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

OPERATIONS RESEARCH
Vol. 66, No. 3, May–June 2018, pp. 597–610
ISSN 0030-364X (print), ISSN 1526-5463 (online)

CROSSCUTTING AREAS

On Information Distortions in Online Ratings

Omar Besbes,[a] **Marco Scarsini**[b]

[a] Graduate School of Business, Columbia University, New York, New York 10027; [b] Dipartimento di Economia e Finanza, LUISS, 00197 Roma, Italy

Contact: ob2105@columbia.edu, http://orcid.org/0000-0002-2982-3794 (OB); marco.scarsini@luiss.it, http://orcid.org/0000-0001-6473-794X (MS)

Received: May 16, 2013
Revised: March 1, 2015; December 16, 2016
Accepted: July 27, 2017
Published Online in Articles in Advance: February 2, 2018

Subject Classifications: probability:stochastic model applications; information systems: analysis and design; decision analysis sequential
Area of Review: Games, Information, and Networks

https://doi.org/10.1287/opre.2017.1676

Abstract. Consumer reviews and ratings of products and services have become ubiquitous on the Internet. This paper analyzes, given the *sequential* nature of reviews and the limited feedback of such past reviews, the information content they communicate to future customers. We consider a model with heterogeneous customers who buy a product of unknown quality and we focus on two different informational settings. In the first setting, customers observe the whole history of past reviews. In the second one they only observe the sample mean of past reviews. We examine under which conditions, in each setting, customers can recover the true quality of the product based on the feedback they observe. In the case of total monitoring, if consumers adopt a fully rational Bayesian updating paradigm, then they asymptotically learn the unknown quality. With access to only the sample mean of past reviews, inference becomes intricate for customers and it is not clear if, when, and how social learning can take place. We first analyze the setting when customers interpret the mean as the proxy of quality. We show that in the long run, the sample mean of reviews stabilizes and, in general, customers overestimate the underlying quality of the product. We establish properties of the bias, stemming from the selection associated with observing only reviews of customers who purchase. Then, we show the existence of a simple non-Bayesian quality inference rule that leads to social learning when all customers use such a rule. The results point to the strong information content of even limited statistics of past reviews as long as customers have minimal sophistication.

Funding: Marco Scarsini is a member of GNAMPA-INdAM. His work was partially supported by PRIN 20103S5RN3 and IDG31300110.

Keywords: online reviews • quality inference • sequential analysis • social learning • Bayesian updating • stochastic approximation algorithms

1. Introduction

1.1. Motivation

The use of reviews has become ubiquitous on the Internet, where websites allow users to comment on the products or services they purchased. Typically the review consists of a grade and a written comment. The grade may be expressed in different ways, the most common one being the five star system, used, among others, by Amazon, Yelp, and Expedia. Such grades, or their aggregate statistics, are often the basis for other users' decisions, who in turn may post their reviews. The impact of reviews on rated services or products has been reported to be significant, for example, attributing 5%–9% increase in revenue to an increase in one star on Yelp (Luca 2011).

Reviews are used by consumers before making a decision to purchase an item. Once a consumer has purchased the item, she may write a review that expresses her level of satisfaction. Typically consumers do not know the exact quality of the object that they intend to purchase, for instance they either ignore or would not be able to interpret the technical features of a new smartphone. In addition, consumers are heterogeneous and may react differently to the purchase of the same object. The satisfaction of a consumer depends both on the intrinsic quality of what she bought and on her idiosyncratic attitude toward the purchase. The same object could meet the needs and standards of one consumer while being inadequate for another.

If consumers were perfectly homogeneous when writing a review, they would implicitly reveal the quality of the object. Heterogeneity of the consumers creates a problem both ex ante and ex post. Ex ante, only consumers who expect to be satisfied by the product purchase the item. Ex post, only the reviews of consumers who bought the product are available. Given this bias, it is not clear what information reviews communicate and whether consumers can learn the quality of the product over time. The present paper aims to investigate this question and what type of review information leads to social learning.

To that end, we consider a model with an infinite population of heterogeneous consumers who sequentially decide whether to buy a product of unknown quality. If they buy the product, they receive a utility that depends on the unknown quality and on their type, which is random private information and is stochastically independent of the quality. Once a consumer decides to buy a product, she uses it and writes a review that declares her utility. The decision to either buy the product or to pass is driven by the expected utility of the product given the available information. A prior on the quality of the product is assumed to be common knowledge.

1.2. Why Ratings

One of the issues that arises when dealing with ratings is why people write reviews. Even if in most cases there is no immediate reward for writing reviews, people write them abundantly. This is not a marginal phenomenon that can be classified as a pathology. One common explanation is that consumers write reviews just out of altruism. They do it to be helpful toward other consumers. If this were the case, then they would write reviews that are as informative and as useful as possible to future consumers. That is, they would write about the intrinsic quality of the product that they bought (if they can identify it), rather than about their own subjective experience when they consumed the product. A second possible explanation is that consumers write reviews for the same reasons they post pictures on Instagram or comments on Facebook or Twitter. They just want to express themselves and declare who they are, what they feel, what they think. In this case they will write more about their subjective experience than about the quality of what they bought.

While we will comment on the altruistic case, it seems intuitive that in such a case, customers should learn over time. This indeed happens in the model we consider here. The focus of the present paper is mainly on the case in which customers report their subjective experience with the product. In such a case, it is not clear how to interpret the reviews one sees or their average, and whether current customers can communicate useful information to future customers.

1.3. Main Contributions

We analyze a hierarchy of informational settings, going from a setting in which customers observe all past reviews and are fully rational to a setting in which customers only observe the average of past reviews and are boundedly rational.

Consumers are assumed to be heterogeneous, so only the ones whose expected utility for the product is positive will make the purchase and then write a review. The main case we analyze is when customers post a review that reflects their personal experience with the product.

In the fully rational model with the availability of all past reviews, we analyze the conditional expectation of the quality estimate of Bayesian customers. We show that despite the selection bias introduced by purchases, customers learn in the long run the true quality of the product. In other words, social learning takes place.

Then, we analyze a setting in which customers base their inference on the mean of past reviews. This can be interpreted as a true information limitation or, alternatively, as a situation in which customers do not have the computational power to process all past reviews and base their inference only on the sample mean of posted reviews. In this case the inference is intricate because the mean is not a sufficient statistic. The first basic question in this setting is how customers should interpret the sample mean of posted reviews and how this relates to the underlying quality of the product.

We first consider the case in which customers are naïve and interpret at face value the average of reviews that they see posted. In this case, we show that the sample mean of reviews stabilizes over time and admits an almost sure limit. If only a fraction of consumers buys the product, which happens if the outside option is appealing to some consumers, then customers consistently overestimate the true quality in the long run. In this case after N customers have considered whether to purchase the product, the customer population incurs welfare losses of order N, because of the lack of knowledge of the quality.

We then ask whether there exists an inference rule that leads to social learning if customers are more "sophisticated." Since only the mean of reviews is observed, a fully Bayesian customer would need to account for all possible paths that lead to the observed mean, a highly intractable and too demanding task. To obviate this computational issue, we develop a simple approximation to the conditional expectation of quality given the observed sample mean of past reviews. We establish that, if, prior to making a purchasing decision, customers use this approximation that corrects for the selection bias associated with observed reviews, then, as long as all customers use the same rule, social learning does take place. In other words, the bias-correcting estimate converges to the true quality almost surely. In addition, the order of magnitude of the customer welfare losses stemming from the lack of knowledge of the true quality is bounded by order $\sqrt{N}$. Quite remarkably, this is the same order of losses one would obtain if customers were altruistic reviewers.

Based on the above framework, we then consider the question pertaining to the interplay between naïve and sophisticated customers. We show that while sophisticated customers can still learn the true quality, the losses per naïve consumer decrease with the fraction of naïve consumers; so naïve consumers benefit from being surrounded by other naïve consumers.

Revisiting the motivation and the question of the informativeness of reviews, the heterogeneity of consumers and the fact that the processes of review generation and purchases are inherently intertwined makes the interpretation task of past reviews and quality inference challenging for consumers in practice. The fact that rational consumers can eventually learn the true quality with an access to the full history of reviews points to the fact that reviews are very informative for future consumers when properly interpreted. Furthermore, even if only the sample mean of past reviews is observable, while taking this mean at face value by naïve consumers leads to overestimating the quality, minimal sophistication of consumers when interpreting the reviews is sufficient to recover the informativeness of past reviews. In that sense, the past reviews, and even just their sample mean, are quite informative. Furthermore, full Bayesian rationality is not required for reviews to benefit consumers as even potentially simple, non-Bayesian rules with access to partial information about past reviews can lead to social learning.

1.4. Literature Review

The literature on social learning can be traced back to the seminal papers of Banerjee (1992) and Bikhchandani et al. (1992) (see also Bikhchandani et al. 1998). In this literature no reviews are present and learning occurs (or fails to occur) only by sequentially observing the behavior of the previous customers. It is assumed that they choose one of two options without knowing for sure which one is better. Each customer is endowed with a signal, namely, a random variable that stochastically depends on the unknown quality of the available options. The choice of one of the two options is based on this signal, which is private information, and on the observed behavior of the previous customers. When the prior probability that one option is better than the other is close to $1/2$ and signals are bounded, it is quite possible that the first consumers make the wrong decision, causing a cascade of wrong decisions, so with positive probability learning does not occur. This phenomenon is the outcome of a Bayesian Nash equilibrium of a game that formalizes the model, where all consumers are fully rational and optimize their behavior; this makes the phenomenon quite interesting and puzzling and has led to a stream of follow-up studies.

For instance, Smith and Sørensen (2000) show that the above-mentioned herd behavior is due to boundedness of the signals. When these are unbounded, so the very strong signal of one customer can overwhelm the observations of the previous customers' behavior, then social learning occurs. Çelen and Kariv (2004) examine the issue of social learning when consumers' information is partial, namely, they observe only the action of their immediate predecessor. Banerjee and Fudenberg (2004) look at a word-of-mouth learning model where each consumer samples a certain number of previous consumers before making her decision. They consider conditions for social learning. Goeree et al. (2006) examine a model where consumers are of different types and the number of possible actions and states is an arbitrary finite number. They find general conditions for convergence of the beliefs to the true state. Herrera and Hörner (2013) look at a model where consumers arrive at random times on the market and only the decision of buying the product is observed. They show that, even under this partial observation, unbounded signals guarantee complete learning, whereas with bounded signals cascades may occur. Conditions for the occurrence of social learning have been recently studied in great generality by Arieli and Mueller-Frank (2017).

A recent stream of literature has introduced a network structure in the learning problem. For instance, in Acemoglu et al. (2011) each customer observes the actions of a stochastically generated neighborhood of individuals before making her own decision. They find conditions under which social learning occurs. A nice review of the main contributions can be found in Acemoglu and Ozdaglar (2011).

