent empirical studies on online reviews, such as Li and Hitt (2008), Godes and Silva (2012), Lee et al. (2015), and Hu et al. (2017). Li and Hitt (2008) find that the reviews follow a declining trend over time. They provide an explanation for this empirical phenomenon: early consumers have a stronger self-selection bias, and later consumers do not fully correct for the bias when making purchasing decisions. In our paper, the customers do not have nonstationary attributes over time, but the source of the bias is similar. Hu et al. (2017) report the existence of both acquisition and underreporting bias in Amazon.com reviews: consumers with a favorable predisposition acquire a product and, hence, write a product review, and the distribution of online reviews is J-shaped. Li and Hitt (2008) and Hu et al. (2017) also empirically show the bounded rationality of consumers in the sense that they do not fully correct the self-selection biases when interpreting the average rating. In the service sector, Fradkin et al. (2015) use field experiments to show that nonreviewers tend to have worse experiences than reviewers on Airbnb, leading to a positive underreporting bias. All these papers have a strong empirical focus. We, in contrast, develop a model to explain how the selection (and underreporting) bias may arise from consumers interpreting previous ratings at their face value and reporting their total experience and what it means for the firm's pricing and assortment decisions.

Several recent papers in economics have studied learning dynamics in a social network theoretically (Acemoglu et al. 2011, Mossel et al. 2015), and interested readers can find a comprehensive overview in Acemoglu and Ozdaglar (2011). Acemoglu et al. (2018) develop a Bayesian model in which customers observe all past reviews or a summary statistic of a single product. The authors identify conditions for asymptotic learning and characterize the learning speed. As opposed to this stream of literature, the focus of our paper is to quantify the bias and not to achieve asymptotic learning, and our paper assumes that customers use a non-Bayesian learning rule. Non-Bayesian learning occurs when customers have bounded rationality, such as persuasion bias in DeMarzo et al. (2003), imperfect recall in Molavi et al. (2018), and self-selection bias in this paper.

Some recent papers in the operations management literature model the dynamics of learning from reviews and study analytically the learning outcome and the firm's optimal decisions. Maldonado et al. (2018) study rating dynamics when customers choose according to the MNL model similar to ours but with social influence given by current market share (popularity), and Abeliuk et al. (2015) take social influence into consideration to develop policies that maximize profit. Crapis et al. (2016) focus on the pricing policies of a monopoly when consumers learn from reviews, and Papanastasiou and Savva (2016) focus on dynamic pricing of a single product when customers are strategic and learn from previous users in a Bayesian fashion. Besbes and Scarsini (2018) show that, without Bayesian rationality, customers may fail to learn the true quality asymptotically. Vaccari et al. (2018) study the implication of social learning on consumer choices over multiple products. In these two papers, the source of bias (the difference between ratings and the true quality) comes from the uncertain quality experienced by customers; in our model, the bias stems from the idiosyncratic preferences of customers (in addition to service quality shocks). Table 1 summarizes the differences in model setups between this paper and the closest ones in the literature.

The psychological reasons behind acquisition bias have been studied extensively. For example, Juslin et al. (2007) posit that people are naïve with respect to the effects of external sampling biases in information

Table 1. Comparison of Model Setups in Li and Hitt (2008), Hu et al. (2017), Acemoglu et al. (2018), Besbes and Scarsini (2018), Vaccari et al. (2018), and Our Paper

	Li and Hitt (2008)	Hu et al. (2017)	Acemoglu et al. (2018)	Besbes and Scarsini (2018)	Vaccari et al. (2018)	Our paper
Acquisition bias	X	X	X	X		X
Underreporting bias		X				X
Bounded rationality	X	X		X	X	X
Rational Bayesian update			X	X		
Asymptotics			X	X	X	X
Multiple products					X	X
Firm's pricing and assortment decision						X

from the environment. In our case, it is the online ratings that have bias but are treated as unbiased estimators by customers. Recent work by Tong et al. (2018) gives experimental evidence for acquisition bias. The subjects are asked to set prices for a set of houses and choose one that they believe sells for the highest price. It is found that the subjects tend to significantly overestimate the price of the house they choose. At a high level, the acquisition bias is related to the winner's curse (e.g., Thaler 1988) or postdecision surprise (e.g., Smith and Winkler 2006). That is, subjects making choices based on estimators or signals of all alternatives tend to experience disappointment afterward because the signal of the best alternative has selection bias and overestimates the true value.

3. The Model

In this section, we introduce the social learning mechanism and the choice model of customers. The firm is offering $d > 0$ products. The true intrinsic quality of product i, $i \in \{1, \dots, d\}$, is denoted by $q_i \in \mathbb{R}$ and the price by $p_i \in \mathbb{R}_+$, both given exogenously.[3] Customers are indexed by $n \in \{1, 2, \dots\}$ and arrive sequentially. Upon the arrival of customer n, the customer observes the *average rating* of all the products from past customers, denoted by $\hat{q}_i(n)$.

Let the ex ante net utility of owning product i for the consumer, *if* the intrinsic qualities of all products were known, be

$$u_i(n) = q_i - p_i + \varepsilon_i(n), \tag{1}$$

where $q_i - p_i$ is the expected net utility corresponding to the intrinsic quality minus price and the random term $\varepsilon_i(n)$ is the idiosyncratic preference (taste) realized before the purchase, taken to be an independent and identically distributed (i.i.d.) mean-zero Gumbel random variable with cumulative distribution function (CDF)

$$\mathsf{P}(\varepsilon_i(n) \le x) = e^{-e^{-(x-\mu)/\beta}}. \tag{2}$$

Here, μ is the location parameter, and β is the scale parameter. Note that $\mathsf{E}[\varepsilon_i(n)] = \mu + \gamma\beta = 0$, where $\gamma \approx 0.5772$ is the Euler constant.

Although the deterministic part of the utility is identical for all customers, some customers may prefer a particular product i with low intrinsic quality or high price simply because of their tastes, that is, a large ε_i. Because $\mathrm{Var}(\varepsilon_i(n)) = \pi^2\beta^2/6$, the scale parameter β reflects the *heterogeneity* of personal tastes. By convention, we use $i = 0$ to denote the no-purchase option, and thus, p_0 and q_0 are normalized to zero.

According to this random utility model, the choice probability is given by the well-known MNL formula (Ben-Akiva and Lerman 1985)

$$\mathsf{P}\left(u_i(n) \ge \max_{j=0,1,\dots,d} u_j(n)\right) = \frac{\exp\left((q_i - p_i)/\beta\right)}{1 + \sum_{j=1}^{d} \exp\left((q_j - p_j)/\beta\right)}. \tag{3}$$

However, the prices p_i's are observed by the customers, but the intrinsic qualities q_i's are *unknown*.[4] When a customer arrives, we assume the customer uses the observed average rating $\hat{q}_i(n)$ as a proxy for the unknown quality to make choices (note that each customer n may, therefore, have a different quality expectation ex ante $\hat{q}_i(n)$). Combining with the idiosyncratic preference $\varepsilon_i(n)$, the customer forms the *perceived net utility* of product i, $\hat{u}_i(n) = \hat{q}_i(n) - p_i + \varepsilon_i(n)$, and then the customer's choice probability of product i as per our choice model is

$$\mathsf{P}\left(\hat{u}_i(n) \ge \max_{j=0,1,\dots,d} \hat{u}_j(n)\right) = \frac{\exp\left((\hat{q}_i(n) - p_i)/\beta\right)}{1 + \sum_{j=1}^{d} \exp\left(\left(\hat{q}_j(n) - p_j\right)/\beta\right)}. \tag{4}$$

Having defined how customers perceive ratings as expected qualities, it remains to introduce how customers report ratings.

After choosing a product, say i, and using the product, the customer's experienced utility is $r_i(n) = q_i(n) + \varepsilon_i(n)$, where $q_i(n) \triangleq q_i + \xi_i(n)$. The term $q_i(n)$ is consumer n's experienced quality of product i, realized after the consumption. For instance, service deliveries (restaurant, hotel stay, cruise line) have an intrinsic variability, and each customer may vary in the ex post perceptions of the quality (the actual delivered experience). The experience shock $\xi_i(n) = q_i(n) - q_i$ is a mean-zero random variable, independent of everything else, with $\mathrm{Var}(\xi_i(n)) = \sigma_i^2$. To summarize, the idiosyncratic preference ε in the model is generated before the purchase and, if purchased, ξ is realized after the purchase/consumption of the product.

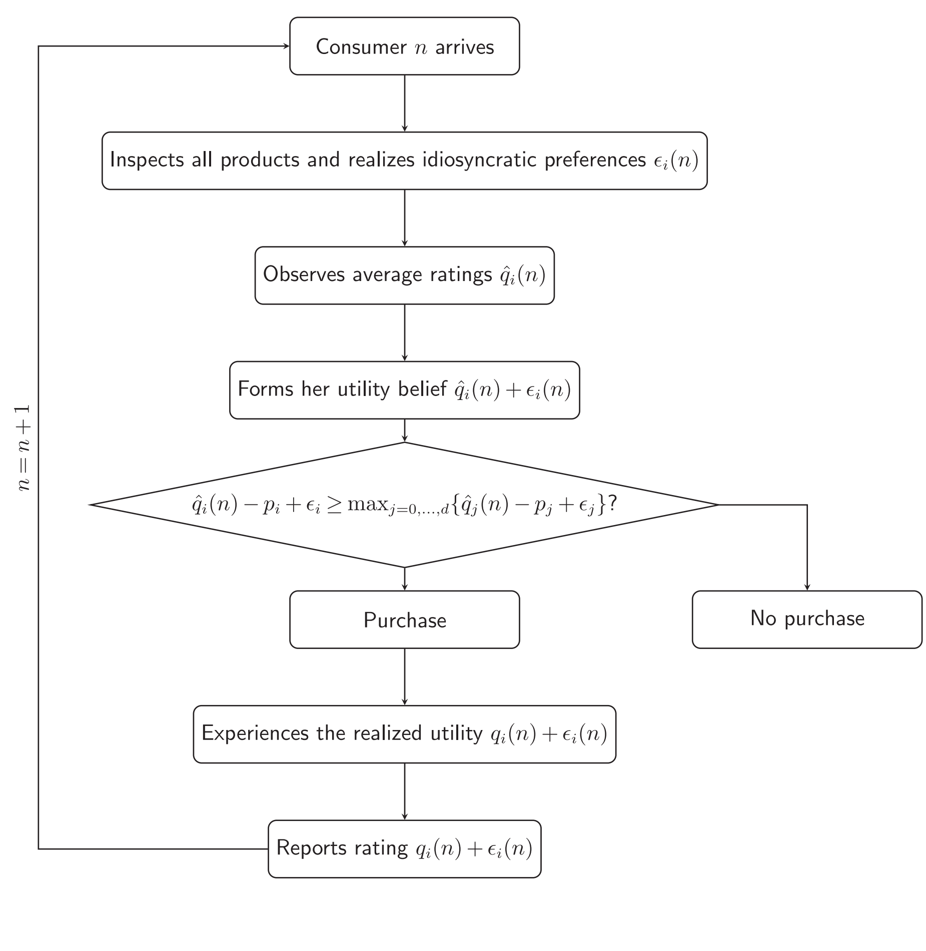
Not being able to distinguish (or bothering to separate) the intrinsic quality q_i from the subjective taste and experience $\varepsilon_i(n) + \xi_i(n)$, the customer reports experienced utility $r_i(n) = q_i + \xi_i(n) + \varepsilon_i(n)$ as the customer's rating of the product. The rating reported by customer n then, in turn, updates the average rating of product i that is observed by the next customer $n+1$, and those of other products are unchanged. The entire process is demonstrated in Figure 1. A list of notations is provided in Appendix A.

In our model, the customers have "bounded rationality" when perceiving ratings before the purchase. Based on the ratings, the customers do not attempt to "debias" and accurately estimate the intrinsic quality using a Bayesian framework. Instead, they simply regard the ratings as proxies for qualities. Bayesian learning is shown to be effective for customers to learn the intrinsic quality (Acemoglu et al. 2018, Besbes and Scarsini 2018). In this paper, we use non-Bayesian learning to capture the source of the acquisition bias (similar to the approach in DeMarzo et al. 2003, Molavi et al. 2018). As we mention in Section 2, the empirical study of Hu et al. (2017) showing that customers cannot infer product quality from biased online reviews supports our assumption. In psychology, such naiveté is well studied; for example, Justin et al. (2007) find that people tend to assume the encountered samples are representative of the population.

Moreover, customers do not adjust for the bias when reporting ratings after the purchase. The customers report $q_i + \xi_i + \varepsilon_i$ instead of $q_i + \xi_i$ as their ratings. Most review platforms do not explicitly instruct customers to report "quality" or "utility," and customers may easily confound objective quality with subjective taste. Empirically, Li and Hitt (2008) find that customers do not fully correct for bias when they report ratings, and Tong et al. (2018) show that individuals tend to overestimate chosen alternatives. Our assumption is similar to the model in Acemoglu et al. (2018).

Although the preceding two points state the "bounded" part of bounded rationality in our model, the ratings are nevertheless updated in a statistically reasonable fashion. Note that $\varepsilon_i(n) + \xi_i(n)$ has mean zero, and

Figure 1. A Consumer's Purchasing and Rating Process



thus, the individual rating $r_i(n)$ is an unbiased signal of the true quality, *unconditionally*. It is reasonable for customers to use the average of the "unbiased" signals to estimate the actual quality. However, the expected value of $\varepsilon_i(n) + \xi_i(n)$ is biased conditional on the event that product i is actually purchased. It is indeed this fact that leads to acquisition bias and renders the average rating a biased signal.

Remark 1. To interpret the different components of the utility, including q, ξ, and ε, it helps to consider the example of Android versus iOS. The quality q measures the perception of the two operating systems of the general population; for example, an average customer may prefer iOS to Android. However, some customers love the customizability of Android, in which their idiosyncratic preference ε, that is, their valuation of Android that deviates from that of the general population, plays a role. To understand ξ, imagine a customer actually bought an iPhone. The phone may happen to have a glitch that drains the battery, which lowers the experience ξ. In other words, if a customer repeatedly purchases the same product with known q, then ε remains the same although ξ is drawn independently every time.