A different stream of literature deals with online reviews, in which the feedback is related to customer satisfaction. Crapis et al. (2017) study social learning of heterogeneous consumers who, before deciding whether to buy a product, observe the number of people who liked it or disliked it, and, after buying, publicly declare whether they themselves liked it or not. Agents are assumed to be boundedly rational and social learning is studied via mean-field approximation techniques. Pricing strategies for the seller are studied.

Our paper relates to that of Ifrach et al. (2013) in that both include a model with reviews and without signals, customers' heterogeneity, and sequential purchasing decisions. However, the information available (past reviews versus aggregate statistics), the unknown quality (binary versus arbitrary), and type of reporting (like/dislike versus full report) are different. In addition, the two papers focus on fundamentally different questions. Their paper studies the details of the Bayesian update dynamics. In contrast, the present paper focuses on the bounded rationality of consumers, and when/if social learning may still take place based on the sample mean of reviews when updating is performed in a non-Bayesian way.

Epstein et al. (2010) propose a framework for non-Bayesian learning. More closely related to our work is the paper by Jadbabaie et al. (2012), who consider a model where consumers communicate over a social network and update their information in a non-Bayesian way. They provide conditions for learning to occur in this setting. A connected issue was studied by Acemoglu et al. (2014), whose model considers agents on a graph who are endowed with private signals about an unknown state of the world and can establish costly communication links before making an action. The agents' payoff depends on their actions and on the unknown state. The authors study asymptotic learning.

In the present paper, we explicitly capture the bias that might be present in reviews. Li and Hitt (2008) study some reasons why online reviews may fail to provide information about quality. One is possible manipulation to create artificially high ratings. Another is that, even if reviews accurately reflect earlier consumers opinions, those opinions may not be representative of the opinions of the broader consumer population in later time periods. Early buyers may then have a bias that can be either positive or negative. Building on such observations, Papanastasiou et al. (2014) consider a particular type of manipulation of early buyers' opinions via short-term stock-outs. Lafky (2014) experimentally deals with the issue of why people rate products. Zhang et al. (2014) use data to interpret review ratings and to study their distribution.

1.5. Organization of the Paper

The paper is organized as follows. Section 2 formulates the model. Section 3 analyzes the setting in which customers observe and can process all past reviews. Section 4 examines the case in which consumers base their decisions only on the average of existing reviews, and it studies learning or lack thereof when customers are either naïve, sophisticated, or a mixture of these two types. Section 5 concludes with some possible extensions and avenues for future research. All the proofs are relegated to the appendix.

2. The Model

We consider a marketplace that aggregates consumer reviews for products and services. We focus on a single product or service being introduced on the market. Consumers arrive sequentially over time, examine the product and some summary statistic of the reviews posted by previous consumers, make a decision of whether to purchase, and, if they purchase, report a review.

2.1. Utility Model

We consider an infinite sequence of customers, who sequentially decide to either buy a product or to pass. If customer n buys the product, then she receives a utility

$$Y_n = Q_n + \Theta_n.$$

Otherwise, the utility she receives is τ. The variable Y_n represents customer n's satisfaction derived from the object. It is the sum of two components. The term Q_n is the evaluation of the quality of the product by consumer n, which is assumed to be observable only *after* the purchase of the product. The term Θ_n is customer n's idiosyncratic attitude toward the product, which is observed *before* the purchase of the product. The sequences $\{Q_n\}_{n\in\mathbb{N}}$ and $\{\Theta_n\}_{n\in\mathbb{N}}$ are assumed to be independent; moreover the Θ_n are i.i.d. and the Q_n are conditionally i.i.d. given Q, with conditional mean Q. We refer to Q as the intrinsic quality of the product and we define $\varepsilon_n := Q_n - Q$. The random variables Q, ε_n and Θ_n all have bounded support, whose convex hulls are the intervals $[\underline{q}, \bar{q}]$, $[\underline{\varepsilon}, \bar{\varepsilon}]$, and $[\underline{\theta}, \bar{\theta}]$, respectively. Just like ε_n, the variables Θ_n are assumed to have zero mean. This is without loss of generality.

The variable Θ_n is private information to customer n. The other customers do not know it, although they know its distribution. All customers share the same prior distribution π_0 for Q. Note that Θ_n does not carry any information about the quality Q. So it is not a signal in the sense of the classical social learning literature.

2.2. Information Structure and Purchasing Decisions

We assume that upon arrival to the system, customer n observes some information on past reviews, decides whether to make a purchase or not, and in the event of a purchase, she posts a review X_n expressing her evaluation of the product.

For $n \in \mathbb{N}$, let Z_n denote the purchasing decision of customer n, i.e.,

$$Z_n = \begin{cases} 1 & \text{if customer } n \text{ buys the product,} \\ 0 & \text{if customer } n \text{ does not buy the product.} \end{cases}$$

Let $b_0 = 0$ and $\{b_n: n \geqslant 1\}$ denote the random subsequence of buying consumers. For $n \geqslant 1$ define

$$B(n) := \max\{k \in \mathbb{N}: b_k < n\}. \tag{1}$$

Therefore $B(n)$ is the number of customers who purchased the product before customer n.

Inference with all past reviews. The first benchmark case we study is one in which customers have access to all past reviews. Since only customers who purchase the product post reviews, the satisfaction becomes measurable only if a customer buys the object. Customer n's evaluation of the product, as seen by future customers, is then given by

$$X_n = \begin{cases} Y_n & \text{if } Z_n = 1, \text{ i.e., } n \text{ buys the product,} \\ * & \text{if } Z_n = 0, \text{ i.e., } n \text{ does not buy the product.} \end{cases}$$

Define

$$\mathscr{F}_n^X := \sigma(X_1, \ldots, X_n)$$

the σ-field generated by the first n evaluations and $\boldsymbol{\mathcal{F}}^X = \{\mathscr{F}_n^X\}_{n\in\mathbb{N}}$ the corresponding filtration.

The sequence $\{b_n\}$ is a sequence of stopping times adapted to the filtration $\boldsymbol{\mathcal{F}}^X$. Define

$$\mathscr{F}_n^b := \sigma(X_{b_1}, \ldots, X_{b_n}) \tag{2}$$

to be the σ-field generated by the satisfactions of the first n buying customers and $\boldsymbol{\mathcal{F}}^b = \{\mathcal{F}^b_n\}_{n\in\mathbb{N}}$ the corresponding filtration.

Since the variables Θ_n are independent of Q, the σ-fields $\mathcal{F}^X_{n-1}$ and $\mathcal{F}^b_{B(n)}$ carry the same information for the conditional distribution of Q. In other words, knowing that a previous customer did not buy the product does not add any valuable information for the following customers.

We assume that customers are risk neutral and have an outside option with utility τ. The optimal strategy for customer n is to buy the product if and only if

$$\mathbb{E}[Q \mid \mathcal{F}^X_{n-1}] + \Theta_n \geqslant \tau.$$

Remark 1. In this setting, one could define a game with a countable number of players (the consumers) where the action space of each player is $\{0,1\}$ (i.e., not to buy or buy), and the payoff of player n is the utility derived from her action Z_n, i.e.,

$$(Q_n + \Theta_n)Z_n + \tau(1 - Z_n).$$

Given that in our setting the players after player n have no influence on her payoff and the players before player n influence her payoff only through the information they reveal in their reviews, a profile of individually optimal decisions gives rise to a pure perfect Bayesian equilibrium. In our case the equilibrium profile $(Z_n)_{n\in\mathbb{N}}$ is given by

$$Z_n = \mathbb{1}\{\mathbb{E}[Q \mid \mathcal{F}^X_{n-1}] + \Theta_n \geqslant \tau\}.$$

A related setting can be found in Acemoglu et al. (2011). There, players sequentially take a binary action and their payoff depends on a binary state of nature. Their model involves a social network and players receive information only from neighbors in this network.

Inference with limited feedback: The mean of past reviews. We will also focus on the case in which the public information available to customer $n+1$ is just the average $\bar{X}_n$ of the reviews written by the previous customers who bought the product, where

$$\bar{X}_n := \begin{cases} * & \text{if } B(n+1) = 0, \\ \dfrac{1}{B(n+1)} \displaystyle\sum_{i=1}^{n} X_i Z_i & \text{if } B(n+1) \geqslant 1. \end{cases} \tag{3}$$

In other words, the customer only observes whether there was at least one review and in such a case, the mean of all reviews. Note that we do not assume that the customer knows her position n or the number of past reviews. One could interpret this setting as one in which the customers have minimal information about the history of reviews or in which customers have bounded rationality and cannot process all reviews.

As earlier, we assume that customers are risk neutral and have an outside option with utility τ. The optimal strategy for customer $n+1$ is to buy the product if and only if

$$\mathbb{E}[Q \mid \bar{X}_n] + \Theta_n \geqslant \tau.$$

In this model with reduced information, the inference is quite intricate for customers. From a technical perspective, there is no filtration that represents the information available to customers, and therefore one cannot resort to typical Doob-martingale results. More importantly, from a customer's perspective, and from a processing power perspective, it is intractable to perform exact Bayesian updating to compute $\mathbb{E}[Q \mid \bar{X}_n]$. In the present paper, we analyze different approximations of $\mathbb{E}[Q \mid \bar{X}_n]$ that customers may use and analyze the resulting learning behavior that may follow.

For both settings with full or limited feedback, we will be interested in understanding under what conditions social learning takes place, i.e., when customers recover the value of Q in the long run.

2.3. Assumptions

Assumption 1. *The random variables Θ_n, $n \geqslant 1$ admit a continuously differentiable density and are such that*

$$\delta := \mathbb{P}\{\Theta_n > \max\{\tau - \underline{q},\, \bar{q} - \underline{q} - \underline{\varepsilon},\, (\tau - \underline{q} - \underline{\varepsilon})/2\}\} > 0.$$

This condition ensures that customers are sufficiently heterogeneous so that there is always a nontrivial probability that a customer purchases, independently of the beliefs about the quality of the product. As a consequence, customers purchase the product infinitely often and hence the number of reviews posted increases to ∞ as $n \uparrow \infty$.

Remark 2 (Altruistic Customers)**.** In the present paper, the fact that customers report their personal evaluations of the product is the main driver of the challenges above. If consumers were altruistic and aimed at conveying the most useful information to the other potential consumers, consumer n, after buying the product, would reveal Q_n in her review. In such a case, the sequence Θ_n becomes simply a driver of who purchases but would have no bearing on the reviews and, therefore, inference could be carried out according to the usual parametric Bayesian paradigm with a sample of conditionally i.i.d. random variables.

In addition, we assume throughout the paper that the reviews submitted are truthful and there is no cost in writing reviews. The former is without loss of generality since players have no incentive to misreport their review in our model.

3. Inference with All Past Reviews

In this section, we assume that the information available to customer n before she makes her purchasing decision is $\mathscr{F}^X_{n-1}$, i.e., she observes all past reviews. Furthermore, we assume that customers are fully rational Bayesian players. Define

$$M_n = \mathbb{E}[Q \mid \mathscr{F}^X_{n-1}]. \tag{4}$$

So, M_n is the conditional expectation of Q that customer n holds after observing the reviews of the previous customers. As mentioned before, this is equal to the conditional expectation of Q after observing only the reviews of customers who purchased the product, that is

$$\mathbb{E}[Q \mid \mathscr{F}^X_{n-1}] = \mathbb{E}[Q \mid \mathscr{F}^b_{B(n)}].$$

Customer n will purchase if and only if $M_n + \Theta_n \geqslant \tau$, that is, if and only if her realization of Θ_n is such that $\Theta_n \geqslant \tau - M_n$. We next analyze the dynamic evolution of M_n.

Theorem 1. *Suppose that Assumption 1 holds. Then $M_n \to Q$ almost surely as $n \uparrow \infty$.*

In other words, Theorem 1 establishes that social learning takes place when all reviews are observed, i.e., customers end up learning the true value of Q in the long run despite the interplay across reviews and purchases. In this setting, reviews indeed play the important role of disseminating the real characteristics of the product, and, as a consequence, allowing customers to make the right purchasing decision in the long run. The fundamental element of the learning process is that customers keep buying the product, and information accumulates over time, that is, almost surely the sequence b_n diverges. This is guaranteed by Assumption 1.

To prove the result, we first note that M_n is a Doob Martingale, so it converges almost surely to a random variable M_∞. The main challenge in the proof is to establish that $M_\infty = Q$. The sequential nature of the purchasing decisions and the consequent lack of exchangeability of the reviews makes the learning proof more intricate than what we would have in a standard statistical Bayesian model. To establish the result, we introduce an auxiliary sequence whose nth element is $\mathscr{F}^X_n$-measurable and prove, using a stochastic approximation analysis, that this sequence converges to the true quality Q and hence the latter is measurable with respect to the filtration $\mathscr{F}^X_\infty$. This then allows to use a Doob-type argument to establish that the almost sure limit of M_n is indeed Q.

Discussion (observing actions versus reviews). Our model drastically differs from most of the literature about social learning, which does not include reviews. Starting with Banerjee (1992) and Bikhchandani et al. (1992), in almost all the articles in this stream of literature customers possess private information about the unknown quality of a product. This private information comes in the form of (typically i.i.d.) random signals. Each customer observes the decision of the previous customers and uses Bayes's rule to combine her signal with her observations and to compute her posterior distribution of the product's quality. Based on this, she decides whether to buy or not to buy the product. Given the private signal, all purchase decisions are informative and no further information comes from any agent beyond the fact that she did or did not purchase the product. In this framework, cascades can happen and with positive probability, even asymptotically, customers do not learn the quality of the object. The notable feature is that herd behavior and the consequent learning failure is not due to any form of limited rationality of the customers; it is the result of a Bayesian equilibrium in a game with incomplete information.

Goeree et al. (2006) consider a model with heterogeneous consumers and finite but not necessarily binary state space. Their model is in the classical stream of social learning literature in that consumers learn from observing the behavior and not the opinions of other consumers. There are nevertheless some analogies with our model. For instance, in their model all actions are chosen with positive probability in every state of the world. This is comparable with our Assumption 1 that, independently of the true quality, some consumers will always buy the product.

In our setting, the purchasing decision is not informative at all. This is because of the fact that agents do not have private signals. The utility component Θ_n in our model is independent of the quality Q, therefore after observing a customer buying the product or passing, the information of the subsequent customers remains the same. Only the reviews provide information. With respect to learning, the challenges stem from heterogeneity of the customers and the purchasing bias: only customers with high enough Θ_n purchase the product and the threshold varies with n and is history dependent. Therefore, the fact that learning can be achieved is not self-evident. Our results show that customers are actually able to filter the sampling bias induced by the Θ_n and asymptotically learn the quality of the object, so that the fraction of customers who make a purchasing mistake tends to zero.

4. Inference with the Average of Past Reviews

In this section, we analyze the case in which customers base their decisions on the mean of past reviews. This setting has a dual motivation: sometimes not all past reviews are available but only aggregate statistics such as the mean. Alternatively, the processing power of customers can be limited and anchoring around the mean of past reviews is very common. In this context, we will be interested in evaluating $\mathbb{E}[Q \mid \bar{X}_n]$, which is the key quantity that customer $n+1$ will compute to make a purchasing decision.

In the current context, purchasing decisions, and therefore reviews are history dependent and as a result, it is not clear how much information the sample mean of past reviews contains. In particular, the sample mean is not even a sufficient statistic for the history of past reviews. This raises a significant question with regard to the computation of $\mathbb{E}[Q \mid \bar{X}_n]$. As highlighted in Section 2, the exact computation of $\mathbb{E}[Q \mid \bar{X}_n]$ is intractable. To achieve it, we should in principle consider all histories that lead to a value of $\bar{X}_n$ and, for each history, compute the corresponding conditional expectation and finally average over all these histories. This intractable task motivates the introduction of approximating procedures for the computation of $\mathbb{E}[Q \mid \bar{X}_n]$.

Next we will introduce two types of possible behavior that customers can adopt to make inference from the observation of $\bar{X}_n$.

The main question we will focus on here is whether learning can take place despite the fact that customers base their decisions only on the mean of past reviews. Furthermore, we will refine our analysis to also measure the customer welfare losses stemming from the lack of knowledge of the product quality Q. For a given realization of Q, we define such losses, after N customers have joined the system, as

$$\mathscr{L}_Q(N) := \sum_{i=1}^{N} \mathbb{E}[(Q+\Theta_i-\tau)^+ \mid Q] - \mathbb{E}[(Q+\Theta_i-\tau)Z_i \mid Q].$$

The first term in the sum corresponds to the expected surplus of a customer at the time of purchase decision when she knows the value of Q. The second term corresponds to the expected surplus when customer i decides according to Z_i. Hence the difference between the two is exactly the loss in customer surplus attributable to the lack of knowledge of the value of Q.

4.1. Average of Reviews as a Proxy for Quality

We start with the case of customers who interpret $\bar{X}_n$ as the underlying quality of the product and approximate $\mathbb{E}[Q \mid \bar{X}_n]$ by $\bar{X}_n$, as long as at least one review has already been posted. We refer to such customers as naïve. In particular, customer $n+1$ forms an estimate $\tilde{q}_{n+1}$ of the true quality of the product Q as follows: for $n \geqslant 0$,

$$\tilde{q}_{n+1} = \begin{cases} \mathbb{E}[Q] & \text{if } B(n+1) = 0, \\ \bar{X}_n & \text{if } B(n+1) \geqslant 1. \end{cases} \tag{5}$$

By convention, when no one has yet purchased, we assume that customers approximate Q by its mean. This is inconsequential and any other starting point in $[\underline{q}, \bar{q}]$ would lead to similar results. In turn, based on the estimate above, customer $(n+1)$'s purchase decision is given by

$$Z_{n+1} = \begin{cases} 1 & \text{if } \tilde{q}_{n+1} + \Theta_{n+1} \geqslant \tau, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

In this context, each customer influences the information and the decision of future customers as her decision to purchase and the review she posts leads to a new average of reviews available to the next customers.

Theorem 2 (Naïve Customers). *Suppose that Assumption 1 holds. Suppose further that all customers use the approximation (5), and make purchase decisions according to (6) for $n \geqslant 1$. Then, we have the following*:

(a) *The sequence $\tilde{q}_n$ converges almost surely to some limit $\tilde{x}(Q) \geqslant Q$, as $n \uparrow \infty$.*

(b) *The limit $\tilde{x}(Q)$ is almost surely the unique solution of $x = Q + \mathbb{E}[\Theta \mid \Theta > \tau - x]$ and is such that $\tilde{x}(Q)$ is nondecreasing in Q and the bias $\tilde{x}(Q) - Q$ is nonincreasing in Q.*

(c) *The long-term average loss is given by*

$$\lim_{N\uparrow\infty} \frac{1}{N} \mathscr{L}_Q(N) = \int_{\tau-\tilde{x}(Q)}^{\tau-Q} [Q+z-\tau]\, \mathrm{d}F_\Theta(z) \textit{ almost surely.}$$

The fact that the quantity $\bar{X}_n$ does not necessarily converge to the true quality Q is something that can easily be expected given the naïve rule used for quality inference and resulting purchasing decisions. The relevance of Theorem 2(a) is that, independently of the underlying distributions of Θ and ε, the sample mean of reviews stabilizes in the long run and almost surely converges. The proof of Theorem 2(a) is based on framing the evolution of the running average through the lens of a Robbins-Monro type stochastic approximation algorithm, which enables us to identify the only candidate limit points and, in turn, to establish almost sure convergence.

Furthermore, part (b) characterizes the behavior of the bias as a function of the realized quality Q. Theorem 2(b) has two important implications. The fact that $\tilde{x}(Q)$ is nondecreasing in Q establishes that, despite the distortions that may occur in the mean of reviews because of the selection bias and the naïveté of consumers, the ordering of the sample mean of reviews (in the long run) will always be in line with the ordering of the true quality of the products. A second important implication of the characterization of $\tilde{x}(Q)$ is that there is a clear dichotomy with regard to social learning for naïve customers:

—If the quality of the product is such that all customers would purchase it with knowledge of Q, i.e., $Q + \underline{\theta} \geqslant \tau$, then $\tilde{x}(Q) = Q$ and social learning takes place in the long run when customers are naïve.

—If the lower bound of the support of Θ is such that $\underline{\theta} < \tau - Q$, then social learning does not take place and customers consistently overestimate the quality of the product in the long run. In this case, Theorem 2(c) shows that the long-run average customer loss is bounded away from zero.

In this sense, part (c) provides a benchmark to compare to with potentially more sophisticated inferential rules, or with a mixed population. We investigate this next.

4.2. Reverse Engineering the Average of Reviews

Now we ask the question of whether there exists a tractable way for customers to approximate $\mathbb{E}[Q \mid \bar{X}_n]$ when making inferences about the true quality Q, while ensuring that social learning takes place.

Suppose for a moment that customers are able to compute $\mathbb{E}[Q \mid \bar{X}_n]$. Notice that when $B(n+1) \geqslant 1$, one may rewrite $\bar{X}_n$ as

$$\bar{X}_n = Q + \frac{1}{B(n+1)} \sum_{i=1}^{B(n+1)} \Theta_{b_i} + \frac{1}{B(n+1)} \sum_{i=1}^{B(n+1)} (Q_{b_i} - Q) \quad (7)$$

and the distribution of Θ_{b_i} is the conditional distribution of $[\Theta \mid \Theta \geqslant \tau - \mathbb{E}[Q \mid \bar{X}_{b_i - 1}]]$.