Remark 2. Although most models on online reviews (including ours) assume that customers post ratings based on perceived qualities instead of perceived values (quality minus price), some empirical evidence (Li and Hitt 2010) suggests that customers in many instances do post ratings as perceived values. Most results in this paper still extend to that case as it is just a matter of subtracting the constant price at the rating stage or the evaluation stage. It does not affect customers' perceived qualities from ratings. It remains a future direction to study the outcome when a mixture of customers rate and perceive ratings by both criteria.

3.1. Ratings Evolution Mechanism

With our model for customer choices and ratings in place, we now introduce how the average ratings of the products evolve over time. Let $N_i(n)$ be the number of customers purchasing product i and reporting their ratings prior to customer n. The rating system is fully characterized by

$$\boldsymbol{X}(n) \triangleq (\hat{q}_1(n), \ldots, \hat{q}_d(n), N_1(n), \ldots, N_d(n)), \tag{5}$$

and evolves as a Markovian stochastic process. More precisely, given $\boldsymbol{X}(n)$, customer n chooses a product according to choice probability (4).

Conditional on the event that product i is chosen, the rating $r_i(n) = q_i + \xi_i(n) + \varepsilon_i(n)$ is reported by the customer, and the average rating $q_i(n)$ is updated by the firm as follows:

$$\hat{q}_i(n+1) = \frac{N_i(n)}{N_i(n)+1}\hat{q}_i(n) + \frac{1}{N_i(n)+1}r_i(n). \tag{6}$$

For all products $j \neq i$, their ratings remain the same: $\hat{q}_j(n+1) = \hat{q}_j(n)$. Similarly, $N_i(n+1) = N_i(n) + 1$ and $N_j(n+1) = N_j(n)$ for $j \neq i$.

We do not specify the initial state of the ratings, that is, $\hat{q}_i(0)$ for $i = 1, \ldots, d$, as by the updating formula, this is quickly overwritten by the customers' given ratings. Therefore, the theoretical properties derived in the paper are independent of the initial state. In the simulation study in Section 6, we start with $\hat{q}_i(0) \equiv 0$.

3.2. Research Questions

We now ask the following research questions:

1. Does bias—defined as the difference between the expected rating and the true quality—exist? If so, how large is it?
2. Do the average ratings $\hat{q}_i(n)$ converge as customers purchase the products and post their ratings? If so, to what do they converge? What are the limit choice probabilities based on these ratings?
3. How do the (potentially biased) limit choice probabilities compare with the probabilities when the true qualities are known to consumers?
4. How would social learning influence the expected consumer surplus? How does the expected consumer surplus compare with the known-quality case?
5. Does the firm have an incentive to reveal intrinsic qualities? In other words, does it make more revenue with the bias in ratings compared with revealing (if a credible way exists to do so) the true qualities?

We address all these questions in turn next.

3.2.1. Do Biases Exist?

We first introduce a formal expression for the bias. Let

$$\begin{aligned} E_i(n) &\triangleq \left\{\hat{u}_i(n) \geq \max_{j=0,1,\ldots,d} \hat{u}j(n)\right\} \\ &= \left\{\hat{q}_i(n) - p_i + \varepsilon_i(n) \geq \max_{j=0,1,\ldots,d} \hat{q}_j(n) - p_j + \varepsilon_j(n)\right\} \end{aligned}$$

denote the event that customer n purchases product i. The expected rating is $\mathsf{E}[r_i(n)|E_i(n)]$, so the expected bias of the customer rating on product i is defined as $\mathsf{E}[r_i(n) - q_i|E_i(n)]$.

Note that, conditional on the purchase of product i ($E_i(n)$), the ex post rating can be decomposed into $q_i + \varepsilon_i(n)$ and $\xi_i(n)$, where the latter is independent of everything else. The conditional distribution of the first component, $q_i + \varepsilon_i(n)$, satisfies

$$\mathsf{P}(q_i + \varepsilon_i(n) \leq x | E_i(n)) = \frac{\mathsf{P}(q_i + \varepsilon_i(n) \leq x, E_i(n))}{\mathsf{P}(E_i(n))}. \tag{7}$$

The probability of $E_i(n)$, which is the choice probability of product i, is denoted as

$$C_i(n) \triangleq \mathsf{P}(E_i(n)) = \frac{\exp((\hat{q}_i(n) - p_i)/\beta)}{1 + \sum_{j=1}^{d} \exp((\hat{q}_j(n) - p_j)/\beta)}.$$

The probability $\mathsf{P}(q_i+\varepsilon_i(n)\le x, E_i(n))$ is equal to

$$\int_{-\infty}^{x-q_i}\underbrace{\frac{1}{\beta}e^{(\mu-y)/\beta}e^{-e^{(\mu-y)/\beta}}}_{(1)}$$
$$\times\prod_{j=0,\ldots,d,j\neq i}\underbrace{\mathsf{P}\Big(\hat{q}_j(n)-p_j+\varepsilon_j(n)\le\hat{q}_i(n)-p_i+y\Big)}_{(2)}dy.$$

Term (1) in the preceding equation is the probability density function (PDF) of $\varepsilon_i(n)$; conditional on $\varepsilon_i(n)=y$, term (2) decomposes the probability of $E_i(n)$ into $\mathsf{P}(\hat{q}_j(n)-p_j+\varepsilon_j(n)\le\hat{q}_i(n)-p_i+y)$ for $j\neq i$ because $\varepsilon_j(n)$'s are independent. After simplifying (7), we have the following.

Proposition 1. *Conditional on the purchase event $E_i(n)$, $q_i+\varepsilon_i(n)$ has a Gumbel distribution with mean $q_i-\beta\log(C_i(n))$ and variance $\pi^2\beta^2/6$:*

$$\mathsf{P}(q_i+\varepsilon_i(n)\le x|E_i(n))=\exp\Big(-e^{(\mu-(x-q_i+\beta\log(C_i(n))))/\beta}\Big).\tag{8}$$

Therefore, conditional on a purchase of product i, the rating $r_i(n)$ of customer n, which is the sum of the independent random variables $q_i+\varepsilon_i(n)$ and the service shock $\xi_i(n)$, satisfies

$$\mathsf{E}\big[r_i(n)|E_i(n)\big]=q_i-\beta\log(C_i(n)),$$
$$\mathrm{Var}\big(r_i(n)|E_i(n)\big)=\frac{\pi^2\beta^2}{6}+\sigma_i^2.$$

The proofs of Proposition 1 and all subsequent results can be found in Appendix B. Proposition 1 confirms the intuition that the bias $\mathsf{E}[r_i(n)-q_i|E_i(n)]$ always exists. More precisely, the bias equals $-\beta\log(C_i(n))$. Because the choice probability $C_i(n)\in(0,1)$, the bias is always positive; that is, the customer reports a rating that is higher than the true quality in expectation.

Moreover, for the same taste heterogeneity measure β, the acquisition bias is proportional to the logarithm of the reciprocal of the choice probability. In other words, a "niche" product, which is associated with smaller choice probability, tends to have a larger bias—a phenomenon also observed empirically in Li and Hitt (2008). On the other hand, the variance of the biased rating is simply the unconditional variance of $\varepsilon_i(n)+\xi_i(n)$.

It is interesting to note that the acquisition bias displays a negative-reinforcing phenomenon. Under the negative-reinforcing mechanism, unpopular products have ratings with a larger bias and, thus, appear more "attractive" relatively. As a result, the difference in the intrinsic attractiveness of the products is diminished when viewed through their ratings.

3.2.2. Do the Average Ratings $\hat{q}_i(n)$ Converge? If so, to What?

We now proceed to the second question. Notice that, unconditionally, the ratings are unbiased signals of quality, so a reasonable conjecture could be that they converge to the true qualities, but this turns out not to be the case. The key is to note that the acquisition bias is systematic, not idiosyncratic like the randomness in experience utilities and, thus, cannot be averaged out. Whenever a customer chooses a product, it always implies that the customer has a personal preference for the product over the others. As more and more customers report their ratings, the bias accumulates over time and does not vanish.

We first show that the ratings indeed converge, which is consistent with practical observations. For example, the ratings for movies on IMDb start to stabilize once the number of users reporting their ratings exceeds a few thousand. The intuition is essentially the law of large numbers: although heterogeneous tastes drive ratings in a stochastic way, in the long run, such randomness is averaged out (but still biased). The technical proof, which is included in the appendix, is more involved as the ratings of customers are not independent, but form a complicated stochastic process.

To gain intuition on to where the ratings converge, consider the following fixed-point heuristic. Suppose the limiting average ratings of the products are $\hat{q}_i^\infty$. Then the limiting choice probability is given by

$$C_i^\infty=\frac{\exp((\hat{q}_i^\infty-p_i)/\beta)}{1+\sum_{j=1}^d\exp\Big(\big(\hat{q}_j^\infty-p_j\big)/\beta\Big)}.$$

The limiting bias, characterized in Proposition 1 based on the limiting choice probability, must, in turn, be consistent with $\hat{q}_i^\infty$. Therefore, the following fixed point seems plausible

$$q_i-\beta\log(C_i^\infty)=\hat{q}_i^\infty.$$

The formal result making this intuition concrete is the following theorem.

Theorem 1. *As $n\to\infty$, $\hat{q}_i(n)\to\hat{q}_i^\infty$ and $C_i(n)\to C_i^\infty$ for all i almost surely, where*

$$C_i^\infty=\frac{2e^{(q_i-p_i)/(2\beta)}}{\sum_{j=1}^d e^{(q_j-p_j)/(2\beta)}+\sqrt{\Big(\sum_{j=1}^d e^{(q_j-p_j)/(2\beta)}\Big)^2+4}},$$
$$\hat{q}_i^\infty=q_i-\beta\log(C_i^\infty).$$

From Theorem 1, it is clear that $\hat{q}_i^\infty>q_i$. Moreover, the limiting choice probability of product i exhibits the independence of irrelevant alternatives property, similar to the MNL model with known quality q_i and price p_i, and the scale parameter of the customers' heterogeneous taste is 2β instead of β.

Theorem 1 also reveals an interesting consequence of social learning: the rating system offsets the intrinsic attractiveness gaps between products by making customers *appear* more heterogeneous than they actually are. To understand it, imagine when the scale parameter is infinite. In this case, the qualities and prices of the products no longer matter as all customers simply follow their own idiosyncratic preferences. Therefore, under social learning, niche products attract more customers than they would have otherwise, and high-quality products attract relatively less. As a result, choice probabilities distribute more uniformly among products. This is a consequence of the negative-reinforcing mechanism we mentioned previously: the products look less dissimilar quality-wise, and as a result, it is equivalent to the effect of larger heterogeneity in customers' tastes.

The convergence of online ratings to a limit is a central topic in the literature on social learning. For example, theorem 2 in Besbes and Scarsini (2018) and proposition 1 in Vaccari et al. (2018) both document the failure of social learning for naïve or non-Bayesian customers. In Theorem 1, we not only demonstrate the bias qualitatively, but also quantify the magnitude of bias and the implication for market share. Interestingly, the setting of multiple products under the MNL framework actually makes the model more tractable.

3.2.3. Comparison of Choice Probabilities if True Qualities Were Known to Consumers. How do the limit choice probabilities (under social learning) compare with those with the true qualities if they were known to all? As shown in Section 3.2.2, our intuition is that social learning makes the quality gap between products less pronounced. We formalize this as follows.

Proposition 2. *Compared with the choice probabilities* (3) *with known qualities, we have that*

1. *The no-purchase probability decreases.*
2. *The ranking of the market shares of the products is preserved (except for the no-purchase option).*
3. *There exists a cutoff $\tau > 0$ (dependent on all p_i's and q_i's) such that, for products with $q_i - p_i \le \tau$, the choice probability increases, and for products with $q_i - p_i \ge \tau$, the choice probability decreases.*

Under social learning, the ratings of all products are positively biased in the long term. Proposition 2, part one says the no-purchase option becomes relatively unattractive and has a lower choice probability. The second part states that, although the ratings are distorted, they do not significantly disrupt the order as the ranking of the market shares is preserved. However, the ratings may not have the same ranking as the quality. Because the prices affect the choice probabilities, products of similar qualities may have very different self-selection bias and, thus, ratings in the limit. This is in contrast with the model of Vaccari et al. (2018), in which the rankings of the qualities are preserved as they don't model the self-selection bias explicitly. The third part of the proposition states that less attractive products (low quality or high price) benefit from social learning, and attractive products hurt. Proposition 2 extends our previous observation on the negative reinforcement effect of social learning, which makes low-quality products more viable.

3.2.4. Comparison of Consumer Surplus. Social learning leads to positive bias in the average ratings of all products. However, it is not clear how it would affect consumer surplus. We first recall the expected consumer surplus under the MNL model when qualities are known (Anderson and de Palma 1992):

$$\mathsf{E}\Big[\max_{j=0,1,\ldots,d} u_i\Big] = \mathsf{E}\Big[\max_{j=0,1,\ldots,d} q_i - p_i + \varepsilon_i\Big] = \beta \log\left(1 + \sum_{i=1}^{d} \exp\Big((q_i - p_i)/\beta\Big)\right). \tag{9}$$

Under social learning, a consumer's rating is closely related to the consumer's surplus, that is, net utility. More specifically, consumer n's surplus of product i should be the consumer's rating of i, $r_i(n)$ net of the price p_i. Consequently, from Proposition 1, the expected surplus of consumer n, conditional on the consumer's purchase of product i, is $\mathsf{E}[r_i(n) - p_i | E_i(n)] = q_i - p_i - \beta\log(C_i(n))$. As a result, the (unconditional) expected consumer surplus is just

$$\sum_{i=0}^{d} \mathsf{E}\big[r_i(n) - p_i | E_i(n)\big]\mathsf{P}(E_i(n)) = \sum_{i=0}^{d} (q_i - p_i - \beta\log(C_i(n)))C_i(n). \tag{10}$$

It is intuitive that (9) $\ge$ (10) as customers are making optimal decisions based on a misspecified environment under social learning (inflated ratings). The bias induces suboptimal choices, which can only lead to lower surplus compared with making decisions under full information. It is unclear how to attribute surplus loss. Is it caused by a good product seeming less good relatively or a product with low quality getting a superficially high rating? Answering these precisely seems very difficult, but we gain some insight when we try to quantify these effects.