We search for an approximation $\hat{q}_{n+1} \approx \mathbb{E}[Q \mid \bar{X}_n]$ where $\hat{q}_{n+1}$ is some function of $\bar{X}_n$. Suppose for a moment that $\hat{q}_n$ converges to an almost sure limit. Since the $\varepsilon_{b_i} = Q_{b_i} - Q$ are i.i.d., one would expect that, by the strong law of large numbers, as n grows large,

$$\frac{1}{B(n+1)} \sum_{i=1}^{B(n+1)} (Q_{b_i} - Q) \approx 0,$$

and in turn, given (7),

$$\bar{X}_n \approx \hat{q}_n + \mathbb{E}[\Theta \mid \Theta \geqslant \tau - \hat{q}_n]. \quad (8)$$

Equation (8) motivates the definition $\hat{q}_n$ as a solution y to the fixed point equation $\bar{X}_n = y + \mathbb{E}[\Theta \mid \Theta \geqslant \tau - y]$. We next formalize this idea. To do that, we will need the following technical assumption.

Assumption 2. *The random variable Θ has decreasing mean residual lifetime, i.e.,* $\mathrm{MRL}(x) := \mathbb{E}[\Theta \mid \Theta > x] - x$ *is decreasing on* $[\underline{\theta}, \bar{\theta}]$.

While this is a technical condition, it is satisfied by a large number of common distributions. A simple sufficient condition for it to hold is that the density function $f_\Theta(\cdot)$ be strictly log-concave (see, e.g., Marshall and Olkin 2007, Proposition B.8, C.1.b, and D.2). Therefore it holds for various distributions such as normal, Laplace, extreme value, or their truncations.

Customer-independent bias correcting approximation for $\mathbb{E}[Q \mid \bar{X}_n]$. As stated above, a simple potential approximation for $\mathbb{E}[Q \mid \bar{X}_n]$ can be obtained by solving a fixed point equation that is independent of the position of the customer and even the number of reviews seen so far. In particular, let

$$\psi(y) = y + \mathbb{E}[\Theta \mid \Theta \geqslant \tau - y].$$

Note that $\psi(y) = \mathrm{MRL}(\tau - y) + \tau$. Under Assumption 2, $\mathrm{MRL}(z)$ is decreasing. We deduce that $\psi(y)$ is increasing on $[\underline{q}, \bar{q}]$ with codomain $\mathcal{D} = [\psi(\underline{q}), \psi(\bar{q})]$. Let $\mathcal{P}(\cdot)$ denote the projection operator on $\mathcal{D}$ and for $x \in \mathbb{R}$, define $h(x)$ as follows:

$$h(x) := \psi^{-1}(\mathcal{P}(x)). \quad (9)$$

By convention, when no one has yet purchased, we assume that customers approximate Q by its mean. We now define formally the approximation of $\mathbb{E}[Q \mid \bar{X}_n]$ as follows. For $n \geqslant 0$,

$$\hat{q}_{n+1} = \begin{cases} \mathbb{E}[Q] & \text{if } B(n+1) = 0, \\ h(\bar{X}_n) & \text{if } B(n+1) \geqslant 1. \end{cases} \quad (10)$$

This is a behaviorally appealing rule as it is relatively simple and does not rely on the number of past reviews.

Social Learning. We next analyze the evolution of $\hat{q}_n$ when each customer n takes $\hat{q}_n$ as a proxy for $\mathbb{E}[Q \mid \bar{X}_n]$ when making a purchase decision.

Theorem 3 (Bias Correcting Rule)**.** *Suppose that Assumptions 1 and 2 hold. Suppose further that all customers use the approximation (10), so that the decision of customer $n+1$ is given by $Z_{n+1} = \mathbb{1}\{\hat{q}_{n+1} + \Theta_{n+1} \geqslant \tau\}$ for $n \geqslant 1$. Then:*

(a) *The sequence $\hat{q}_n$ converges to Q almost surely as $n \uparrow \infty$.*

(b) *There exists some constant $C > 0$ such that, for all $N \geqslant 1$,*

$$\mathcal{L}_Q(N) \leqslant C\sqrt{N} \text{ almost surely.}$$

Theorem 3 has many implications. A first implication is that there exist non-Bayesian inference rules that are only based on the mean of past reviews and lead to social learning, as long as all customers use the same rule. This is quite remarkable as it is not even clear how to compute $\mathbb{E}[Q \mid \bar{X}_n]$ in the first place. The proof of part (a) is based on a Robbins-Monro type stochastic approximation analysis, which enables us to establish almost sure convergence to the true quality.

A second implication is that these rules do not require the knowledge of the number of past reviews posted or the position of the customer in the population. A general question is, What is the minimal information required to ensure social learning? Here, we establish that the sample mean, in conjunction with the distribution of the Θ_n's, is sufficient to ensure social learning.

A third implication is that even under the very limited information that customers observe (sample mean of past reviews, which is not a sufficient statistic in this case), the true quality parameter is approached at a rate $1/\sqrt{N}$ (part (b)), which is the fastest rate one could obtain even if customers were able to observe all past reviews and all customers were altruistic. Indeed, in the latter case, one would be back in a classical statistics setup with i.i.d. observations, in which there is a fundamental limit for the quality of the estimates that one could obtain for Q, based on a given number of observations.

4.3. Mixed Population of Customers

In this section, we investigate social learning for a heterogeneous population in terms of their inference rule (which could be an indication of their computational capabilities). In particular, we analyze the interplay between naïve and more "sophisticated" customers. We assume that each customer can have one of two types: naïve or sophisticated. Customer n is naïve with probability β and sophisticated with probability $1-\beta$, independently of all else. We assume that β belongs to $(0,1]$ in what follows. Sophisticated customers, who are assumed to know β as well as the distribution of Θ, use an approximation $\check{q}_n \approx \mathbb{E}[Q \mid \bar{X}_n]$ where $\check{q}_n$ is some function of $\bar{X}_n$ and naïve customers use $\bar{X}_n$ as a proxy for $\mathbb{E}[Q \mid \bar{X}_n]$. Suppose for a moment that $\check{q}_n$ and $\bar{X}_n$ converge to an almost sure limit, then one would expect that, as n grows large,

$$\bar{X}_n \approx \check{q}_n + \beta\mathbb{E}[\Theta \mid \Theta \geqslant \tau - \bar{X}_n] + (1-\beta)\mathbb{E}[\Theta \mid \Theta \geqslant \tau - \check{q}_n].$$

Motivated by this and the previous section, we now define an inference and purchasing rule for sophisticated customers in this setting. For any $x \in \mathbb{R}$, we define

$$\psi_\beta(y) = y + (1-\beta)\mathbb{E}[\Theta \mid \Theta \geqslant \tau - y].$$

ψ_β is strictly increasing on $[\underline{q}, \bar{q}]$ with codomain $\mathscr{D}_\beta = [\psi_\beta(\underline{q}), \psi_\beta(\bar{q})]$. Furthermore, under Assumption 2, ψ'_β is well defined and is bounded below by β. Let $\mathscr{P}_\beta(\cdot)$ denote the projection operator on $\mathscr{D}_\beta$. For $x \in \mathbb{R}$, define $h_\beta(x)$ as follows:

$$h_\beta(x) := \begin{cases} \psi_\beta^{-1}(\mathscr{P}_\beta(x - \beta\mathbb{E}[\Theta \mid \Theta \geqslant \tau - x])), & \text{if } x \geqslant \underline{q} \\ \underline{q}, & \text{if } x < \underline{q}. \end{cases}$$

In particular, we formally define the approximation of $\mathbb{E}[Q \mid \bar{X}_n]$ that sophisticated customers use as follows. For $n \geqslant 0$,

$$\check{q}_{n+1} = \begin{cases} \mathbb{E}[Q] & \text{if } B(n+1) = 0, \\ h_\beta(\bar{X}_n) & \text{if } B(n+1) \geqslant 1. \end{cases} \tag{11}$$

Theorem 4. *Suppose that Assumptions 1 and 2 hold. Suppose further that all naïve customers use the update rule (5) and all sophisticated customers use the approximation (11) when making purchase decisions. Then, we have the following*:

(a) *The sequence $\check{q}_n$ converges almost surely to Q as $n \uparrow \infty$ and $\bar{X}_n$ converges almost surely to some $\tilde{x}_\beta(Q) \geqslant Q$, where $\tilde{x}_\beta(Q)$ is the unique solution to*

$$x = Q + \beta\mathbb{E}[\Theta \mid \Theta \geqslant \tau - x] + (1-\beta)\mathbb{E}[\Theta \mid \Theta \geqslant \tau - Q].$$

(b) *The limit $\tilde{x}_\beta(Q)$ is nondecreasing in Q and nonincreasing in β. Furthermore, the bias $\tilde{x}_\beta(Q) - Q$ is nonincreasing in Q.*

(c) *The welfare losses satisfy*

$$\lim_{N\uparrow\infty} \frac{1}{N}\mathscr{L}_Q(N) = \beta\int_{\tau-\tilde{x}_\beta(Q)}^{\tau-Q} [Q+z-\tau]\,\mathrm{d}F_\Theta(z) \text{ almost surely.}$$

Theorem 4(a) first establishes that convergence still occurs for each population, even when they interact. More notably, the result shows that sophisticated customers do not need to know who has written which reviews and can learn the true quality of the product even in an environment with customers that use different rules for inference and purchasing and where they only the sample mean of past reviews. The only information required is the distribution of the Θ_n's and the proportion of naïve customers. Without this information the quality Q would not be identifiable.

Theorem 4(b) is the parallel of Theorem 2(b) and shows that inferred quality by the subset of naïve customers has similar properties as in the case a homogeneous naïve population.

The integral $\int_{\tau-\tilde{x}_\beta(Q)}^{\tau-Q}[Q+z-\tau]\,\mathrm{d}F_\Theta(z)$ presented in Theorem 4(c) may be interpreted as the loss per naïve customer and can be seen to be nonincreasing in β, given part (b) of the theorem. Hence, an individual naïve customer, asymptotically, is negatively affected by the presence of sophisticated customers that precede him. This stems from the fact that naïve customers tend to overestimate the quality, and in turn induce a weaker selection bias in reviews, which limits errors for future naïve customers.

5. Extensions and Conclusions

Reports versus sincere evaluations. Various studies have documented that observed statistics may have an important impact on consumer ratings. For example, Talwar et al. (2007) show that a user's rating partly reflects the difference between true quality and prior expectation of quality as inferred from previous reviews through some empirical analysis. Moe and Trusov (2011) show, through an analysis of sales data across time, that the current average of reports has a significant effect on consumer rating and sales. This raises the question of whether past reviews not only impact purchasing decisions (as captured in the previous sections) but also impact what a customer ultimately reports. In turn, this leads to the question of whether social learning can take place if reviews are not truthful.