We use C_i^* to denote the choice probabilities under known qualities. Clearly plugging in $C_i(n) = C_i^*$ in (10) gives (9). We would study the loss relative to (9) when $C_i(n)$ slightly deviates from C_i^*, say $C_i(n) - C_i^* = \delta_i$. The results are summarized in Proposition 3.

Proposition 3. *Comparing the expected consumer surplus with/without social learning, we have*

- *The expected consumer surplus under social learning is always lower,* (9) $\ge$ (10) *unless* $C_i(n) = C_i^*$, $i = 1, \ldots, n$.

- *The surplus loss is of the order*

$$\sum_{i=0}^{d} \frac{\beta\delta_i^2}{C_i^*} + O\left(\delta_i^3\right). \tag{11}$$

The proposition implies that the misinformation or bias carried by the average rating does incur a cost to the consumers. Moreover, the bias for products with low C_i^*, that is, the niche products that few customers would have chosen under known qualities, is most responsible for the surplus loss. This can be seen from the coefficient in (11) as a slight inflation in the choice probability δ_i is magnified by the coefficient β/C_i^* when C_i^* is small.

3.2.5. Comparison of Revenues. If the firm can truthfully commit and has the credibility to reveal the product qualities up front, then is it worthwhile for it to do so? Given a fixed set of prices and an assortment of products, the revenues can go either way: there are instances in which the firm makes more revenue under social learning and other instances in which it is better off revealing true qualities as we show in the two following examples ($\beta = 1$ for both of them).

Example 3. In this example, for a fixed assortment and prices, revealing true qualities gives higher revenue for the firm. Consider two products with true qualities $q_1 = 100, q_2 = 2$ and prices $p_1 = 90, p_2 = 1$, leading to choice probabilities $C_1 = 0.9998, C_2 = 0.0001$ and a revenue of 90.0 under known qualities. Under social learning, the limit ratings lead to $C_1^\infty = 0.989, C_2^\infty = 0.011$ and a revenue of 89.0.

Example 4. In this example, social learning gives higher revenue for the firm. Consider two products with true qualities $q_1 = 100, q_2 = 105$ and prices $p_1 = 93, p_2 = 100$, leading to choice probabilities $C_1 = 0.880, C_2 = 0.119$ and a revenue of 93.8 under known qualities. Under social learning, the limit choice probabilities are $C_1^\infty = 0.731, C_2^\infty = 0.269$ and a revenue of 94.8.

On the other hand, we show the following interesting fact in Section 4.3: if the firm either (a) optimizes prices for a fixed assortment or (b) optimizes the assortment given prices, both with the seller aware of the true qualities, then it always makes more revenue with the social learning mechanism as compared with revealing the true qualities of the products to the consumers. So the firm may have no incentives to reveal true qualities, but instead let the biases remain in the ratings.

4. Pricing and Assortment Optimization

In this section, we turn to the practical use of our findings. We consider an online retailer selling products with a rating system and explore how the social learning mechanism affects the firm's optimal operational decisions in assortment planning and pricing. We assume that the seller is aware of the true qualities and the social learning mechanism, and customers behave according to the model of Section 3, that is, base their purchase decision on the average rating.

For optimal pricing and assortment, we focus on the limiting choice probability. The optimal assortment problem is given by $\max_{S\subseteq\{1,\ldots,d\}} \Sigma_{i\in S}\, p_i C_i^\infty(S)$, where $C_i^\infty(S)$ is the limiting choice probability presented in Theorem 1 when an assortment of products S is offered instead of $\{1,\ldots,d\}$. The optimal pricing problem is given by $\max_{p_1,\ldots,p_d} \Sigma_{i=1}^d p_i C_i^\infty$. As we see in Section 6, the convergence of the average ratings to the limit occurs after a few hundred customer arrivals for less than 10 products. Therefore, it is reasonable to consider the limiting market shares instead of the transient ones.

4.1. Assortment Optimization

Without loss of generality, let us label products in decreasing order of their prices p_i's, that is, $p_1 \geq p_2 \geq \ldots \geq p_d$. When the qualities are known, the firm tries to find an assortment of product $S\subseteq\{1,2,\ldots,d\}$ to offer to maximize the revenue:

$$\max_{S\subseteq\{1,2,\ldots,d\}} \frac{\sum_{i\in S} p_i \exp\left((q_i - p_i)/\beta\right)}{1 + \sum_{i\in S} \exp\left((q_i - p_i)/\beta\right)}.$$

The assortment optimization problem with known qualities is well studied in the literature. For example, Talluri and Van Ryzin (2004) find that the optimal assortment S^* under the MNL model is a revenue-ordered set, that is, $S^* = \{1,2,\ldots,k\}$ for some k.

Under social learning, the firm tries to maximize

$$\begin{aligned}&\max_{S\subseteq\{1,2,\ldots,d\}} \sum_{i\in S} p_i C_i^\infty(S)\\ &= \max_{S\subseteq\{1,2,\ldots,d\}} \frac{\sum_{i\in S} 2p_i \exp\left((q_i - p_i)/(2\beta)\right)}{\sum_{i\in S} e^{(q_i-p_i)/(2\beta)} + \sqrt{\left(\sum_{i\in S} e^{(q_i-p_i)/(2\beta)}\right)^2 + 4}}.\end{aligned} \tag{12}$$

In this case, however, the optimal assortment might be a set that is not ordered by the revenue as we show in the following example.

Example 5. In this example, $d = 3$, $\beta = 1$, $p_1 = 8.1$, $p_2 = 4.5$, $p_3 = 4$, $\exp((q_1 - p_1)/(2\beta)) = 1$, $\exp((q_2 - p_2)/(2\beta)) = 10$, and $\exp((q_3 - p_3)/(2\beta)) = 1$. By enumerating, the optimal assortment under social learning is $S^* = \{1,3\}$. In contrast, when qualities are known, the optimal assortment is $S^* = \{1\}$.

To understand why the optimal assortment is $\{1,3\}$ under social learning, note that the product with the highest revenue, product 1, is always included by Proposition 4. It is not optimal to include product 2 because its revenue is much lower although it is substantially more attractive than product 1; thus, the market share (choice probability) of product 1 would be cannibalized if product 2 were included. But, then, why is it optimal to include product 3, which has even lower revenue? This is because, although product 3 cannibalizes some market share from product 1, it attracts even more customers who would have not purchased anything. The new source of revenue offsets the cannibalization effect. More precisely, although the market share of product 1 is reduced slightly after including product 3, the bias of product 1 becomes larger, following Proposition 1, and it ends up with a higher rating. It offsets the decrease in market share, and thus, the cannibalization effect is not significant. On the other hand, the no-purchase probability greatly decreases. This explains why the revenue is increased after including product 3.

Next, we show some structural results of the optimal assortment. Denote S^* the optimal assortment under social learning.

Proposition 4. *Suppose a product $i \in S^*$.*

1. *If $p_j e^{(q_j-p_j)/(2\beta)} > p_i e^{(q_i-p_i)/(2\beta)}$ and $e^{(q_j-p_j)/(2\beta)} > e^{(q_i-p_i)/(2\beta)}$, then $j \in S^*$.*
2. *If $p_j > 2p_i$, then $j \in S^*$.*
3. *The product with the highest revenue is always in the optimal assortment.*

The first claim implies a partial order of products: if a product has high revenue but low quality, it is, in general, a good idea to include it in the assortment. The second and third claims show that, although the optimal assortment is not strictly revenue ordered, it is so to a large extent. In particular, the second claim implies that there exists a cutoff p^* such that the optimal assortment includes all products with revenues higher than $2p^*$ and no products with revenues lower than p^*; however, it has to be computed case by case for products with revenues in $[p^*, 2p^*]$.

Based on Proposition 4, one would expect that the computation of the optimal assortment is tractable. Unfortunately, this is not the case, and we show this in Proposition 5.

Proposition 5. The assortment optimization problem (12) is NP-hard.

Although the optimal assortment is not revenue ordered and the assortment optimization problem is NP-hard, Proposition 4 seems to imply that the optimal assortment may be "almost" revenue ordered. Note that the complexity of finding the best revenue-ordered assortment is $O(d)$, so if a revenue-ordered assortment can generate revenue close to that of the optimal assortment, then it may be worthwhile to sacrifice a small amount of revenue for computational efficiency. Our next result gives a bound on the difference.

Proposition 6. *The best revenue-ordered assortment can generate at least $1/2$ of the optimal revenue.*

Proposition 6 implies that the procedure to find the optimal assortment in the classic MNL model still performs well. In particular, the firm can compute the expected revenues of d revenue-ordered assortments and choose the best among them. It is, of course, significantly cheaper computationally than to examine all $2^d - 1$ assortments, and the expected revenue is still guaranteed to be within a factor of $1/2$.

4.2. Optimal Pricing

Next, we consider the pricing problem of the firm. That is, given the assortment of the products $i \in \{1, \dots, d\}$, the objective of the firm is to determine the prices p_i for all i so that the revenue $\sum_{i=1}^{d} p_i C_i^{\infty}$ is maximized:

$$\max_{p_1,\dots,p_d} \frac{\sum_{i=1}^{d} 2p_i \exp\left((q_i - p_i)/(2\beta)\right)}{\sum_{i=1}^{d} e^{(q_i-p_i)/(2\beta)} + \sqrt{\left(\sum_{i=1}^{d} e^{(q_i-p_i)/(2\beta)}\right)^2 + 4}}.$$

Note that setting a price $p_i = \infty$ is equivalent to excluding product i from offer. So the price optimization also solves the joint pricing and assortment optimization problem. It is well known that, under the classic MNL model, the markup (defined as price minus cost) is constant (see Anderson and de Palma 1992). That is, when the qualities are known, the optimal prices satisfy $p_1 = \dots = p_d$ when the costs are normalized to zero. Surprisingly, the same property holds in the presence of social learning.

Proposition 7. *The optimal pricing under social learning satisfies $p_1 = \dots = p_d < \infty$ (costs normalized to zero). Moreover, the optimal (common) price under social learning is higher than that with known qualities.*

Proposition 7 provides valuable insights for firms transitioning from brick and mortar to online services and starting to implement a rating system. It implies that prices should increase after the transition to maximize revenue.

4.3. Revenue and Surplus When the Firm Optimizes Under Social Learning

In Section 3.2.5, we gave examples showing that, if the assortments or prices are fixed, the firm's revenue can be higher or lower when consumers know true qualities. In this section, we compare the revenues when the assortment of products or the prices are optimized by the firm.

We prove analytically that revenues under social learning dominate true quality revelation as long as the firm optimizes its prices for a fixed assortment or optimizes its assortment for a fixed set of prices. Therefore, we conclude that the firm essentially has no incentives to reveal true qualities even if it has a credible mechanism to do so.

Proposition 8. *The firm achieves more revenue under social learning (i.e., consumers deciding based on $\hat{q}_i^{\infty}$) as compared with revealing true qualities (q_i) to the consumer under either of the following two conditions (optimizations done with firm knowing true qualities):*

1. *For a fixed assortment, the firm sets optimal prices.*
2. *For a fixed set of prices, the firm optimizes the assortment.*

Condition 1 of Proposition 8 is straightforward, following Proposition 2, which states that the no-purchase probability decreases under social learning, and Proposition 7, which shows that the optimal prices are all the same (with costs normalized to zero). So, fixing this common price, social learning boosts the entire market share, resulting in more revenue. Condition 2 of Proposition 8 seemingly contradicts Example 3, which shows that the expected revenue under social learning can be lower for a given assortment. This is because social learning can cannibalize the demand for more profitable products. However, Proposition 8, part 2 shows that it will not happen if we can properly choose the optimal assortment. We point out that the first condition of Proposition 8 covers the case when assortment and pricing are optimized simultaneously as we allow for $p_i = \infty$.

A relevant question we ask is how the expected consumer surplus compares with the known-quality case when the firm can optimize its assortment or the prices. Proposition 3 shows that, when the assortment and the prices are fixed, the expected consumer surplus is lower under social learning than under known qualities. However, we need to give answers independently for the optimized assortment or prices.

When the prices are fixed and the firm optimizes the assortment, depending on the parameters, social learning can either benefit or hurt consumers. For example, if the firm has only one product, the optimal assortment under both scenarios would always be the same, that is, offering the single product. By Proposition 3, the consumer surplus is lower under social learning. We can also construct examples such that the consumer surplus is higher under social learning. Recall that in Example 5 the optimal assortment under MNL is $\{1\}$, which generates consumer surplus 0.69, and the optimal one under social learning is $\{1,3\}$, which generates higher consumer surplus 1.03. In this example, the firm offers a larger set of products to maximize revenue, which happens to benefit consumers.

When the assortment is fixed and the firm optimizes prices, consumers are always hurt by social learning as the firm always charges higher prices by Proposition 7.

5. Extension: Underreporting Bias

Besides the acquisition bias studied in the model, there is another type of self-selection bias documented in the literature (Hu et al. 2017) referred to as the underreporting bias. That is, consumers with extreme, either positive or negative, ratings are more likely to write reviews than consumers with moderate product ratings. One explanation for the underreporting bias is the social satisfaction from word-of-mouth communication: when the experienced quality is much lower than the prior expectation, the customer is motivated to "moan" and warn others or, in the opposite case, to "brag" and share the surprising good news. In this section, we examine the impact of the underreporting bias and resulting rating evolutions.

To isolate the underreporting bias, we intentionally remove the acquisition bias and the experience shock from the model: given the current ratings $(\hat{q}_1(n), \ldots, \hat{q}_d(n))$, we assume the arriving customer n is intelligent enough to debias the rating so that, if the customer chooses product i, the *potential* rating $r_i(n)$ is a Gumbel random variable with mean q_i and scale β. In contrast to Proposition 1, the acquisition bias $-\beta\log(C_i(n))$ has been removed.

To account for the underreporting bias, suppose the customer only posts the rating if $|r_i(n) - \hat{q}_i(n)| > R$. That is, the impression of the product ($r_i(n)$) is significantly different from what the online rating indicates ($\hat{q}_i(n)$) for some threshold R. More precisely, given that product i is chosen, the evolution of the ratings (6) is replaced by

$$\hat{q}_i(n+1) = \begin{cases} \dfrac{N_i(n)}{N_i(n)+1}\hat{q}_i(n) + \dfrac{1}{N_i(n)+1}r_i(n) & |r_i(n) - \hat{q}_i(n)| > R \\ \hat{q}_i(n) & |r_i(n) - \hat{q}_i(n)| \le R. \end{cases}$$

The ex post expected value of the customer's rating is $\mathsf{E}[r_i(n) \| r_i(n) - \hat{q}_i(n)| < R]$. The underreporting bias can, thus, be quantified as $\mathsf{E}[r_i(n) \| r_i(n) - \hat{q}_i(n)| < R] - \mathsf{E}[r_i(n)] = \mathsf{E}[r_i(n) \| r_i(n) - \hat{q}_i(n)| < R] - q_i$.

To analyze the impact of the under-reporting bias on the rating system, we first introduce the following technical lemma.

Lemma 1. *Suppose X is a Gumbel random variable with mean zero and scale β.*

1. *As $|\mu| \to \infty$, $\mathsf{E}[X \| X - \mu| < R]$ converges to zero.*
2. *There exists a unique $\eta < 0$ so that $\mathsf{E}[X \| X - \eta| < R] = 0$, $\mathsf{E}[X \| X - \mu| < R] < 0$ when $\mu > \eta$ and $\mathsf{E}[X \| X - \eta| > R] < 0$ when $\mu < \eta$.*

To apply Lemma 1, let $\mu = \hat{q}_i(n) - q_i$ and $X = r_i(n) - q_i$. The underreporting bias is equal to $\mathsf{E}[X \| X - \mu| > R]$ after the shift. The first property states that, if the current rating deviates from the true quality by a large margin (large $|\mu|$), then the underreporting bias has little impact because all customers are likely to leave reviews after observing a large discrepancy between the expectation and the experience. The underreporting bias, thus, disappears. The second property dictates the pattern of the underreporting bias: if $(\hat{q}_i(n) - q_i > \eta)$, then the underreporting bias causes the new rating to be positively (negatively) biased.

To understand the long-run effect of the underreporting bias, we study the limit of the rating dynamics as $n \to \infty$, similar to our analysis in Section 3.2.2. Suppose the limiting average ratings of the products are $\hat{q}_i^{\infty}$. Then, the limiting choice probability is given by

$$C_i^{\infty} = \frac{\exp\left((\hat{q}_i^{\infty} - p_i)/\beta\right)}{1 + \sum_{j=1}^{d} \exp\left(\left(\hat{q}_j^{\infty} - p_j\right)/\beta\right)}.$$

The ex post expected value of a new rating of product i in the limit is $\mathsf{E}[r_i \| r_i - \hat{q}_i^{\infty}| > R]$, where r_i is a Gumbel random variable with mean q_i and scale β. Applying the fixed-point argument, we have $\mathsf{E}[r_i \| r_i - \hat{q}_i^{\infty}| > R] = \hat{q}_i^{\infty}$, which is equivalent to $\mathsf{E}[X \| X - (\hat{q}_i^{\infty} - q_i)| > R] = \hat{q}_i^{\infty} - q_i$, where $X = r_i - q_i$. By Lemma 1, if $\hat{q}_i^{\infty} - q_i < 0 < \eta$ or $\hat{q}_i^{\infty} - q_i > \eta$, the equation does not hold.

Corollary 1. *As $n \to \infty$, the underreporting bias satisfies $\hat{q}_i^{\infty} - q_i \in (0, \eta)$ with probability one, where η is defined in* Lemma 1.

As with the acquisition bias, the underreporting bias also introduces a positive effect to the ratings of all the products in the limit compared with their true qualities. The quantity η depends on R and is approximately zero when R is either very small or very large from the proof of Lemma 1.

6. Numerical Studies

We next conduct Monte Carlo simulations of the dynamics to illustrate our results. The main objectives of the section are to demonstrate the transient dynamics of the ratings and market shares, validate our theory, and test the robustness of our main model assumption, that is, the MNL choice model.

6.1. Monte Carlo Simulation

In this section, we demonstrate the dynamics of the average ratings by a simulation study. To avoid biases in selecting parameters, we generate random instances for $d \in \{2, 5, 10, 20\}$ products. For each d, we generate 1,000 instances independently. In each instance, the qualities q_i, prices p_i, and the standard deviation of the experience shock σ_i are drawn independently from the uniform distribution between 0.5 and 2. The value of β is also drawn from the same distribution. We simulate 1,000 customer arrivals according to the random utility model for each instance. The purpose of the randomized instances is to sample model parameters from a reasonable range.

We investigate the following quantities averaged over the instances: the absolute difference (averaged over the products) between the rating and the limit after the 200th/1,000th customer (*RConv200*/*RConv1000*). These two quantities measure the convergence of the ratings to the limit specified in Theorem 1. We investigate the average revenue earned from the 1,000 customers under social learning (*RevSL*) and with known qualities (*RevMNL*), the average consumer surplus of the 1,000 customers under social learning (*SurSL*) and with known qualities (*SurMNL*). We also compute the difference between the limit ratings and the actual quality (*RBias*). Moreover, we compute the optimal assortment under social learning and known qualities and evaluate their expected revenues under social learning. The percentage difference (*RevDiff*) helps us gauge the benefit of taking social learning into account when making operational decisions. We evaluate the revenue of the optimal revenue-ordered assortment and compute the revenue gap in percentage (*RevOrder*). It is worth pointing out that it is computationally intensive to find the optimal assortment under social learning for $d = 20$ as there are more than a million assortments to enumerate (and we have 1,000 instances). Proposition 4 allows us to eliminate a majority of the assortments that don't satisfy the structural properties and facilitate the computation significantly. These quantities are averaged over the 1,000 instances and presented in Table 2.

From the simulation study, we can see that the convergence to the limit may get slower when d increases as each product, on average, has fewer ratings for a fixed number of customers. The revenues under social learning are slightly higher. This is because we generate qualities and prices independently: because an expensive product tends to have lower choice probability in expectation, it gets a boost in demand under social learning. In practice, products of high quality are usually associated with a high price. In this case, the more profitable products might be more attractive, and social learning may even hurt the revenue if the firm does not optimize its operations. Consumer surplus is higher under known qualities as shown in Proposition 3, and the gap doesn't seem to scale with the number of products. The bias of social learning is significant, especially compared with the magnitude

Management Science, 2021, vol. 67, no. 12, pp. 7472–7492, © 2021 INFORMS

Table 2. Results of the Monte Carlo Simulation

d	*RConv200*	*RCon1000*	*RevMNL*	*RevSL*	*SurMNL*	*SurSL*	*RBias*	*RevDiff*, %	*RevOrder*, %
2	0.151	0.079	1.129	1.192	1.697	1.152	1.125	0.1	0.0
5	0.394	0.125	1.103	1.158	2.290	2.092	2.094	0.9	0.0
10	1.118	0.286	1.108	1.146	3.041	2.866	2.931	1.8	0.0
20	2.497	1.003	1.094	1.103	3.910	3.509	3.791	2.4	0.0

of the qualities ([0.5, 2]). Failing to account for the self-selection bias may mislead the firm when designing assortments. The loss accounts for more than 0.9% of the total revenue when $d \geq 5$, and it scales with the number of products. On the other hand, the heuristic provided by Proposition 6 is performing surprisingly well. The average revenue loss is less than 0.1%.

6.2. Robustness of the MNL Assumption

The main theoretical results in this paper are based on a crucial distributional assumption: the utilities of a random customer for the products follow the Gumbel distribution. Although the MNL model is widely used in academia and industry, we test the insights when the assumption fails. In particular, suppose there are four products ($d = 4$). The actual qualities of the products are (1.13, 1.58, 0.50, 0.95). The prices are (0.72, 0.64, 0.78, 1.02). The experience shocks are i.i.d. normal with $\{\sigma_i\}_{i=1}^{d} = (1.10, 1.31, 1.13, 1.53)$. Instead of Gumbel-distributed random utilities in (2) and thus (3), we test two alternative distributions: (1) the ε_i's follow an i.i.d. normal distribution with mean zero and standard deviation one, (2) the ε_i's follow an i.i.d. uniform distribution in the interval $[-2, 2]$. In either case, it is not possible to derive the distribution of the biased ratings as in Proposition 1 and Theorem 1. By Monte Carlo simulation, we want to test whether the main insight still holds: the ratings of all products are biased although niche products benefit more. We simulate 1,000 customer arrivals and record the limiting average ratings (they have converged from the numerics). They can be used to compute the limiting biases because the true qualities are known. Moreover, we can use separate Monte Carlo simulations to approximately compute the choice probabilities when the qualities are known as they do not have closed-form expressions. Table 3 shows that the MNL choice model is not required for our model insights. When a product is more popular (having a higher choice probability), the bias tends to be lower.

Table 3. Limiting Bias and Choice Probabilities with Uniform and Normal Distributed Random Utilities

	Normal		Uniform	
Product	Bias	Choice probability	Bias	Choice probability
1	1.024	0.229	1.204	0.231
2	0.781	0.456	0.994	0.457
3	1.301	0.078	1.349	0.078
4	1.221	0.112	1.175	0.111

7. Conclusion

In this paper, we quantify the acquisition bias when customers report ratings on a platform. The acquisition bias serves as a negative-reinforcing mechanism and makes customer tastes appear more heterogeneous than they actually are. As an unintended consequence, it benefits niche products and hurts popular products in terms of the market share. We point out the implications of this for the firm in its pricing and assortment optimization decisions. Our results give managerial insights into quality, pricing, and assortment decisions and customer feedback and how they are intimately tied together and have to be analyzed jointly. Our research raises many interesting questions and directions for future research on the links between reviews, quality, and the firm's operational and managerial decisions.

Appendix A. Table of Notations

Table A.1. Comprehensive List of Notations Used in the Paper

Notation	Definition
γ	≈ 0.5772, the Euler constant
β	The scale parameter of the Gumbel distribution
$\hat{q}_i(n)$	The average rating of product i observed by customer n
$N_i(n)$	The number of customers purchasing product i prior to customer n
$E_i(n)$	The event that customer n purchases product i
$C_i(n)$	The choice probability $\mathrm{P}(E_i(n))$
$\varepsilon_i(n)$	The random Gumbel utility of product i for customer n
$\xi_i(n)$	The random noise of customer n when experiencing product i
$r_i(n)$	The rating of product i from customer n conditional on $E_i(n)$
$\boldsymbol{X}(n)$	$(\hat{q}_1(n), \ldots, \hat{q}_d(n), N_1(n), \ldots, N_d(n))$
$\boldsymbol{Y}(n)$	$(\hat{q}_1(n), \ldots, \hat{q}_d(n))$
$\lVert\boldsymbol{x}\rVert$	The Euclidean norm of $\boldsymbol{x}$
i.o.	Infinitely often
$\mathbf{1}$	A vector of all ones
$\mathbf{0}$	A vector of all zeros
$\mathbb{I}$	The identity matrix

Appendix B. Proofs

Proof of Proposition 1. Without ambiguity, we remove (n) that is associated with customer n. By the CDF of ε_i in (2), the conditional CDF is

$$
\begin{aligned}
&\mathrm{P}\Big(q_i+\varepsilon_i \le x|E_i(n)\Big)\\
&=\left(\frac{e^{(\hat{q}_i-p_i)/\beta}}{1+\sum_{j=1}^{d}e^{(\hat{q}_j-p_j)/\beta}}\right)^{-1}\int_{-\infty}^{x-q_i}\frac{1}{\beta}e^{(\mu-\varepsilon_i)/\beta}e^{-e^{(\mu-\varepsilon_i)/\beta}}\prod_{j\neq i}e^{-e^{(\mu-\varepsilon_i-\hat{q}_i+p_i+\hat{q}_j-p_j)/\beta}}d\varepsilon_i\\
&=\left(\frac{e^{(\hat{q}_i-p_i)/\beta}}{1+\sum_{j=1}^{d}e^{(\hat{q}_j-p_j)/\beta}}\right)^{-1}\int_{-\infty}^{x-q_i}\frac{1}{\beta}e^{(\mu-\varepsilon_i)/\beta}\prod_{j=0}^{d}e^{-e^{(\mu-\varepsilon_i-\hat{q}_i+p_i+\hat{q}_j-p_j)/\beta}}d\varepsilon_i\\
&=\left(\frac{e^{(\hat{q}_i-p_i)/\beta}}{1+\sum_{j=1}^{d}e^{(\hat{q}_j-p_j)/\beta}}\right)^{-1}\int_{-\infty}^{x-q_i}\frac{1}{\beta}e^{(\mu-\varepsilon_i)/\beta}e^{-e^{(\mu-\varepsilon_i)/\beta}\Sigma_{j=0}^{d}e^{(-\hat{q}_i+p_i+\hat{q}_j-p_j)/\beta}}d\varepsilon_i\\
&=\left(\frac{e^{(\hat{q}_i-p_i)/\beta}}{1+\sum_{j=1}^{d}e^{(\hat{q}_j-p_j)/\beta}}\right)^{-1}\int_{-\infty}^{x-q_i}-e^{-e^{(\mu-\varepsilon_i)/\beta}\Sigma_{j=0}^{d}e^{(-\hat{q}_i+p_i+\hat{q}_j-p_j)/\beta}}de^{(\mu-\varepsilon_i)/\beta}.
\end{aligned}
$$

Letting $y=e^{(\mu-\varepsilon_i)/\beta}\Sigma_{j=0}^{d}e^{(-\hat{q}_i+p_i+\hat{q}_j-p_j)/\beta}$ or, equivalently, $\varepsilon_i=\mu-\beta\log(y)+\beta\log\Big(\Sigma_{j=0}^{d}e^{(-\hat{q}_i+p_i+\hat{q}_j-p_j)/\beta}\Big)$ we have that the preceding expression

$$
\begin{aligned}
&=\left(\frac{e^{(\hat{q}_i-p_i)/\beta}}{1+\sum_{j=1}^{d}e^{(\hat{q}_j-p_j)/\beta}}\right)^{-1}\int_{e^{(\mu-x+q_i)/\beta}\Sigma_{j=0}^{d}e^{(-\hat{q}_i+p_i+\hat{q}_j-p_j)/\beta}}^{+\infty}\\
&\quad\times\frac{1}{\sum_{j=0}^{d}e^{(-\hat{q}_i+p_i+\hat{q}_j-p_j)/\beta}}e^{-y}dy\\
&=e^{-e^{(\mu-x+q_i)/\beta}\Sigma_{j=0}^{d}e^{(-\hat{q}_i+p_i+\hat{q}_j-p_j)/\beta}}\\
&=\exp\left(-e^{\left(\mu+q_i+\log\left(\Sigma_{j=0}^{d}e^{(-\hat{q}_i+p_i+\hat{q}_j-p_j)/\beta}\right)-x\right)/\beta}\right).
\end{aligned}
$$

Compared with the CDF of a standard Gumbel distribution in (2), we can see that the conditional distribution of $r_i(n)$ is a standard Gumbel distribution shifted by $q_i+\log\Big(\Sigma_{j=0}^{d}e^{(-\hat{q}_i+p_i+\hat{q}_j-p_j)/\beta}\Big)$. Therefore, the expected value of $r_i(n)$ conditional on $E_i(n)$ is $q_i+\log\Big(\Sigma_{j=0}^{d}e^{(-\hat{q}_i+p_i+\hat{q}_j-p_j)/\beta}\Big)$, and the conditional variance is the same as a standard Gumbel random variable, which is $\pi^2\beta^2/6$.