One possible way to account for such behavioral models of reporting is to introduce a *reported* rating variable Z_n that may depend on the true (sincere) rating X_n but also on the average of past reported ratings, to which the consumer has access. In particular, we may incorporate such a feature as follows:

$$Z_{b_1} = X_{b_1}, \quad Z_{b_k} = \varphi(X_{b_k}, \bar{Z}_{b_{k-1}}), \quad k \geqslant 2,$$

where the function $\varphi: \mathbb{R}^2 \to \mathbb{R}$ is nondecreasing in its first argument. Such a formulation would capture behavioral models that indicate that consumers tend to account, directly or indirectly, for the information available to them when they decide what to report. The case in which consumers simply report their sincere ratings, analyzed in the present paper, corresponds to $\varphi(x,z) = x$ for all x, z in $\mathbb{R}^2$.

One may again use the approach presented earlier to exhibit a bias correcting rule that rationalizes the observation of $\bar{Z}_n$, given the reporting mechanism and selection bias observed in the reported reviews. Investigating the distributional conditions and the types of functions $\varphi(\cdot,\cdot)$ under which such a rule yields social learning constitutes an interesting research direction.

Manipulation of reviews. The fact that customers rely more and more on reviews in their evaluation process of products and services has led to significant manipulation attempts through automated review generation algorithms. While there is research on the detection of fake reviews, another line of questions that emerges is what type of fake review generation would be more effective and what is the best way to counter such efforts.

Reviews of competing products and product characteristics. The present paper provides a crisp characterization of how consumer learning may take place through online reviews. In particular, even when customers are not able to process all past reviews and base their purchasing decision only on the aggregate statistic corresponding to the sample mean of past reviews, we establish that a simple correction rule for inference leads to social learning. Interesting avenues of future research include the understanding of the interplay of product reviews and purchasing decisions for all products in the consideration set of customers. Similarly, one would also like to better delineate what is possible to learn about products under richer information structures. In the present paper, the quality is modeled as a single parameter and social learning takes place, even under very limited feedback. Is such learning possible if customers are attempting to uncover various characteristics of the product based on reviews?

At a more general level, important questions include the delineation of the class of behavioral models that lead to social learning but also quantifying the information contained in reviews, given the fact that customers may report reviews based on earlier reviews, the fact that various entities post fake reviews, and the heterogeneity in the customers interested in a given product.

Acknowledgments

The authors are grateful to Johannes Hörner, Asu Ozdaglar, an associate editor, and two referees for their comments, which greatly helped improve the paper.

Appendix. Proofs

Preliminaries. In the proofs, many statements involving Q should be understood in the almost sure sense. We do not explicitly write the almost sure qualification at all junctions to avoid repetition. Furthermore, given the boundedness of Θ and to avoid ill defined objects, we set $\mathbb{E}[\phi(\Theta) \mid \Theta > \bar{\theta}] = \phi(\bar{\theta})$ for every integrable function ϕ.

Proof of Theorem 1. Define $\mathcal{F}_\infty^X = \sigma(\bigcup_{n\geqslant 1} \mathcal{F}_n^X)$ and $M_\infty = \mathbb{E}[Q \mid \mathcal{F}_\infty^X]$. Given (4), by Levy's convergence theorem, $M_n \to M_\infty$ almost surely.

Next, we want to establish that $M_\infty = Q$ almost surely. We first show that customers purchase infinitely often, irrespectively of the value of Q. We then show that there exists a $\mathcal{F}_\infty^X$-measurable mapping from the set of sequences of reviews into $[\underline{q}, \bar{q}]$ that is equal to Q almost surely. Then we apply a Doob-type result to prove that $M_\infty = Q$ almost surely.

Step 0. We first establish that customers purchase the product infinitely often.

As a first substep, we show that $\mathbb{P}\{Z_n = 1\} \geqslant \delta$ for all $n \geqslant 1$, where δ was defined in Assumption 1. Remember that $Z_n = 1$ iff $M_{n-1} + \Theta_n \geqslant \tau$. By definition of M_n we have $M_n \geqslant \underline{q}$. Hence

$$\mathbb{P}(Z_n = 1) = \mathbb{P}(\Theta_n \geqslant \tau - M_{n-1}) \geqslant \mathbb{P}(\Theta_n \geqslant \tau - \underline{q}) \geqslant \delta,$$

where the last inequality follows from Assumption 1.

One may then use a coupling argument (to a sequence of i.i.d. Bernoulli random variables with success probability δ) in conjunction with the converse of the Borel Cantelli lemma to conclude that $Z_n = 1$ infinitely often. In turn, $B(n+1) \uparrow \infty$ almost surely as $n \uparrow \infty$.

Step 1. For any $k \geqslant 1$, note that $M_{b_k} \geqslant \underline{q}$ and hence by Assumption 1, $\tau - M_{b_k} \leqslant \bar{\theta}$. Furthermore, define

$$X'_{b_k} = X_{b_k} - \mathbb{E}[\Theta \mid \Theta > \tau - M_{b_k}],$$
$$\bar{X}'_{b_k} = \frac{1}{k} \sum_{i=1}^{k} X'_{b_i}.$$

Note that for all $k \geqslant 1$,

$$\mathbb{E}[X'_{b_k} \mid Q] = Q, \tag{A.1}$$
$$\bar{X}'_{b_{k+1}} = \bar{X}'_{b_k} + \frac{1}{k+1}[X'_{b_{k+1}} - \bar{X}'_{b_k}]. \tag{A.2}$$

Equation (A.2) describes the dynamics of a stochastic approximation algorithm with step size sequence $\{\gamma_{k+1} = 1/(k+1): k \geqslant 1\}$. Furthermore define,

$$\begin{aligned} g_Q(x) &:= \mathbb{E}[X'_{b_{k+1}} \mid \bar{X}'_{b_k} = x, Q] - x \\ &= \mathbb{E}[Q + \Theta_{b_{k+1}} + \varepsilon_{b_{k+1}} - \mathbb{E}[\Theta \mid \Theta > \tau - M_{b_k}] \mid \bar{X}'_{b_k} = x, Q] - x \\ &= Q + \mathbb{E}[\Theta_{b_{k+1}} - \mathbb{E}[\Theta \mid \Theta > \tau - M_{b_k}] \mid \bar{X}'_{b_k} = x, Q] - x \\ &= Q + \mathbb{E}[\mathbb{E}[\Theta_{b_{k+1}} - \mathbb{E}[\Theta \mid \Theta > \tau - M_{b_k}] \mid M_{b_k}] \mid \bar{X}'_{b_k} = x, Q] - x \\ &= Q - x. \end{aligned}$$

Hence, conditionally on Q, the only candidate (almost sure) limit of $\bar{X}_{b_k}$ as $k \uparrow \infty$ is a solution of the equation

$g_Q(x) = 0$, which is Q. We next establish that almost sure convergence takes place.

Here, we do not have a basic stochastic approximation algorithm as, conditional on $\mathcal{F}^X_{b_k}$, the distribution of $X'_{b_{k+1}} - \bar{X}'_{b_k}$ depends not only on $\bar{X}'_{b_k}$ but also on M_{b_k}, so the so-called Robbins-Monro condition is not satisfied. However, we will adapt the proof of Theorem 1 in Benveniste et al. (1990, Section 5.1) to establish almost sure convergence. Following the latter, we define

$$T_k = \bar{X}'_{b_k} - Q, \quad \text{and} \quad \zeta_k = |T_k|^2.$$

Noting that $\zeta_{k+1} = \zeta_k + 2\gamma_{k+1}T_k(X'_{b_{k+1}} - \bar{X}'_{b_k}) + \gamma^2_{k+1}(X'_{b_{k+1}} - \bar{X}'_{b_k})^2$, we have

$$\begin{aligned}\mathbb{E}[\zeta_{k+1} \mid \mathcal{F}^X_k, Q] &= \zeta_k + 2\gamma_{k+1}T_k\mathbb{E}[X'_{b_{k+1}} - \bar{X}'_{b_k} \mid \mathcal{F}^X_k, Q] \\ &\quad + \gamma^2_{k+1}\mathbb{E}[(X'_{b_{k+1}} - \bar{X}'_{b_k})^2 \mid \mathcal{F}^X_k, Q] \\ &= \zeta_k + 2\gamma_{k+1}T_k(Q - \bar{X}'_{b_k}) \\ &\quad + \gamma^2_{k+1}\mathbb{E}[(X'_{b_{k+1}} - \bar{X}'_{b_k})^2 \mid \mathcal{F}^X_k, Q].\end{aligned}$$

Observe that

$$\begin{aligned}&\mathbb{E}[(X'_{b_{k+1}} - \bar{X}'_{b_k})^2 \mid \mathcal{F}^X_k, Q] \\ &\quad = \mathbb{E}[(Q + \Theta_{b_{k+1}} + \varepsilon_{b_{k+1}} - \mathbb{E}[\Theta \mid \Theta > \tau - M_{b_k}] - \bar{X}'_{b_k})^2 \mid \mathcal{F}^X_k, Q] \\ &\quad = (Q - \bar{X}'_{b_k})^2 + \mathbb{E}[(\Theta_{b_{k+1}} - \mathbb{E}[\Theta \mid \Theta > \tau - M_{b_k}])^2] + \mathbb{E}[\varepsilon^2_{b_{k+1}}].\end{aligned}$$

Since the supports of $Q, \Theta_n, \varepsilon_n$ are all bounded, we have

$$\mathbb{E}[(\Theta_{b_{k+1}} - \mathbb{E}[\Theta \mid \Theta > \tau - M_{b_k}])^2] \leqslant C(1 + (\tau - M_{b_k})^2) \leqslant C'$$

and we deduce that

$$\mathbb{E}[(X'_{b_{k+1}} - \bar{X}'_{b_k})^2 \mid \mathcal{F}^X_k, Q] \leqslant C''(1 + \zeta_k).$$

We are now in a position to apply Lemma 2 in Benveniste et al. (1990, Section 5.1) to conclude that conditional on Q

$$\zeta_k \to \zeta < \infty \ \text{a.s.} \quad \text{and} \quad \sum_{k\geqslant 1} \gamma_{k+1}\zeta_k < \infty \ \text{a.s.}$$

This implies that ζ_k converges to zero almost surely. In other words, we have established that

$$\bar{X}'_{b_k} \to Q \text{ almost surely as } k \uparrow \infty.$$

This, in turn, implies that

$$\bar{X}'_n \to Q \text{ almost surely.} \tag{A.3}$$

***Step* 2.** Next we apply a result from Le Cam and Yang (2000, Chapter 8) where they consider the following general model: There is a sequence $\{\mathcal{E}_n\}$ of experiments $\mathcal{E}_n = \{P_{\theta,n}\colon q \in \mathcal{Q}\}$ given by measures on σ-fields $\mathcal{A}_n$ such that we have the following:

(1) There is on $\mathcal{Q}$ a σ-field $\mathcal{B}$ such that all functions $q \mapsto P_{q,n}(A), A \in \mathcal{A}_n$ are $\mathcal{B}$-measurable.