Proof of Theorem 1, Part One. To simplify the notation, we set $p_i=q_i=0$ for $i=1,\dots,d$ and $\beta=1$. The general case can be proved similarly. In other words, given $\boldsymbol{X}(n)$, the conditional CDF of $q_i+\varepsilon_i(n)$ is

$$
\mathrm{P}(q_i+\varepsilon_i(n)\le x|E_i(n),\boldsymbol{X}(n)) = \exp\left(-\frac{1+\sum_{j=1}^{d}\exp\left(\hat{q}_j(n)\right)}{\exp\left(\hat{q}_i(n)\right)}e^{(\mu-x)/\beta}\right).
$$

The dynamics of $\boldsymbol{X}(n)$ are still given by (6), and $N_i(n+1)=N_j(n)+1$; $\hat{q}_j(n+1)=\hat{q}_j(n)$, and $N_j(n+1)=N_j(n)$ for $j\neq i$.

The system can be regarded as a special case of stochastic approximation with random step sizes. To make the connection, we can rewrite our stochastic system (given $\boldsymbol{X}(n)$) by

$$
\begin{aligned}
\hat{q}_i(n+1)&=\begin{cases}\hat{q}_i(n)+\dfrac{1}{N_i(n)+1}(r_i(n)-\hat{q}_i(n)) & \textit{if } E_i(n)\\ \hat{q}_i(n) & \textit{if } E_i^c(n)\end{cases}\\
&=\hat{q}_i(n)+\gamma_i(n)(-\log(C_i(n))-\hat{q}_i(n)+\bar{r}_i(n))\\
&=\hat{q}_i(n)+\gamma_i(n)\left(\log(1+\sum_{j=1}^{d}e^{\hat{q}_j(n)}\right)-2\hat{q}_i(n)+\bar{r}_i(n)\Bigg).
\end{aligned}
\tag{B.1}
$$

Here, $\gamma_i(n)$ is a binary-valued random variable that is independent of $r_i(n)$ (again, conditional on $\boldsymbol{X}(n)$) with $\mathrm{P}(\gamma_i(n)=1/(N_i(n)+1))=C_i(n)$ and $\mathrm{P}(\gamma_i(n)=0)=1-C_i(n)$. We decompose $r_i(n)$ into $-\log(C_i(n))+\bar{r}_i(n)$, where $\bar{r}_i(n)$ is a mean-zero random term. We can regard $\gamma_i(n)$ as the random step size and $\log\Big(1+\Sigma_{j=1}^{d}e^{\hat{q}_j(n)}\Big)-2\hat{q}_i(n)+\bar{r}_i(n)$ as the update for stochastic approximation. In other words, denote $\boldsymbol{Y}(n)=(\hat{q}_1(n),\dots,\hat{q}_d(n))\in\mathbb{R}^d$, (B.1) can be rewritten in a vector form:

$$
\boldsymbol{Y}(n+1)=\boldsymbol{Y}(n)+\gamma(n)(h(\boldsymbol{Y}(n))+\bar{\boldsymbol{r}}(n)), \tag{B.2}
$$

where $h(\cdot):\mathbb{R}^d\mapsto\mathbb{R}^d$ is defined as $h_i(Y(n))=\log\Big(1+\Sigma_{j=1}^{d}e^{\mathbf{Y}_j(n)}\Big)-2\mathbf{Y}_i(n)$ and $\gamma(n)$ is the vectorized step size.

The most common and powerful technique to prove the convergence of stochastic approximation is through the ordinary differential equation (ODE) associated with the mean field. The readers can refer to chapter 5 in Kushner and Yin (2003) for a complete treatment and the online note by Combes (2013) for a simplified proof. In particular, consider $\boldsymbol{y}\in\mathbb{R}^d$ and the ODE

$$
\frac{dy_i(t)}{dt}=\log\left(1+\sum_{j=1}^{d}e^{y_j(t)}\right)-2y_i(t)=h(x(t)),\quad -\infty<t<+\infty, \tag{B.3}
$$

which represents the continuous-time analogue to (B.2) when the random term $\bar{\mathbf{r}}(n)$ is simply replaced by its mean zero. The ODE method typically establishes the convergence of $(\boldsymbol{Y}_1(n),\dots,\boldsymbol{Y}_d(n))$ to the equilibrium point $\boldsymbol{y}_\infty$ of the ODE, which is defined to be a point in $\mathbb{R}^d$ such that $\boldsymbol{y}(t)\to\boldsymbol{y}_\infty$ as $t\to\infty$ regardless of the initial position $\boldsymbol{y}(0)$. The convergence usually breaks down to several conditions that the stochastic system has to satisfy (see, e.g., Combes 2013):

1. Lipschitz continuity: there exists $L\ge 0$ such that, for all $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^d$, $\|h(\boldsymbol{x})-h(\boldsymbol{y})\|\le L\|\boldsymbol{x}-\boldsymbol{y}\|$, where $\|\cdot\|$ denotes the Euclidean norm.
2. Diminishing step sizes: for all $i=1,\dots,d$, $\Sigma_{n=1}^{+\infty}\gamma_i(n)=+\infty$ and $\Sigma_{n=1}^{+\infty}\gamma_i(n)^2<+\infty$ almost surely.
3. Martingale difference noise: there exists $K>0$ such that, for all n, we have $\mathrm{E}[\bar{r}_i(n)|\mathbf{X}(n)]=0$ and $\mathrm{E}[\bar{r}_i^2(n)|\mathbf{X}(n)]\le K(1+\|\mathbf{Y}(n)\|)$.
4. Boundedness: $\sup_n\|\mathbf{Y}\|<+\infty$ almost surely.
5. Lyapunov function: there exists a positive (except for $\boldsymbol{y}_\infty$), radially unbounded, continuously differentiable Lyapunov function $V(\boldsymbol{y}):\mathbb{R}^d\mapsto\mathbb{R}$ such that, for all $\boldsymbol{y}\in\mathbb{R}^d$, $\langle\nabla V(\boldsymbol{y}),h(\boldsymbol{y})\rangle\le 0$, and the equality is only attained at $\boldsymbol{y}=\boldsymbol{y}_\infty$.

Note that, different from Combes (2013), we adapt condition two for random step sizes. This is without loss of generality according to Kushner and Yin (2003).

Next, we verify those conditions one by one for our problem. To show condition one, note that, for $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^d$,

$$
\|h(\boldsymbol{x})-h(\boldsymbol{y})\|\le\left|\log\left(1+\sum_{j=1}^{d}e^{x_j}\right)-\log\left(1+\sum_{j=1}^{d}e^{y_j}\right)\right|\sqrt{d}+2\|\boldsymbol{x}-\boldsymbol{y}\|
$$

Because each entry of the gradient of the function $\log\left(1+\Sigma_{j=1}^{d}\exp(x_j)\right)$ is between zero and one, by the mean value theorem, we can show that

$$\left\|\log\left(1+\sum_{j=1}^{d}e^{x_j}\right)-\log\left(1+\sum_{j=1}^{d}e^{y_j}\right)\right\|\le\sum_{i=1}^{d}|x_i-y_i|\le\sqrt{d}\|x-y\|.$$

Therefore, $h(\cdot)$ is uniformly Lipschitz continuous, and by the Picard–Lindelöf theorem, the solution to the ODEs is well defined and unique given the initial condition.

For condition two, note that Lemma C.2 in Appendix C implies that index i is chosen infinitely often. Therefore, $\{1/n\}_{n=1}^{+\infty}$ a subsequence of $\{\gamma_i(n)\}_{n=1}^{+\infty}$, and other entries are zero. It implies condition two.

For condition three, note that, by definition, $\mathsf{E}[\bar{r}_i(n)|\mathbf{X}(n)]=0$. Moreover, by Proposition 1, $\mathsf{E}[\bar{r}_i(n)^2|\mathbf{X}_n]=\pi^2/6+\sigma_i^2$ (recall that $\beta=1$). Clearly, condition three is satisfied.

Condition four is a direct result of Lemma C.1 from Appendix C.

For condition five, we use the globally asymptotic stability established in Lemma C.3 from Appendix C. Then, the condition is guaranteed by the converse Lyapunov theorem (Kellett 2015). Because all five conditions hold in our problem, the convergence is guaranteed. Thus, we have proved part one of the theorem.

Proof of Theorem 1, Part Two. By part one, $\boldsymbol{X}(n)$ must converge to the fixed point, which solves

$$C_i^{\infty}=\frac{e^{\left(\hat{q}_i^{\infty}-p_i\right)/\beta}}{1+\sum_{j=1}^{n}e^{\left(\hat{q}_j^{\infty}-p_j\right)/\beta}}=\frac{e^{(q_i-p_i)/\beta-\log\left(C_i^{\infty}\right)}}{1+\sum_{j=1}^{n}e^{(q_j-p_j)/\beta-\log\left(C_j^{\infty}\right)}}$$
$$\Rightarrow\left(\frac{C_i^{\infty}}{C_j^{\infty}}\right)^2=\frac{e^{(q_i-p_i)/\beta}}{e^{(q_j-p_j)/\beta}}. \tag{B.4}$$

In the first equality, notice that the limiting market share satisfies the MNL model when the quality is replaced by the limiting average rating. The second equality follows the fixed-point argument before Theorem 1. It implies that there exists a constant $C>0$ such that $C_i^{\infty}=Ce^{(q_i-p_i)/(2\beta)}$. Plugging it back into (B.4) gives

$$C_i^{\infty}=\frac{e^{(q_i-p_i)/(2\beta)}}{C+\sum_{j=1}^{n}e^{(q_j-p_j)/(2\beta)}}.$$

Combining the two gives an equation for C:

$$C^2+C\sum_{j=1}^{n}e^{(q_j-p_j)/(2\beta)}-1=0.$$

Because $C>0$, its value is uniquely determined by the equation. Plugging C into the formula, we have completed the proof.

Proof of Proposition 2, Part One. From Theorem 1, the no-purchase probability with known/unknown qualities are

$$\frac{1}{1+\sum_{j=1}^{n}e^{(q_j-p_j)/\beta}}\quad\text{and}\quad 1-C\sum_{j=1}^{n}e^{(q_j-p_j)/(2\beta)}$$
$$=\frac{C}{C+\sum_{j=1}^{n}e^{(q_j-p_j)/(2\beta)}},$$

respectively, where C is defined after (B.4). To show that the former is less, it suffices to show that

$$\sum_{j=1}^{n}e^{(q_j-p_j)/(2\beta)}\ge C\sum_{j=1}^{n}e^{(q_j-p_j)/\beta}.$$

If we can show that $C\le e^{-(q_i-p_i)/(2\beta)}$ for all i, then $Ce^{(q_j-p_j)/\beta}\le e^{(q_j-p_j)/(2\beta)}$, and the claim holds. Note that, from the definition in Theorem 1,

$$C=\frac{-\sum_{j=1}^{n}e^{(q_j-p_j)/(2\beta)}+\sqrt{\left(\sum_{j=1}^{n}e^{(q_j-p_j)/(2\beta)}\right)^2+4}}{2}$$
$$=\frac{2}{\sum_{j=1}^{n}e^{(q_j-p_j)/(2\beta)}+\sqrt{\left(\sum_{j=1}^{n}e^{(q_j-p_j)/(2\beta)}\right)^2+4}}$$
$$<\frac{1}{\sum_{j=1}^{n}e^{(q_j-p_j)/(2\beta)}}<e^{-(q_i-p_i)/(2\beta)}.$$

Therefore, we have proved part one of the proposition.

Proof of Proposition 2, Part Two. From Theorem 1, it is clear from the numerator that C_i^{∞} has the same ranking as C_i when q_i's are known except for the no-purchase option. Therefore, the ranking of the market shares is preserved.

Proof of Proposition 2, Part Three. From Theorem 1, the choice probabilities for product i with known/unknown qualities are

$$\frac{e^{(q_i-p_i)/\beta}}{1+\sum_{j=1}^{n}e^{(q_j-p_j)/\beta}}\quad\textit{and}\quad Ce^{(q_i-p_i)/(2\beta)}=\frac{e^{(q_i-p_i)/(2\beta)}}{C+\sum_{j=1}^{n}e^{(q_j-p_j)/(2\beta)}}.$$

The first term is greater if and only if

$$\frac{e^{(q_i-p_i)/\beta}}{1+\sum_{j=1}^{n}e^{(q_j-p_j)/\beta}}\ge\frac{e^{(q_i-p_i)/(2\beta)}}{C+\sum_{j=1}^{n}e^{(q_j-p_j)/(2\beta)}}$$
$$\Longleftrightarrow e^{(q_i-p_i)/(2\beta)}\ge\frac{1+\sum_{j=1}^{n}e^{(q_j-p_j)/\beta}}{C+\sum_{j=1}^{n}e^{(q_j-p_j)/(2\beta)}}.$$

Note that the term in the right-hand side of the preceding inequality is uniform for all products. Therefore, the choice probability is less under social learning if and only if q_i-p_i is above a cutoff.