(2) A fixed finite measure μ on $(\mathcal{Q}, \mathcal{B})$ has been chosen.

Let $\mathcal{A}_n$ be a filtration on $\mathcal{X}$ and call $\mathcal{A}_\infty = \sigma(\vee_n \mathcal{A}_n)$.

Assume that $(\mathcal{Q}, \mathcal{B})$ is a Borel subset of a complete separable metric space with its σ-field of Borel subsets. Disintegrate the measures $P_{q,n}(\mathrm{d}x)\mu(\mathrm{d}q)$ in the form $F_{x,n}(\mathrm{d}q)S'_n(\mathrm{d}x)$. Call the $F_{x,n}$ consistent at q if for every neighborhood V of q the posterior measure $F_{x,n}(V^c)$ tends to zero almost surely for $P_{q,\infty}$.

Proposition 1 (Doob 1949). *Assume that $P_{q,n}$ is the restriction to $\mathcal{A}_n$ of a measure $P_{q,\infty}$ defined on $\mathcal{A}_\infty$. Assume also that there is a measurable function f from $(\mathcal{X}, \mathcal{A}_\infty)$ to $(\mathcal{Q}, \mathcal{B})$ such that*

$$\iint |q - f(x)| P_{q,n}(\mathrm{d}x)\,\mu(\mathrm{d}q) = 0. \tag{A.4}$$

Then for μ-almost all $q \in \mathcal{Q}$, the posterior measures $F_{x,n}$ are consistent.

The above result is stated in the language of Le Cam and Yang (2000, Chapter 8, Proposition 3).

Coming back to the present setting, let $(\mathcal{Q}, \mathcal{B})$ be the parameter space, where $\mathcal{Q} = [\underline{q}, \bar{q}]$ and $\mathcal{B}$ is the Borel σ-field, let $\mathcal{X}$ be the space of review sequences $\{X_n\}_{n\in\mathbb{N}}$, let the σ-field $\mathcal{A}_n$ be $\mathcal{F}^b_n$, as defined in (2), and let $\mathcal{A}_\infty = \mathcal{F}^b_\infty$. For each value q a measure $P_{q,\infty}(\mathrm{d}x)$ can be defined on $\mathcal{A}_\infty$ since the conditions for the Kolmogorov extension theorem are satisfied. Then, given that $\mathcal{Q}$ is bounded, (A.3) proves the existence of a measurable function from $(\mathcal{X}, \mathcal{A}_\infty)$ to $(\mathcal{Q}, \mathcal{B})$ that satisfies (A.4). Hence Proposition 1 shows that $M_\infty = Q$ almost surely.

Proof of Theorem 2. First note that $\mathbb{E}[\Theta \mid \Theta > \tau - x]$ is nonincreasing in x. The equation $x = Q + \mathbb{E}[\Theta \mid \Theta > \tau - x]$ admits a unique solution since the left-hand side (LHS) is increasing and the right-hand side is nonincreasing and the functions are all continuous. Hence $\tilde{x}(Q)$ is well defined. (a) The proof is organized as follows. We first map the evolution of the subsequence $\bar{X}_{b_k}$ to that of a stochastic approximation algorithm; we then apply a general stochastic approximation result for Robbins-Monro schemes to establish almost sure convergence of $\bar{X}_n$; we conclude by establishing that the almost sure limit of $\tilde{q}_n$ is exactly $\tilde{x}(Q)$.

***Step* 0.** We establish that customers purchase the product infinitely often.

We first show that $\mathbb{P}\{Z_n = 1\} \geqslant \delta$ for all $n \geqslant 1$, where δ was defined in Assumption 1. To that end, we first establish by induction that, if k purchases took place, then $\bar{X}_{b_k} \geqslant (\tau + \underline{q} + \underline{\varepsilon})/2$ for all $k \geqslant 2$.

If two purchases took place, one must have $\Theta_{b_2} \geqslant \tau - X_1$ and as a result

$$\bar{X}_{b_2} = \tfrac{1}{2}(X_1 + X_2) \geqslant \tfrac{1}{2}(X_1 + Q + \tau - X_1 + \varepsilon_{b_2}) \geqslant \tfrac{1}{2}(\underline{q} + \tau + \underline{\varepsilon}),$$

so the result clearly holds for $k = 2$.

Suppose now that the result holds for some $k \geqslant 2$. Suppose that a $(k+1)$st purchase takes place. We have

$$\begin{aligned}\bar{X}_{b_{k+1}} &= \frac{1}{k+1}(k\bar{X}_{b_k} + X_{b_{k+1}}) \geqslant \frac{1}{k+1}(k\bar{X}_{b_k} + Q + \tau - \bar{X}_{b_k} + \varepsilon_{b_{k+1}}) \\ &\geqslant \frac{1}{k+1}((k-1)\bar{X}_{b_k} + \underline{q} + \tau + \underline{\varepsilon}).\end{aligned}$$

By the induction hypothesis, we obtain that

$$\bar{X}_{b_{k+1}} \geqslant \tfrac{1}{2}(\underline{q} + \tau + \underline{\varepsilon}).$$

Given the above, we have

$$\begin{aligned}\mathbb{P}\{Z_n = 1\} &= \mathbb{P}\{Z_n = 1 \mid B(n) = 0\}\mathbb{P}\{B(n) = 0\} \\ &\quad + \mathbb{P}\{Z_n = 1 \mid B(n) = 1\}\mathbb{P}\{B(n) = 1\} \\ &\quad + \sum_{k=2}^{n} \mathbb{P}\{Z_n = 1 \mid B(n) = k\}\mathbb{P}\{B(n) = k\} \\ &= \mathbb{P}\{\Theta_n \geqslant \tau - \mathbb{E}[Q]\}\mathbb{P}\{B(n) = 0\}\end{aligned}$$

$$
\begin{aligned}
&+\mathbb{P}\{\Theta_n > \tau - X_{b_1} \mid B(n)=1\}\mathbb{P}\{B(n)=1\}\\
&+\sum_{k=2}^{n}\mathbb{P}\{\Theta_n \geqslant \tau - \bar{X}_{b_k} \mid B(n)=k\}\mathbb{P}\{B(n)=k\}\\
\geqslant\ &\mathbb{P}\{\Theta_n \geqslant \tau - \bar{q}\}\mathbb{P}\{B(n)=0\} + \mathbb{P}\{\Theta_n > \tau\\
&\quad - (Q+\tau-\bar{q}+\varepsilon_1) \mid B(n)=1\}\mathbb{P}\{B(n)=1\}\\
&+\sum_{k=2}^{n}\mathbb{P}\{\Theta_n \geqslant (\tau - \underline{q} - \underline{\varepsilon}) \mid B(n)=k\}\mathbb{P}\{B(n)=k\}\\
\geqslant\ &\delta,
\end{aligned}
$$

where the last inequality follows from Assumption 1.

One may then use a coupling argument (to a sequence of i.i.d. Bernoulli random variables with success probability δ) in conjunction with the converse of the Borel Cantelli lemma to conclude that $Z_n = 1$ infinitely often. In turn, $B(n+1)\uparrow\infty$ almost surely as $n\uparrow\infty$. Note that $\bar{X}_n = \bar{X}_{b_{B(n+1)}}$, and hence to analyze the almost sure convergence of $\bar{X}_n$, it suffices to analyze $\bar{X}_{b_k}$ as k grows to ∞.

***Step* 1.** Given (3), one has that, for $k \geqslant 1$,

$$\bar{X}_{b_{k+1}} = \bar{X}_{b_k} + \frac{1}{k+1}[X_{b_{k+1}} - \bar{X}_{b_k}]. \tag{A.5}$$

Let

$$\tilde{g}_Q(x) := \mathbb{E}[X_{b_{k+1}} \mid \bar{X}_{b_k} = x, Q] - x.$$

Equation (A.5) describes the dynamics of a stochastic approximation algorithm with step size sequence $\{1/(k+1)\colon k \geqslant 1\}$. Conditionally on Q, the only candidate (almost sure) limit of $\bar{X}_{b_k}$ as $k\uparrow\infty$ is a solution of the equation $\tilde{g}_Q(x) = 0$. We next establish that convergence takes place and characterize the limit.

***Step* 2.** We first analyze $\tilde{g}_Q(x)$. First, we observe that $\tilde{g}_Q(x)$ is given by

$$\tilde{g}_Q(x) = Q + \mathbb{E}[\Theta \mid \Theta \geqslant \tau - x] - x.$$

Note that $\mathbb{E}[\Theta \mid \Theta \geqslant \tau - x]$ is nonincreasing in x and hence for any value of Q in its support, $\tilde{g}_Q(x)$ is continuous and strictly decreasing. Furthermore, $\tilde{g}_Q(Q) \geqslant 0$ and $\lim_{x\uparrow+\infty}\tilde{g}_Q(x) = -\infty$.

This implies that there exists a unique $\tilde{x}(Q)\in\mathbb{R}$ such that $\tilde{g}_Q(x) > 0$ for $x < \tilde{x}(Q)$ and $\tilde{g}_Q(x) < 0$ for $x > \tilde{x}(Q)$.

If $x \geqslant x'$, then

$$
\begin{aligned}
\tilde{g}_Q(x) - \tilde{g}_Q(x') &= \mathbb{E}[\Theta \mid \Theta \geqslant \tau - x] - x\\
&\quad - \mathbb{E}[\Theta \mid \Theta \geqslant \tau - x'] + x' \leqslant -(x - x').
\end{aligned}
$$

Similarly, if $x \leqslant x'$, then $\tilde{g}_Q(x) - \tilde{g}_Q(x') \geqslant -(x - x')$. We deduce that

$$(x - x^*)\tilde{g}_Q(x) = (x - x^*)(\tilde{g}_Q(x) - \tilde{g}_Q(x^*)) \leqslant -(x - x^*)^2.$$

The latter implies that for all $\eta > 0$,

$$\sup_{x\colon \eta\leqslant|x-x^*|\leqslant 1/\eta}(x - x^*)\tilde{g}_Q(x) \leqslant -\eta^2 < 0.$$

The so-called stability condition is satisfied.

Boundedness of ε_n, Θ_n, and Q implies

$$
\begin{aligned}
\sigma_Q^2(x) &:= \int_{\tau-x}^{\bar{\theta}}\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} |Q+z+w|^2\,\mathrm{d}F_{\Theta\mid\Theta\geqslant\tau-x}(z)\,\mathrm{d}F_{\varepsilon}(w)\\
&\leqslant |Q|^2 + \mathbb{E}[\varepsilon^2] + Q\mathbb{E}[\Theta \mid \Theta > \tau - x] + \mathbb{E}[\Theta^2 \mid \Theta > \tau - x]\\
&\leqslant C,
\end{aligned} \tag{A.6}
$$

where C is an appropriate positive constant.

The inequality (A.6) implies that the conditions of Theorem 1 of Benveniste et al. (1990, Section 5.1) are satisfied, which implies that

$$\bar{X}_{b_k} \to \tilde{x}(Q) \text{ almost surely as } k\uparrow\infty.$$

This, in turn, implies that $\bar{X}_n$ converges to $\tilde{x}(Q)$ almost surely.

(b) The proof of part (b) is a direct corollary of the proof of Theorem 4(b) with $\beta = 1$.

(c) We have that

$$
\begin{aligned}
\mathscr{L}_Q(N) &= \sum_{i=1}^{N}\mathbb{E}[(Q+\Theta_i-\tau)\mathbb{1}\{\Theta_i > \tau - Q\} \mid Q]\\
&\quad - \mathbb{E}[(Q+\Theta_i-\tau)Z_i \mid Q]\\
&= \sum_{n=1}^{N}\mathbb{E}[|Q+\Theta_i-\tau|\mathbb{1}\{\tau - \max\{Q,\tilde{q}_n\} \leqslant \Theta_n\\
&\qquad \leqslant \tau - \min\{Q,\tilde{q}_n\}\} \mid Q]\\
&= \sum_{n=1}^{N}\mathbb{E}\left[\int_{\tau-\max\{Q,\tilde{q}_n\}}^{\tau-\min\{Q,\tilde{q}_n\}} |Q+z-\tau|\,\mathrm{d}F_\Theta(z)\,\middle|\, Q\right].
\end{aligned}
$$

Given the result of part (a), the fact that $\tilde{x}(Q) \geqslant Q$ and noting that Θ does not have any atom, one may use the continuous mapping theorem to conclude that

$$\int_{\tau-\max\{Q,\tilde{q}_n\}}^{\tau-\min\{Q,\tilde{q}_n\}} |Q+z-\tau|\,\mathrm{d}F_\Theta(z) \to \int_{\tau-\tilde{x}(Q)}^{\tau-Q}[Q+z-\tau]\,\mathrm{d}F_\Theta(z)$$
$$\text{almost surely as } n\uparrow\infty.$$

Noting that

$$\int_{\tau-\max\{Q,\tilde{q}_n\}}^{\tau-\min\{Q,\tilde{q}_n\}} |Q+z-\tau|\,\mathrm{d}F_\Theta(z) \leqslant |Q-\tau| + \mathbb{E}[|\Theta|] < \infty$$
$$\text{almost surely,}$$

we deduce by the dominated convergence theorem that

$$
\begin{aligned}
&\mathbb{E}\left[\int_{\tau-\max\{Q,\tilde{q}_n\}}^{\tau-\min\{Q,\tilde{q}_n\}} |Q+z-\tau|\,\mathrm{d}F_\Theta(z) \mid Q\right]\\
&\quad\to \int_{\tau-\tilde{x}(Q)}^{\tau-Q}[Q+z-\tau]\,\mathrm{d}F_\Theta(z) \quad \text{almost surely as } n\uparrow\infty.
\end{aligned}
$$

The result then follows through a classical Cesàro sum argument.

Proof of Theorem 3. (a) The proof follows a similar structure as that of Theorem 2. Below we highlight and expand only the parts that differ.

We first prove a preliminary lemma that pertains to $h(\cdot)$.

Lemma 1. *Under Assumptions 1 and 2, the function h is nondecreasing and Lipschitz continuous on $\mathbb{R}$.*

Proof of Lemma 1. Under Assumptions 1 and 2, the function $\psi(y)$ is continuously differentiable and (strictly) increasing on $[\underline{q}, \bar{q}]$.

Let $\rho = \min_{q\in[\underline{q},\bar{q}]}\psi'(q)$. By the above, $\rho > 0$ and, for all $x, x' \in \mathbb{R}$,

$$
\begin{aligned}
|h(x) - h(x')| &= |\psi^{-1}(\mathscr{P}(x)) - \psi^{-1}(\mathscr{P}(x'))|\\
&\leqslant \rho^{-1}|\mathscr{P}(x) - \mathscr{P}(x')| \leqslant \rho^{-1}|x - x'|.
\end{aligned}
$$

The Lipschitz continuity follows and the monotonicity of $h(\cdot)$ is direct.

***Step* 0.** As in Theorem 2, one may establish that customers purchase the product infinitely often.

Note that for any $n \geqslant 0$, customer $n+1$ purchases with probability $\mathbb{P}\{\hat{q}_{n+1}+\Theta_{n+1}\geqslant\tau\}\geqslant\mathbb{P}\{q+\Theta_{n+1}\geqslant\tau\}\geqslant\delta>0$. As a result, it is possible to establish that $Z_n=1$ infinitely often and $B(n+1)\uparrow\infty$ almost surely as $n\uparrow\infty$. Note that $\bar{X}_n=\bar{X}_{b_{B(n+1)}}$, and hence to analyze the almost sure convergence of $\bar{X}_n$, it suffices to analyze $\bar{X}_{b_k}$ as k grows to ∞.

***Step* 1.** The evolution of $\bar{X}_{b_k}$ may be written as in (A.5), describing the dynamics of a stochastic approximation algorithm with step size $1/k$. Let

$$g_Q(x) := \mathbb{E}[X_{b_{k+1}} \mid \bar{X}_{b_k}=x, Q]-x.$$

Note that this function now differs from $\tilde{g}_Q(\cdot)$ defined in Theorem 2. Conditionally on Q, the only candidate (almost sure) limit of $\bar{X}_{b_k}$ as $k\uparrow\infty$ is a solution of the equation $g_Q(x)=0$. We next establish that convergence takes place and characterize the limit.

***Step* 2.** We establish properties of $g_Q(x)$. First, we observe that $g_Q(x)$ is given by

$$g_Q(x)=Q+\mathbb{E}[\Theta\mid\Theta\geqslant\tau-h(x)]-x.$$

Noting that $h(x)$ is continuous and nondecreasing, we have that $\mathbb{E}[\Theta\mid\Theta\geqslant\tau-h(x)]$ is nonincreasing in x. We deduce that, for any value of Q in $[\underline{q},\bar{q}]$, $g_Q(x)$ is continuous and strictly decreasing. Furthermore, $g_Q(\underline{Q})\geqslant 0$ and $\lim_{x\uparrow+\infty} g_Q(x)=-\infty$. This implies that there exists a unique $x^*\in\mathbb{R}$ such that $g_Q(x)>0$ for $x<x^*$ and $g_Q(x)<0$ for $x>x^*$.

Given that $h(\cdot)$ is nondecreasing, one may establish through a similar reasoning as the one used for $\tilde{g}_Q(\cdot)$ in the proof of Theorem 2 that for all $\eta>0$,

$$\sup_{x:\,\eta\leqslant|x-x^*|\leqslant 1/\eta}(x-x^*)g_Q(x)\leqslant-\eta^2<0,$$

and hence the stability condition is satisfied.

On another hand, for some $C>0$,

$$\sigma_Q^2(x):=\int_{\tau-h(x)}^{\bar{\theta}}\int_{\underline{\varepsilon}}^{\bar{\varepsilon}}|Q+z+w|^2\,\mathrm{d}F_{\Theta\mid\Theta\geqslant\tau-h(x)}(z)\,\mathrm{d}F_{\varepsilon}(w)\leqslant C,$$

for some appropriate $C>0$. We deduce that the conditions of Theorem 1 of Benveniste et al. (1990, Section 5.1) are satisfied, and hence

$$\bar{X}_{b_k}\to x^* \ \text{ almost surely as } k\uparrow\infty.$$

As noted above, this directly implies that $\bar{X}_n$ converges to x^*. We are left to characterize x^* and $h(x^*)$.

***Step* 3.** Notice that $g_Q(x^*)=0$ is equivalent to $Q+\mathbb{E}[\Theta\mid\Theta\geqslant\tau-h(x^*)]-x^*=0$. By definition of $h(\cdot)$ (see (9)), one has that $h(x^*)=Q$. The continuity of $h(\cdot)$ implies that $\hat{q}_n=h(\bar{X}_n)$ converges almost surely to Q. This completes the proof of part (a).

(b) The proof of this part is organized as follows. We first bound system losses as a function of how closely $\hat{q}_n$ approaches Q. Then, we characterize the speed of convergence of $\bar{X}_n$ to x^*, which in turn provides a bound on the speed of convergence of $\hat{q}_n$ to Q. The latter bound yields in turn a bound on the losses.

***Step* 1.** We have that

$$\mathscr{L}_Q(N)=\sum_{n=1}^{N}\mathbb{E}\left[\int_{\tau-\max\{Q,\hat{q}_n\}}^{\tau-\min\{Q,\hat{q}_n\}}|Q+z-\tau|\,\mathrm{d}F_{\Theta}(z)\,\middle|\,Q\right]$$
$$\leqslant\sum_{n=1}^{N}\mathbb{E}[|\hat{q}_n-Q|\mid Q]\leqslant\sum_{n=1}^{N}\sqrt{\mathbb{E}[|\hat{q}_n-Q|^2\mid Q]}.$$

We continue to denote by x^* the almost sure limit of $\bar{X}_{b_k}$. As seen in the proof of part (a), one has that $h(x^*)=Q$. We deduce from Lemma 1 that

$$\mathbb{E}[|\hat{q}_n-Q|^2\mid Q]=\mathbb{E}[|h(\bar{X}_n)-h(x^*)|^2\mid Q]\leqslant\rho^{-1}\mathbb{E}[|\bar{X}_n-x^*|^2\mid Q].$$

***Step* 2.** Bounding the mean squared error $\mathbb{E}[|\bar{X}_n-x^*|^2\mid Q]$.

As seen in the proof of part (a), if b_k denotes the kth customer who purchases, the evolution of $\bar{X}_{b_k}$, conditional on Q, may be seen as that of a Robbins-Monro algorithm.

Continuing with the notation introduced in part (a) and as noticed there, we have that $(x-x^*)g_Q(x)\leqslant-(x-x^*)^2$.