Proof of Proposition 3, Part One. This is obvious because the product selected under social learning leads to a lower or equal realized utility than the product with the highest utility, which is chosen when the true qualities are known. As it is true for every realization of the random components, taking expectations, we can easily see that consumer's expected surplus is hurt. Moreover, if we replace the $C_i(n)$ s in (10) by C_i^* s, we can that see (9) = (10).

Proof of Proposition 3, Part Two. To prove part two, we take $C_i(n)$s as decision variables for all $i=0,\ldots,d$ and optimize over them in (10):

$$\begin{aligned}\max_x\quad&\sum_{i=0}^{d}f(x_0,\ldots,x_d)\triangleq(q_i-p_i-\beta\log(x_i))x_i\\ \text{s.t.}\quad&\sum_{i=0}^{d}x_i=1\\ &x_i\ge 0\quad\forall\, i=0,\ldots,d.\end{aligned}\tag{B.5}$$

Dualize the first constraint (B.5) by introducing the Lagrangian multiplier λ. The optimization can be reformulated as

$$\min_{\lambda}\max_{x\ge 0}\sum_{i=0}^{d}(q_i-p_i-\beta\log(x_i))x_i+\lambda\sum_{i=0}^{d}x_i-\lambda. \tag{B.6}$$

The necessary Karush–Kuhn–Tucker (KKT) condition indicates that, at optimality, we need

$$x_i = \exp\left((q_i - p_i + \lambda - \beta)/\beta\right) \quad \forall\, i = 0, \ldots, d.$$

And

$$\sum_{i=0}^{d} x_i = 1.$$

Note that Taylor's expansion of (B.6) leads to

$$f(x_0, \ldots, x_d) - f(C_0^*, \ldots, C_d^*) = \sum_{i=0}^{d} (\lambda - \beta + q_i - p_i - \beta \log(C_i^*))\delta_i + \sum_{i=0}^{d} -\frac{\beta}{C_i^*}\delta_i^2 + O\left(\delta_i^3\right).$$

By the constraint, $\Sigma_{i=0}^{d}\delta_i = 0$. By the KKT condition, $\lambda - \beta + q_i - p_i - \beta\log(C_i^*) = \lambda - \beta + q_j - p_j - \beta\log(C_j^*)$ for all i, j. Therefore, the surplus loss is $\Sigma_{i-0}^{d}\frac{\beta}{C_i^*}\delta_i^2 + O(\delta_i^3)$, which completes the proof.

Proof of Proposition 4. To show the first claim, suppose $i \in S^*$ and $j \notin S^*$. If we replace i in S^* by j, then the numerator of the revenue in (12) is increasing and the denominator is decreasing. This contradicts the fact that S^* is the optimal assortment. Therefore, $j \in S^*$.

To show the second claim, we use the formulation in Rusmevichientong et al. (2010). The optimal assortment problem can be equivalently formulated as

$$\begin{aligned} \max \quad & \lambda \\ \text{subject to} \quad & \frac{\sum_{i\in S} 2p_i \exp\left((q_i - p_i)/(2\beta)\right)}{\sum_{i\in S} e^{(q_i - p_i)/(2\beta)} + \sqrt{\left(\sum_{i\in S} e^{(q_i-p_i)/(2\beta)}\right)^2 + 4}} = \lambda, \end{aligned} \tag{B.7}$$

Or, equivalently,

$$\begin{aligned} \max \quad & \lambda \\ \text{subject to} \quad & \sum_{i\in S}\left(p_i - \frac{\lambda}{2}\right)\exp\left((q_i - p_i)/(2\beta)\right) \\ & = \frac{\sqrt{\left(\sum_{i\in S} e^{(q_i-p_i)/(2\beta)}\right)^2 + 4}}{2}\lambda. \end{aligned} \tag{B.8}$$

Suppose λ^* and S^* attain the optimality. We show that, if $p_i < \lambda^*/2$, then $i \notin S^*$. Otherwise, removing such i from S^* increases the left-hand side of (B.8) and decreases the coefficient of the right-hand side of (B.8). Thus, λ^* can be further increased, which contradicts the optimality of λ^*. Therefore, if $p_i < \lambda^*/2$, then $i \notin S^*$. On the other hand, (B.7) is also equivalent to

$$\begin{aligned} \max \quad & \lambda \\ \text{subject to} \quad & \sum_{i\in S}(p_i - \lambda)\exp\left((q_i - p_i)/(2\beta)\right) \\ & = \frac{2}{\sqrt{\left(\sum_{i\in S} e^{(q_i-p_i)/(2\beta)}\right)^2 + 4} + \sum_{i\in S} e^{(q_i-p_i)/(2\beta)}}\lambda. \end{aligned} \tag{B.9}$$

We show that, if $p_i > \lambda^*$, then $i \in S^*$. Otherwise, adding such i to S^* increases the left-hand side of (B.9) and decreases the coefficient of the right-hand side of (B.9). Again, λ^* can be further increased, and the claim holds by contradiction. Combining these results, if $p_j < 2p_i$, then it implies that $p_j > 2p_i \geq \lambda^*$. Therefore, $j \in S^*$.

To show the third claim, simply note that the product with the highest revenue must satisfy $p_1 > \lambda^*$ from (B.7).

Proof of Proposition 5. We prove this by a reduction from the well-known two-partition problem, which is defined as follows:

Definition B.1 (Two-Partition). Given a set of d nonnegative rational numbers $v_1, v_2, \ldots, v_d$, determine whether there is a set $S \subseteq \{1, 2, \ldots, d\}$ such that $\Sigma_{i\in S} v_i = \Sigma_{i\in\{1,2,\ldots,d\}\setminus S} v_i$.

According to Garey and Johnson (1979), the two-partition problem is NP-complete.

The reduction works as follows. Starting with any instance of two-partition, we design an instance of Problem (12). We show that the solution to (12) takes a certain value if and only if there is a solution to the two-partition problem. Therefore, solving (12) equivalently solves the two-partition problem. Because two-partition is NP-complete, this relationship establishes the NP-hardness of (12).

Given any instance of two-partition, that is, áv_i for $i = 1, \ldots, d$, let $T = \frac{1}{2}\Sigma_{i\in\{1,2,\ldots,d\}} v_i$. Note that $\Sigma_{i\in S} v_i = \Sigma_{i\in\{1,2,\ldots,d\}\setminus S} v_i$ if and only if $\Sigma_{i\in S} v_i = T$. Because scaling all v_i by a constant does not change the solution to two-partition, we let $v_i' \leftarrow v_i \frac{100}{51}\frac{1}{T}$, then $\Sigma_{i\in S} v_i = \Sigma_{i\in\{1,2,\ldots,d\}\setminus S} v_i$ if and only if $\Sigma_{i\in S} v_i' = \frac{100}{51}$.

Next, we construct an instance of Problem (12) as follows. Consider $d+1$ products indexed by $1, 2, \ldots, d, d+1 >$ and set the revenues of the products as

$$p_i = \begin{cases} \frac{9}{5} & i = 1, \ldots, d \\ \frac{7}{2} & i = d+1, \end{cases}$$

and set $\exp((q_i - p_i)/(2\beta))$ as

$$\exp\left((q_i - p_i)/(2\beta)\right) = \begin{cases} v_i' & i = 1, \ldots, d \\ \frac{12}{17} & i = d+1, \end{cases}$$

and now, we show that the two-partition problem has a solution if and only if there exists an assortment $S \subseteq \{1, \ldots, d, d+1\}$ whose expected revenue is at least two. Note that $r_{d+1} > r_i$ for $i = 1, \ldots, d$. According to claim three in Proposition 4, product $d+1$ is always in the optimal assortment. Therefore, having an assortment $S \subseteq \{1, \ldots, d, d+1\}$ whose expected revenue is at least two is equivalent to having

$$\max_{S\subset\{1,2,\ldots,d\}} \frac{\sum_{i\in S}\frac{9}{5}v_i' + \frac{7}{2}\cdot\frac{12}{17}}{\sum_{i\in S} v_i' + \frac{12}{17} + \sqrt{\left(\sum_{i\in S} v_i' + \frac{12}{17}\right)^2 + 4}} \geq 1.$$

It is straightforward to verify that this inequality holds if and only if we can find an assortment $S \subseteq \{1, 2, \ldots, d\}$ such that

$$-\left(\frac{3}{5}\sum_{i\in S} v_i' - \frac{20}{17}\right)^2 \geq 0.$$

And note that the preceding inequality holds if and only if $\Sigma_{i\in S}v_i' = \frac{100}{51}$, which is equivalent to $\Sigma_{i\in S}v_i = \Sigma_{i\in\{1,2,\ldots,d\}\setminus S}v_i$. Therefore, we show the NP-hardness of the assortment optimization problem.

Proof of Proposition 6. We first introduce some notations. Because p_i and q_i are exogenously given in the assortment optimization problem, we let $v_i \triangleq \exp((q_i - p_i)/(2\beta))$. Introduce $R(S)$ as the revenue of assortment S and $V(S) \triangleq \Sigma_{i\in S}v_i$.

Let S^* be the optimal assortment and $i^* = \max_j\{j\in S\}$; that is, i^* is the index for the product with the lowest price in S^*. From the proof of Proposition 4, we have shown that $p_{i^*} \geq R(S^*)/2$. Next, we construct a revenue-ordered assortment that guarantees at least1/2 of the optimal revenue. More precisely, consider $\bar{S} = \{1,2,\ldots,i^*\}$. That is, we include all the products with prices higher than p_{i^*} into the assortment. Clearly, $\bar{S}$ is revenue ordered and is a superset of S^*. Its expected revenue is

$$\begin{aligned}
R(\bar{S}) &= \frac{2\sum_{i=1}^{i^*}p_iv_i}{\sum_{i=1}^{i^*}v_i + \sqrt{\left(\sum_{i=1}^{i^*}v_i\right)^2+4}}\\
&= \frac{2\sum_{i\in S^*}p_iv_i + 2\sum_{i\in\bar{S}\setminus S^*}p_iv_i}{V(\bar{S}) + \sqrt{V(\bar{S})^2+4}}\\
&= \frac{R(S^*)\left(V(S^*)+\sqrt{V(S^*)^2+4}\right) + \sum_{j\in S\setminus S^*}2p_jv_j}{V(\bar{S}) + \sqrt{V(\bar{S})^2+4}}\\
&\geq \frac{R(S^*)\left(V(S^*)+\sqrt{V(S^*)^2+4}\right) + R(S^*)V(\bar{S}\setminus S^*)}{V(\bar{S}) + \sqrt{V(\bar{S})^2+4}}\\
&= R(S^*)\frac{V(\bar{S})+\sqrt{V(S^*)^2+4}}{V(\bar{S}) + \sqrt{V(\bar{S})^2+4}}\\
&\geq R(S^*)\frac{V(\bar{S})+2}{V(\bar{S}) + \sqrt{V(\bar{S})^2+4}} \geq \frac{1}{2}R(S^*).
\end{aligned}$$

The first inequality follows from the fact that $p_i \geq R(S^*)$ for $i \geq i^*$. Therefore, we have found a revenue-ordered assortment that generates at least $\frac{1}{2}$ of the optimal revenue, and so does the best revenue-ordered assortment.

Proof of Proposition 7. Denote the revenue $\Sigma_{i=1}^d p_iC_i^\infty$ by $R(\mathbf{p})$. The first-order condition yields

$$\begin{aligned}
\frac{\partial R(\mathbf{p})}{\partial p_i} &= \frac{\partial}{\partial p_i}\frac{2\sum_{j=1}^d p_je^{(q_j-p_j)/(2\beta)}}{\sum_{j=1}^d e^{(q_j-p_j)/(2\beta)} + \sqrt{\left(\sum_{j=1}^d e^{(q_j-p_j)/(2\beta)}\right)^2+4}}\\
&= \frac{2e^{(q_i-p_i)/(2\beta)}}{a}\left(1-\frac{p_i}{2\beta}+\frac{p_i\sum_{j=1}^d p_je^{(q_j-p_j)/(2\beta)}}{2\beta\sqrt{\left(\sum_{j=1}^d e^{(q_j-p_j)/(2\beta)}\right)^2+4}}\right) = 0
\end{aligned}$$

where we denote the denominator $\Sigma_{j=1}^d e^{(q_j-p_j)/(2\beta)} + \sqrt{\left(\Sigma_{j=1}^d e^{(q_j-p_j)/(2\beta)}\right)^2+4}$ by a. The first-order condition implies that either $p_i = +\infty$ or $p_i = c_0$ for some constant C_0. By a similar argument to Gallego and Wang (2014), we conclude that $p_i = +\infty$ is not optimal and all prices must be equal.

To show the second claim, note that, by Gallego and Wang (2014), the optimal prices for all the products are equal when the qualities are known. Therefore, the first-order condition yields

$$\frac{d}{dp}\left(\frac{\sum_{i=1}^d pe^{(q_i-p)/\beta}}{1+\sum_{i=1}^d e^{(q_i-p)/\beta}}\right) = 0 \Rightarrow e^{p/\beta} - e^{p/\beta}\frac{p}{\beta} + \sum_{i=1}^d e^{q_i/\beta} = 0. \quad \text{(B.10)}$$

In the presence of social learning, by denoting $a = \Sigma_{j=1}^d e^{q_i/(2\beta)}$, we have

$$\begin{aligned}
&\frac{d}{dp}\frac{2pe^{-p/(2\beta)}\sum_{j=1}^d e^{q_i/(2\beta}}{e^{-p/(2\beta)}\sum_{j=1}^d e^{q_j/(2\beta)} + \sqrt{e^{-p/\beta}\left(\sum_{j=1}^d e^{q_j/(2\beta)}\right)^2+4}}\\
&= 2a\left(\frac{1}{a+\sqrt{a^2+4e^{p/\beta}}} - \frac{2e^{p/\beta}p}{\beta\sqrt{a^2+4e^{p/\beta}}(a+\sqrt{a^2+4e^{p/\beta}})^2}\right)\\
&= \frac{2a}{a+\sqrt{a^2+4e^{p/\beta}}}\left(1-\frac{2p}{\beta\sqrt{e^{-p/\beta}a^2+4}(ae^{-p/(2\beta)}+\sqrt{a^2e^{-p/\beta}+4})}\right).
\end{aligned} \quad \text{(B.11)}$$

Because the term in the parentheses of (B.11) is decreasing in p, it suffices to show that, when we plug the solution to (B.10) into (B.11), the term is positive. Denote the solution to (B.10) by p_1. The positivity is equivalent to

$$\begin{aligned}
&a+\sqrt{a^2+4e^{p_1/\beta}} - \frac{2e^{p_1/\beta}p_1}{\beta\sqrt{a^2+4e^{p_1/\beta}}} > 0\\
&\Longleftrightarrow a+\sqrt{a^2+4e^{p_1/\beta}} - \frac{2e^{p_1/\beta}+2\sum_{i=1}^d e^{q_i/\beta}}{\sqrt{a^2+4e^{p_1/\beta}}} > 0\\
&\Longleftrightarrow a\sqrt{a^2+4e^{p_1/\beta}} + a^2 + 4e^{p_1/\beta} > 2e^{p_1/\beta} + 2\sum_{i=1}^d e^{q_i/\beta}.
\end{aligned}$$

Because $a^2 = \left(\Sigma_{i=1}^d e^{q_i/(2\beta)}\right)^2 > \Sigma_{i=1}^d e^{q_i/\beta}$, the inequality holds. Therefore, we have proved the result.

Proof of Proposition 8.

1. Fix the optimal price given by revealing true qualities. By the first part of Proposition 2, the total probability of purchase under social learning is higher and, therefore, achieves more revenue (because of the single common price). If now we optimize for prices under social learning, we just improve the revenue even more.

2. Now, consider the assortment optimization under known qualities. Denote the optimal revenue by R^k, that is,

$$R^k \triangleq \max_S \frac{\sum_{i\in S}p_i\exp\left(\frac{q_i-p_i}{\beta}\right)}{1+\sum_{i\in S}\exp\left(\frac{q_i-p_i}{\beta}\right)}. \quad \text{(B.12)}$$

At optimality, suppose the optimal assortment is S^k. Then, we should have

$$R^k = \sum_{i\in S^k}\left(p_i - R^k\right)\exp\left(\frac{q_i - p_i}{\beta}\right). \tag{B.13}$$

This can be derived by multiplying both sides of (B.12) by the denominator and combining common terms. Notice that, for S^k to be optimal, it must be $p_i \geq R^k$.

Now, let us show that, under social learning without information about the true qualities, the total expected revenue from offering S^k must be no smaller than the revenue under known qualities, R^k; that is, $R(S^k) \geq R^k$, where

$$R(S^k) \triangleq \frac{\sum_{i\in S^k} 2p_i \exp\left(\frac{q_i-p_i}{2\beta}\right)}{\sum_{i\in S^k}\exp\left(\frac{q_i - p_i}{2\beta}\right) + \sqrt{\left(\sum_{i\in S^k}\exp\left(\frac{q_i-p_i}{2\beta}\right)\right)^2 + 4}} \tag{B.14}$$

is the total expected revenue from assortment S^k through social learning.

Let us multiply both sides of Equation (B.14) and combine common terms. Then, we would have

$$R(S^k) \triangleq \frac{\sum_{i\in S^k}\left(2p_i - R(S^k)\right)\exp\left(\frac{q_i-p_i}{2\beta}\right)}{\sqrt{\left(\sum_{i\in S^k}\exp\left(\frac{q_i-p_i}{2\beta}\right)\right)^2 + 4}}. \tag{B.15}$$

In other words, the expected revenue $R(S^k)$ from assortment S^k must satisfy equilibrium Equation (B.15).

Let us define the right-hand side of Equation (B.15) as a function with respect to a value r, that is,

$$F(r) \triangleq \frac{\sum_{i\in S^k}(2p_i - r)\exp\left(\frac{q_i-p_i}{2\beta}\right)}{\sqrt{\left(\sum_{i\in S^k}\exp\left(\frac{q_i-p_i}{2\beta}\right)\right)^2 + 4}}.$$

We can see that $F(r)$ is monotonically decreasing in r, and $R(S^k)$ is the unique root of $F(R(S^k)) = R(S^k)$.

As a result, to show $R(S^k) \geq R^k$, we only need to show

$$F(R^k) \geq R^k.$$

Equivalently,

$$\frac{\sum_{i\in S^k}\left(2p_i - R^k\right)\exp\left(\frac{q_i-p_i}{2\beta}\right)}{\sqrt{\left(\sum_{i\in S^k}\exp\left(\frac{q_i-p_i}{2\beta}\right)\right)^2 + 4}} \geq R^k,$$

which is equivalent to

$$\frac{\sum_{i\in S^k}2\left(p_i - R^k\right)\exp\left(\frac{q_i-p_i}{2\beta}\right) + R^k\sum_{i\in S^k}\exp\left(\frac{q_i-p_i}{2\beta}\right)}{\sqrt{\left(\sum_{i\in S^k}\exp\left(\frac{q_i-p_i}{2\beta}\right)\right)^2 + 4}} \geq R^k.$$

Denote $X_i \triangleq (p_i - R^k)\exp\left(\frac{q_i-p_i}{\beta}\right)$. Then the preceding inequality can be written as

$$\frac{\sum_{i\in S^k}2X_i\exp\left(-\frac{q_i-p_i}{2\beta}\right) + R^k\sum_{i\in S^k}\exp\left(\frac{q_i-p_i}{2\beta}\right)}{\sqrt{\left(\sum_{i\in S^k}\exp\left(\frac{q_i-p_i}{2\beta}\right)\right)^2 + 4}} \geq R^k.$$

Because $\Sigma_{i\in S^k}X_i = R^k$ because of Equation (B.13), let us replace R^k by $\Sigma_{i\in S^k}X_i$. In the meantime, to ease notation, let us denote $Y_i \triangleq \exp\left(\frac{q_i-p_i}{2\beta}\right)$. So, equivalently, we want to show

$$\sum_{i\in S^k}2X_i\frac{1}{Y_i} + \left(\sum_{i\in S^k}X_i\right)\left(\sum_{i\in S^k}Y_i\right) \geq \left(\sum_{i\in S^k}X_k\right)\sqrt{\left(\sum_{i\in S^k}Y_i\right)^2 + 4}.$$

This is trivially true as the coefficients before X s are comparable as

$$2\frac{1}{Y_i} + \sum_{i\in S^k}Y_i \geq \sqrt{\left(\sum_{i\in S^k}Y_i\right)^2 + 4}.$$

Thus, $R(S^k) \geq R^k$ gets proved. As S^k is a suboptimal assortment under social learning, we can ascertain that the optimal revenue under social learning must be greater than it is under known qualities.

Proof of Lemma 1. To prove the first property, note that

$$E\left[X \| X-\mu| > R\right] = \frac{E\left[XI_{\{|X-\mu|>R\}}\right]}{P(|X-\mu| > R)}.$$

As $\mu \to \infty$, $P(|X-\mu| > R) \to 1$ and $\mathrm{E}[XI_{\{|X-\mu|<R\}}] \to \mathrm{E}[X] = 0$. Thus, the result follows naturally.

To prove the second property, without loss of generality, we set $\beta = 1$. The CDF and PDF of X are

$$F(x) = \exp(-\exp(-x-\gamma)), \quad f(x) = \exp(-x-\gamma-\exp(-x-\gamma)),$$

where γ is the Euler constant. The underreported bias can be expressed as

$$\mathrm{E}\left[X \| X-\mu| > \mathrm{R}\right] = \frac{\mathrm{E}\left[XI_{\{|X-\mu|>R\}}\right]}{\mathrm{P}(|X-\mu|>R)} = \frac{-\int_{\mu-R}^{\mu+R} xf(x)dx}{F(\mu-R) + 1 - F(\mu+R)}.$$

It suffices to show the same property for the numerator; that is, $-\int_{\mu-R}^{\mu+R} xf(x)dx = 0$ has a unique solution $\mu = \eta$ and η>0. Because $xf(x)$ is negative for $x < 0$ and positive for $x > 0$, if $\mu \notin (-R, R)$, then $-\int_{\mu-R}^{\mu+R} xf(x)dx \neq 0$. Moreover, by continuity, there exists $\eta \in (-R, R)$ such that

$$\int_{\eta-R}^{0} xf(x)dx + \int_{0}^{\eta+R} xf(x)dx = 0.$$

If $\mu \in (-R, \eta)$, then $\int_{\mu-R}^{0} xf(x)dx < \int_{\eta-R}^{0} xf(x)dx$ and $\int_{0}^{\mu+R} xf(x)dx < \int_{0}^{\eta+R} xf(x)dx$ and $-\int_{\mu-R}^{\mu+R} xf(x)dx > 0$. A similar argument can be applied to $\mu \in (\eta, R)$. Hence, such η is unique.

Next, we show that η>0. By the uniqueness of η, it suffices to show that $-\int_{-R}^{R} xf(x)dx > 0$ for all $R > 0$. Define $g(t) = \int_{-t}^{t} xf(x)dx$ for $t \geq 0$. We have

$$g'(t) = t(f(t) - f(-t)) = t\exp(-\gamma)(\exp(-t-\exp(-t-\gamma)) - \exp \times (t - \exp(t-\gamma))).$$

We examine the sign of $g'(t)$ for $t > 0$ t \gt 0, which is always equal to the sign of $\exp(-t-\exp(-t-\gamma)) - \exp(t - \exp(t-\gamma))$ and, thus, the sign of $-2t - \exp(-t-\gamma) + \exp(t-\gamma)$. It is easy to see that, as t increases, $-2t - \exp(-t -\gamma) + \exp(t-\gamma)$ is first negative and then positive. Therefore, $g(t)$ is first decreasing and then increasing. Combined with the fact that $g(0) = 0$ and $\lim_{t\to+\infty} g(t) = 0$, we have $g(t) < 0$ for all $t > 0$.

Appendix C. Auxiliary Lemmas and Proofs

Lemma C.1. Recall $\boldsymbol{Y}(n) = (\hat{q}_1(n), \dots, \hat{q}_d(n))$ in the proof of Theorem 1. We have $\sup_n \|\boldsymbol{Y}(n)\| < +\infty$ almost surely.

Proof of Lemma C.1. Define $c = \max\{1, \log(d+1)\}$. We first show that the stochastic process $Z(n) \triangleq \max\{c, \max_{1\le j\le d} Y_j(n)\}$ is a supermartingale. Conditional on $\boldsymbol{X}(n)$ and $E_i(n)$, $Z(n+1)$ can either be $\max\{c, \max_{j\ne i} \boldsymbol{Y}_j(n)\}$ (because $\boldsymbol{Y}_j(n+1)$ are not updated for $j \ne i$) or $\mathsf{E}[\boldsymbol{Y}_i(n+1)|\boldsymbol{X}(n), E_i(n)]$. The former is less than $Z(n)$. For the latter, we have

$$\begin{aligned}\mathsf{E}[\boldsymbol{Y}_i(n+1)|\boldsymbol{X}(n), E_i(n)] &= \boldsymbol{Y}_i(n) + \frac{1}{N_i(n)+1}(-\log(C_i(n)) - \boldsymbol{Y}_i(n))\\ &= \boldsymbol{Y}_i(n) + \frac{1}{N_i(n)+1}\left(\log\left(1+\sum_{j=1}^{d} e^{\boldsymbol{Y}_j(n)}\right) - 2\boldsymbol{Y}_i(n)\right).\end{aligned} \tag{C.1}$$

Because $\boldsymbol{Y}_j(n) \le Z(n)$ for all j and $1 \le Z(n)$, we have

$$\begin{aligned}(28) &\le \boldsymbol{Y}_i(n) + \frac{1}{N_i(n)+1}\Big(\log(d+1) + Z(n) - 2\boldsymbol{Y}_i(n)\Big)\\ &= \left(1 - \frac{2}{N_i(n)+1}\right)\boldsymbol{Y}_i(n) + \frac{1}{N_i(n)+1} \times \Big(\log(d+1) + Z(n)\Big) \le Z(n),\end{aligned}$$

where the last inequality is because $N_i(n) \ge 1$, $\boldsymbol{Y}_i(n) \le Z(n)$, and $\log(d+1) \le Z(n)$. This implies that $\mathsf{E}[Z(n+1)|\boldsymbol{X}(n), E_i(n)] \le Z(n)$. Summing over the index i, we have proved that $\mathsf{E}[Z(n+1)|\boldsymbol{X}(n)] \le Z(n)$ and $Z(n)$ is a supermartingale. Because $Z(n)$ is nonnegative, by the martingale convergence theorem (2.10, section 4.2, Durrett 2007), $Z(n)$ converges almost surely to a finite random variable.

Similarly, we can show that $\min\{0, \min_{1\le j\le n} \boldsymbol{Y}_j(n)\}$ is a nonpositive submartingale and, thus, converges almost surely to a finite random variable. Therefore,

$$\sup_n \|\boldsymbol{Y}(n)\| < \sqrt{d} \sup_n \max \times \left\{Z(n), -\min\left\{0, \min_{1\le j\le d} \boldsymbol{Y}_j(n)\right\}\right\} + \infty.$$

Lemma C.2. For all $1 \le i \le d$, $\lim_{n\to+\infty} N_i(n) = +\infty$ almost surely.

Proof of Lemma C.2. The event $\{\omega : \lim_{n\to+\infty} N_i(n) = +\infty\}$ is equivalent to $\{A_n\, i.o.\}$, where A_n denotes the event that index i is chosen at n and i.o. represents infinitely often. By the second Borel–Cantelli lemma (3.2, section 4.3, Durrett 2007),

$$\begin{aligned}\{A_n\, i.o.\} &= \left\{\sum_{n=1}^{+\infty} \mathsf{P}(i \text{ being chosen at } n+1|\mathsf{F}_n) = +\infty\right\}\\ &= \left\{\sum_{n=1}^{+\infty} C_i(n+1) = +\infty\right\}\\ &= \left\{\sum_{n=1}^{+\infty} \frac{e^{X_i(n)}}{1+\sum_{j=1}^{d} e^{X_j(n)}} = +\infty\right\}\\ &\supseteq \left\{\sup_n \|\mathbf{X}(n)\| < +\infty\right\}.\end{aligned}$$

The last inequality is because of the following fact: if $\boldsymbol{Y}(n)$ is bounded, then $\Sigma_{n=1}^{+\infty} \frac{e^{Y_i(n)}}{1+\Sigma_{j=1}^{d} e^{Y_j(n)}}$ diverges. By Lemma C.1, the last event is of probability one. Therefore, $\lim_{n\to+\infty} N_i(n) = +\infty$ almost surely.