Noting that the Robbins-Monro algorithm has step size $1/k$, Theorem 22 in Benveniste et al. (1990, Section II 1.10.1) implies that for some $\lambda>0$

$$\mathbb{E}[|\bar{X}_{b_k}-x^*|^2\mid Q]\leqslant\lambda k^{-1}.$$

Next, we need to connect $\bar{X}_n$ to the sequence $\bar{X}_{b_k}$. Noting that $\bar{X}_n=\bar{X}_{b_{B(n+1)}}$, one has that

$$\mathbb{E}[|\bar{X}_n-x^*|^2\mid Q,B(n+1)=k]=\mathbb{E}[|\bar{X}_{b_k}-x^*|^2\mid Q]\leqslant\lambda k^{-1}.$$

Noting that $\bar{X}_{b_0}=\mathbb{E}[Q]$ and letting $C=|\mathbb{E}[Q]-x^*|$, one has, by conditioning on values of $B(n+1)$,

$$\mathbb{E}[|\bar{X}_n-x^*|^2\mid Q]$$
$$=\sum_{k=0}^{n}\mathbb{E}[|\bar{X}_n-x^*|^2\mid Q,B(n+1)=k]\mathbb{P}\{B(n+1)=k\mid Q\}$$
$$\leqslant\lambda\sum_{k=1}^{n}\frac{1}{k}\mathbb{P}\{B(n+1)=k\mid Q\}+C\mathbb{P}\{B(n+1)=0\mid Q\}.$$

Hence for any $n'\in[1,n]$, one has that

$$\mathbb{E}[|\bar{X}_n-x^*|^2\,|\,Q]\leqslant\max\{\lambda,C\}\mathbb{P}\{B(n+1)\leqslant n'\mid Q\}+\frac{\lambda}{n'+1}.$$

Note that

$$\mathbb{P}\{B(n+1)\leqslant n'\mid Q\}\leqslant\mathbb{P}\left\{\sum_{i=1}^{n}Z_i\leqslant n'\,\middle|\,Q\right\}.$$

Let $p=\mathbb{P}\{\underline{q}+\Theta\geqslant\tau\}$, which is bounded away from zero by Assumption 1. Let $I_1,I_2,\ldots$ denote a sequence of i.i.d. random variables having a Bernoulli distribution with parameter p. Then, by a coupling argument, one has that

$$\mathbb{P}\{B(n+1)\leqslant n'\mid Q\}=\mathbb{P}\left\{\sum_{i=1}^{n}I_i\leqslant n'\,\middle|\,Q\right\}.$$

Now, Hoeffding's inequality implies that for any $\Delta\in[0,p)$,

$$\mathbb{P}\left\{n^{-1}\sum_{i=1}^{n}I_i\leqslant(p-\Delta)\right\}\leqslant e^{-2n\Delta^2}.$$

Fix $\Delta \in [0, p)$ and let $n' = \lfloor n(p-\Delta) \rfloor$. Then one obtains that

$$\begin{aligned}\mathbb{E}[|\bar{X}_n - x^*|^2 \mid Q] &\leqslant \max\{\lambda, C\} e^{-2n\Delta^2} + \frac{\lambda}{\lfloor n(p-\Delta)\rfloor + 1}\\ &\leqslant \max\{\lambda, C\} e^{-2n\Delta^2} + \frac{\lambda}{n(p-\Delta)}\\ &\leqslant \frac{\lambda'}{n(p-\Delta)},\end{aligned}$$

where in the last inequality, λ' is an appropriately selected constant. Returning to the losses, one has that

$$\mathscr{L}_Q(N) \leqslant \sum_{n=1}^{N} \sqrt{\mathbb{E}[|\bar{X}_n - Q|^2 \mid Q = q]} \leqslant \frac{\sqrt{\lambda'}}{\sqrt{p-\Delta}} \sum_{n=1}^{N} \frac{1}{\sqrt{n}}.$$

We conclude that there exists some $C > 0$ such that

$$\mathscr{L}_Q(N) \leqslant C N^{1/2} \text{ almost surely.}$$

This concludes the proof of part (b).

Proof of Theorem 4. The proofs of parts (a) and (c) follow along similar lines as those of Theorems 2 and 3 and are omitted.

Here, we focus on part (b). The fact that the equation admits a unique solution follows directly since the LHS is strictly increasing from $(-\infty, \infty)$ into $(-\infty, \infty)$.

The fact that $\tilde{x}_\beta(Q) \geqslant Q$ is direct from the equation that $\tilde{x}_\beta(Q)$ solves and the fact that the conditional expectations are nonnegative.

We first analyze the dependence on Q. The fact that $\tilde{x}_\beta(\lambda)$ is nondecreasing in λ follows from an application of the implicit function theorem. For the bias, note that

$$\tilde{x}_\beta(\lambda) - \lambda = \beta \mathbb{E}[\Theta \mid \Theta > \tau - \tilde{x}_\beta(\lambda)] + (1-\beta)\mathbb{E}[\Theta \mid \Theta > \tau - \lambda].$$

Note that $\mathbb{E}[\Theta \mid \Theta > \tau - \lambda]$ is nondecreasing in λ. On another hand, $\tilde{x}_\beta(\lambda)$ is nondecreasing in λ and hence $\mathbb{E}[\Theta \mid \Theta > \tau - \tilde{x}_\beta(\lambda)]$ is nonincreasing in λ. We deduce that $\tilde{x}_\beta(\lambda) - \lambda$ is nonincreasing in λ.

To establish monotonicity with respect to β, fix $\beta' > \beta$, both in $[0, 1]$. Let

$$\phi_\beta(x) = x - \beta \mathbb{E}[\Theta \mid \Theta > \tau - x].$$

Then

$$\begin{aligned}\phi_\beta(\tilde{x}_{\beta'}(Q)) &= \phi_{\beta'}(\tilde{x}_{\beta'}(Q)) + (\beta' - \beta)\mathbb{E}[\Theta \mid \Theta > \tau - \tilde{x}_{\beta'}(Q)]\\ &= Q + (1-\beta')\mathbb{E}[\Theta \mid \Theta > \tau - Q]\\ &\quad + (\beta' - \beta)\mathbb{E}[\Theta \mid \Theta > \tau - \tilde{x}_{\beta'}(Q)]\\ &= \phi_\beta(\tilde{x}_\beta(Q)) + (\beta' - \beta)[\mathbb{E}[\Theta \mid \Theta > \tau - \tilde{x}_{\beta'}(Q)]\\ &\quad - \mathbb{E}[\Theta \mid \Theta > \tau - Q]]\\ &\leqslant \phi_\beta(\tilde{x}_\beta(Q)),\end{aligned}$$

where the last inequality follows from $\tilde{x}_\beta(Q) \geqslant Q$ and hence

$$\mathbb{E}[\Theta \mid \Theta > \tau - \tilde{x}_{\beta'}(Q)] - \mathbb{E}[\Theta \mid \Theta > \tau - Q] \geqslant 0.$$

The fact that $\phi_\beta(\cdot)$ is increasing implies that $\tilde{x}_\beta(Q) \leqslant \tilde{x}_{\beta'}(Q)$.

References

Acemoglu D, Ozdaglar A (2011) Opinion dynamics and learning in social networks. *Dynam. Games Appl.* 1(1):3–49.

Acemoglu D, Bimpikis K, Ozdaglar A (2014) Dynamics of information exchange in endogenous social networks. *Theor. Econ.* 9(1):41–97.

Acemoglu D, Dahleh MA, Lobel I, Ozdaglar A (2011) Bayesian learning in social networks. *Rev. Econom. Stud.* 78(4):1201–1236.

Arieli I, Mueller-Frank M (2017) A general analysis of sequential learning. IESE Business School Working Paper No. WP-1119-E. SSRN:https://ssrn.com/abstract=2466903.

Banerjee A, Fudenberg D (2004) Word-of-mouth learning. *Games Econom. Behav.* 46(1):1–22.

Banerjee AV (1992) A simple model of herd behavior. *Quart. J. Econom.* 107(3):797–817.

Benveniste A, Métivier M, Priouret P (1990) *Adaptive Algorithms and Stochastic Approximations* (Springer, Berlin).

Bikhchandani S, Hirshleifer D, Welch I (1992) A theory of fads, fashion, custom, and cultural change in informational cascades. *J. Polit. Econ.* 100(5):992–1026.

Bikhchandani S, Hirshleifer D, Welch I (1998) Learning from the behavior of others: Conformity, fads, and informational cascades. *J. Econom. Perspect.* 12(3):151–170.

Çelen B, Kariv S (2004) Observational learning under imperfect information. *Games Econom. Behav.* 47(1):72–86.

Crapis D, Ifrach B, Maglaras C, Scarsini M (2017) Monopoly pricing in the presence of social learning. *Management Sci.* 63(11): 3586–3608.

Doob JL (1949) Application of the theory of martingales. *Le Calcul des Probabilités et ses Applications*, Colloques Internationaux du Centre National de la Recherche Scientifique, Vol. 13 (Centre National de la Recherche Scientifique, Paris), 23–27.

Epstein LG, Noor J, Sandroni A (2010) Non-Bayesian learning. *B. E. J. Theor. Econ.* 10(1):Article 3.

Goeree JK, Palfrey TR, Rogers BW (2006) Social learning with private and common values. *Econom. Theory* 28(2):245–264.

Herrera H, Hörner J (2013) Biased social learning. *Games Econom. Behav.* 80:131–146.

Ifrach B, Maglaras C, Scarsini M (2013) Bayesian social learning from consumer reviews. Technical Report 2293158. SSRN, http://dx.doi.org/10.2139/ssrn.2293158.

Jadbabaie A, Molavi P, Sandroni A, Tahbaz-Salehi A (2012) Non-Bayesian social learning. *Games Econom. Behav.* 76(1):210–225.

Lafky J (2014) Why do people rate? Theory and evidence on online ratings. *Games Econom. Behav.* 87:554–570.

Le Cam L, Yang GL (2000) *Asymptotics in Statistics*, Second ed. (Springer, New York).

Li X, Hitt LM (2008) Self-selection and information role of online product reviews. *Inf. Syst. Res.* 19(4):456–474.

Luca M (2011) Reviews, reputation, and revenue: The case of yelp.com. Technical Report 12–016, Harvard Business School, Boston.

Marshall AW, Olkin I (2007) *Life Distributions* (Springer, New York).

Moe WW, Trusov M (2011) Measuring the value of social dynamics in online product ratings forums. *J. Marketing Res.* 48(3):444–456.

Papanastasiou Y, Bakshi N, Savva N (2014) Scarcity strategies under quasi-Bayesian social learning. Working paper, London Business School.

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

Talwar A, Jurca R, Faltings B (2007) Understanding user behavior in online feedback reporting. MacKie-Mason J, Parkes DC, Resnick P, eds. *Proc. 8th ACM Conf. Electronic Commerce, EC '07* (ACM, New York), 134–142.

Zhang Y, Lappas T, Crovella M, Kolaczyk ED (2014) Online ratings: Convergence towards a positive perspective? *Proc. IEEE Internat. Conf. Acoustics, Speech and Signal Processing, ICASSP* (IEEE, Piscataway, NJ), 4788–4792.

Omar Besbes is an associate professor at the Graduate School of Business, Columbia University. His research focuses on data-driven decision making and its applications in revenue management, operations, and service systems.

Marco Scarsini is a professor in the Department of Economics and Finance at LUISS, Rome, Italy. His research interests include applied probability and game theory, with a current focus on congestion games and social learning.