Lemma C.3. Let $\boldsymbol{y} \in \mathbb{R}^d$ >. The following system of ODEs

$$\frac{dy_i(t)}{dt} = \log\left(1 + \sum_{j=1}^{d} e^{\boldsymbol{y}_j(t)}\right) - 2\boldsymbol{y}_i(t), \quad -\infty < t < +\infty$$

is globally asymptotically stable in the Lyapunov sense. That is, for every trajectory $\boldsymbol{y}(t)$, we have $\boldsymbol{y}(t) \to \boldsymbol{y}_\infty$ as $t \to \infty$ for a unique equilibrium point $\boldsymbol{y}_\infty$.

Proof of Lemma C.3. Typically, the asymptotic stability of ODEs is shown by constructing a Lyapunov function explicitly. However, for the nonlinear ODEs that we are considering, finding a Lyapunov function is not straightforward. Therefore, we prove the globally asymptotic stability directly.

We first define $\boldsymbol{y}_\infty$ to be the solution to$h(\boldsymbol{y}) = \boldsymbol{0}$, where $h(\boldsymbol{y})$ is defined in (B.2). Note that $h(\boldsymbol{y}) = 0$ implies that $2\boldsymbol{y}_i = \log\left(1 + \Sigma_{j=1}^{d} e^{\boldsymbol{y}_j(t)}\right)$, which implies that $\boldsymbol{y}_i$ are all equal for $i = 1, \dots, d$. Therefore, the components of $\boldsymbol{y}_\infty$ are all equal, which are denoted y_∞. As a result, we have $e^{2y_\infty} = 1 + de^{y_\infty}$, which has a unique solution. It implies that $\boldsymbol{y}_\infty$ is uniquely defined.

For $d \ge 2$, consider the transformation $z \in \mathbb{R}^{d-1}$:$z_i = \boldsymbol{y}_i - \boldsymbol{y}_d$ for $i = 1, \dots, d-1$. Clearly, $\boldsymbol{y}$ satisfies the following ODE:

$$\frac{dz_i(t)}{dt} = -2z_i(t). \tag{C.2}$$

Therefore, $z_i(t) = e^{-2t} z_i(0)$. As a result, for any $\varepsilon > 0$, there exists $T(\varepsilon, \boldsymbol{y}(0))$ such that, for $t > T(\varepsilon, \boldsymbol{y}(0))$, we have $z_i(t)| < \varepsilon$ for all i.

Once $\boldsymbol{y}(t)$ reaches the region $C_\varepsilon \triangleq \{|\boldsymbol{y}_i - \boldsymbol{y}_d| < \varepsilon, \forall\, i\}$, define the function

$$V(\boldsymbol{y}) = \sum_{i=1}^{d}\left(\log\left(1 + \sum_{j=1}^{d} e^{\boldsymbol{y}_j(t)}\right) - 2\boldsymbol{y}_i(t)\right)^2.$$

We show that it is a Lyapunov function for $\boldsymbol{y} \in C_\varepsilon$ when $\varepsilon = \log(1.5)/2$. Clearly, $V(\boldsymbol{y})$ is continuously differentiable, radially unbounded, and $V(\boldsymbol{y}) > 0$ except for $\boldsymbol{y} = \boldsymbol{y}_\infty$. To show it is a Lyapunov function, we only need to show that $\langle \nabla V(\boldsymbol{y}), h(\boldsymbol{y})\rangle < 0$ except for $\boldsymbol{y} = \boldsymbol{y}_\infty$. Note that

$$\begin{aligned}\langle \nabla V(\boldsymbol{y}), h(\boldsymbol{y})\rangle &= \sum_{i=1}^{d}\left(\frac{e^{\boldsymbol{y}_i}}{1+\sum_{j=1}^{d} e^{\boldsymbol{y}_j}} - 2\right)\left(\log\left(1+\sum_{j=1}^{d} e^{\boldsymbol{y}_j}\right) - 2\boldsymbol{y}_i\right)^2\\ &\quad + \sum_{i=1}^{d}\sum_{j\ne i} \frac{e^{\boldsymbol{y}_i}}{1+\sum_{j=1}^{d} e^{\boldsymbol{y}_j}}\left(\log\left(1+\sum_{l=1}^{d} e^{\boldsymbol{y}_l}\right) - 2\boldsymbol{y}_i\right)\\ &\quad \times \left(\log\left(1+\sum_{l=1}^{d} e^{\boldsymbol{y}_l}\right) - 2\boldsymbol{y}_j\right)\\ &= \boldsymbol{w}^T M \boldsymbol{w}.\end{aligned}$$

Here, $\boldsymbol{w} \in \mathbb{R}^d$ and $\boldsymbol{w}_i$ denotes $\log\left(1 + \Sigma_{j=1}^{d} e^{\boldsymbol{y}_j}\right) - 2\boldsymbol{y}_i$; $M \in \mathbb{R}^{d\times d}$ is a symmetric matrix and satisfies

$$\begin{aligned}M_{ii} &= \frac{e^{\boldsymbol{y}_i}}{1+\sum_{j=1}^{d} e^{\boldsymbol{y}_j}} - 2;\\ M_{ij} &= \frac{e^{\boldsymbol{y}_i} + e^{\boldsymbol{y}_j}}{2\left(1+\sum_{j=1}^{d} e^{\boldsymbol{y}_j}\right)}, \quad \forall\, i \ne j.\end{aligned}$$

For $\boldsymbol{y} \in C_\varepsilon$, where $\varepsilon = \log(1.5)/2$, we can show that

$$\begin{aligned} M_{ii} + \sum_{j \neq i} |M_{ij}| &= -2 + \frac{de^{y_i} + \sum_{j=1}^{d} e^{y_j}}{2\left(1 + \sum_{j=1}^{d} e^{y_j}\right)} \\ &\leq -2 + \frac{2de^{y_d+\varepsilon}}{2(1 + de^{y_d-\varepsilon})} \\ &< -2 + e^{2\varepsilon} = -0.5. \end{aligned}$$

By the Gershgorin circle theorem, the maximal eigenvalue of M is less than or equal to -0.5. Therefore, M is negative definite, and $\langle \nabla V(\boldsymbol{y}), h(\boldsymbol{y}) \rangle < 0$ except for $\boldsymbol{y} = \boldsymbol{y}_\infty$. Thus, the two facts established, (1) the global attraction of the set C_ε and (2) the existence of a Lyapunov function for $\boldsymbol{y} \in C_\varepsilon$, lead to the globally asymptotic stability of the ODE for $d \geq 2$.
When d = 1, we have

$$\log(1 + e^y) - 2y = \begin{cases} < 0 & y > y_\infty \\ > 0 & y < y_\infty. \end{cases}.$$

As a result, the unique y_∞ is globally asymptotically stable.

Endnotes

[1] Informally, *quality* is an intrinsic attribute reflecting the value of the product that may or may not be known to the actors. Garvin (1987) in an influential article dissects the various dimensions of quality. For our purposes, it is sufficient to treat it as an abstract, fixed, item-specific constant positively correlated to value of the product. We introduce a formal definition in Section 3.

[2] We also discuss briefly in Section 5 an extension to a different type called underreporting bias.

[3] We consider the optimization over prices in Section 4.

[4] The "quality" of the no-purchase option $q_0 = 0$

References

Abeliuk A, Berbeglia G, Cebrian M, Van Hentenryck P (2015) The benefits of social influence in optimized cultural markets. *PLoS One* 10(4):1–20.
Acemoglu D, Ozdaglar A (2011) Opinion dynamics and learning in social networks. *Dynamic Games Appl.* 1(1):3–49.
Acemoglu D, Dahleh MA, Lobel I, Ozdaglar A (2011) Bayesian learning in social networks. *Rev. Econom. Stud.* 78(4):1201–1236.
Acemoglu D, Makhdoumi A, Malekian A, Ozdaglar A (2018) Fast and slow learning from reviews. Working paper. Massachusetts Institute of Technology, Cambridge, MA.
Anderson SP, de Palma A (1992) Multiproduct firms: A nested logit approach. *J. Indust. Econom.* 40(3):261–276.
Banerjee AV (1992) A simple model of herd behavior. *Quart. J. Econom.* 107(3):797–817.
Ben-Akiva M, Lerman S (1985) *Discrete-Choice Analysis: Theory and Application to Travel Demand* (MIT Press, Cambridge, MA).
Besbes O, Scarsini M (2018) On information distortions in online ratings. *Oper. Res.* 66(3):597–610.
Bikhchandani S, Hirshleifer D, Welch I (1992) A theory of fads, fashion, custom, and cultural change as informational cascades. *J. Political Econom.* 100(5):992–1026.
Combes R (2013) An introduction to stochastic approximation. Accessed June 28, 2018, https://people.kth.se/~alepro/DistriOptCourse/lecture_stoch_approx.pdf.
Crapis D, Ifrach B, Maglaras C, Scarsini M (2016) Monopoly pricing in the presence of social learning. *Management Sci.* 63(11):3586–3608.
Dellarocas C, Wood CA (2008) The sound of silence in online feedback: Estimating trading risks in the presence of reporting bias. *Management Sci.* 54(3):460–476.
Dellarocas C, Zhang XM, Awad NF (2007) Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *J. Interactive Marketing* 21(4):23–45.
DeMarzo PM, Vayanos D, Zwiebel J (2003) Persuasion bias, social influence, and unidimensional opinions. *Quart. J. Econom.* 118(3):909–968.
Durrett R (2007) *Probability: Theory and Examples*, 3rd ed. (Cambridge University Press, New York, NY).
Engler TH, Winter P, Schulz M (2015) Understanding online product ratings: A customer satisfaction model. *J. Retailing Consumer Services* 27:113–120.
Fradkin A, Grewal E, Holtz D, Pearson M (2015) Bias and reciprocity in online reviews: Evidence from field experiments on Airbnb. *Proc. 16th ACM Conf. Econom. Comput.* (EC '15) (Association for Computing Machinery, New York, NY), 641.
Gallego G, Wang R (2014) Multiproduct price optimization and competition under the nested logit model with product-differentiated price sensitivities. *Oper. Res.* 62(2):450–461.
Garey MR, Johnson DS (1979) *Computers and Intractability: A Guide to the Theory of NP-Completeness* (W.H. Freeman & Co., New York).
Garvin DA (1987) Competing on the eight dimensions of quality. *Harvard Bus. Rev.* 65(6):101–109.
Godes D, Silva JC (2012) Sequential and temporal dynamics of online opinion. *Marketing Sci.* 31(3):448–473.
Hu N, Pavlou PA, Zhang J (2017) On self-selection biases in online product reviews. *Management Inform. Systems Quart.* 41(2):449–471.
Juslin P, Winman A, Hansson P (2007) The naive intuitive statistician: A naive sampling model of intuitive confidence intervals. *Psych. Rev.* 114(3):678–703.
Kahneman D (2003) Maps of bounded rationality: Psychology for behavioral economics. *Amer. Econom. Rev.* 93(5):1449–1475.
Kellett CM (2015) Classical converse theorems in Lyapunov's second method. *Discrete Continuous Dynamical Systems B* 20(8):2333–2360.
Koh NS, Hu N, Clemons EK (2010) Do online reviews reflect a product's true perceived quality? An investigation of online movie reviews across cultures. *Electronic Commerce Res. Appl.* 9(5):374–385.
Kushner H, Yin GG (2003) *Stochastic Approximation and Recursive Algorithms and Applications*, vol. 35 (Springer Science & Business Media, New York).
Lee YJ, Hosanagar K, Tan Y (2015) Do I follow my friends or the crowd? Information cascades in online movie ratings. *Management Sci.* 61(9):2241–2258.
Li X, Hitt LM (2008) Self-selection and information role of online product reviews. *Inform. Systems Res.* 19(4):456–474.
Li X, Hitt LM (2010) Price effects in online product reviews: An analytical model and empirical analysis. *Management Inform. Systems Quart.* 34(4):809–831.
Maldonado F, Hentenryck PV, Berbeglia G, Berbeglia F (2018) Popularity signals in trial-offer markets with social influence and position bias. *Eur. J. Oper. Res.* 266(2):775–793.
Molavi P, Tahbaz-Salehi A, Jadbabaie A (2018) A theory of non-Bayesian social learning. *Econometrica* 86:445–490.
Mossel E, Sly A, Tamuz O (2015) Strategic learning and the topology of social networks. *Econometrica* 83(5):1755–1794.
Muchnik L, Aral S, Taylor SJ (2013) Social influence bias: A randomized experiment. *Sci.* 341(6146):647–651.
Papanastasiou Y, Savva N (2016) Dynamic pricing in the presence of social learning and strategic consumers. *Management Sci.* 63(4):919–939.
Podium (2017) 2017 state of online reviews. Accessed June 28, 2018, http://learn.podium.com/rs/841-BRM-380/images/2017-SOOR-Infographic.jpg.

Rusmevichientong P, Shen ZM, Shmoys DB (2010) Dynamic assortment optimization with a multinomial logit choice model and capacity constraint. *Oper. Res.* 58(6):1666–1680.

Salganik MJ, Dodds PS, Watts DJ (2006) Experimental study of inequality and unpredictability in an artificial cultural market. *Sci.* 311(5762):854–856.

Smallwood DE, Conlisk J (1979) Product quality in markets where consumers are imperfectly informed. *Quart. J. Econom.* 93(1): 1–23.

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

Smith JE, Winkler RL (2006) The optimizer's curse: Skepticism and postdecision surprise in decision analysis. *Management Sci.* 52(3): 311–322.

Talluri K, Van Ryzin G (2004) Revenue management under a general discrete choice model of consumer behavior. *Management Sci.* 50(1):15–33.

Thaler RH (1988) Anomalies: The winner's curse. *J. Econom. Perspect.* 2(1):191–202.

Tong J, Feiler D, Ivantsova A (2018) Good choice, bad judgment: How choice under uncertainty generates overoptimism. *Psych. Sci.* 29(2):254–265.

Vaccari S, Maglaras C, Scarsini M (2018) Social learning from online reviews with product choice. Working paper, Columbia University, New York.

Ying Y, Feinberg F, Wedel M (2006) Leveraging missing ratings to improve online recommendation systems. *J. Marketing Res.* 43(3):355–365.