Submitted to *Management Science*
manuscript

Dynamic Pricing with Online Reviews

(Authors' names blinded for peer review)

This paper investigates how the presence of product reviews can help inform a dynamic-pricing monopolist. A salient feature of our problem is that the customers' willingness-to-pay, and hence the demand function, evolves over time in conjunction with the dynamics of the average rating generated by online reviews. The monopolist strives to maximize its total expected revenue over a finite horizon by adjusting prices in response to the evolving review dynamics. To formulate the problem in tractable form, we study a fluid model which serves as a good approximation of the system dynamics when the volume of sales is large. This formulation lends itself to key structural insights, which are leveraged to design a well-performing dynamic pricing policy for the underlying revenue maximization problem. The proposed policy admits a closed-form expression for price and its performance is near-optimal in a suitable asymptotic sense. We illustrate the effectiveness of the proposed policy via simulation and counterfactual analysis in an online marketplace.

Key words:

1. Introduction

1.1. Overview of the Problem

Product review platforms have emerged as viable mechanisms for sharing opinions and experiences on a wide range of products (or services) in online marketplaces. In these markets, the quality of a newly launched product is initially uncertain, but the reviews generated by buyers help inform subsequent customers. This form of social learning reduces the uncertainty pertaining to product quality, and hence supports better-informed purchase decisions. Clearly this learning dynamic, and said review process, affects the demand trajectory over time. The main question studied in this paper is how to optimally set prices within the context of these dynamics.

To address this question, we formulate a stylized model of a marketplace where a monopolist sells a single product to sequentially arriving customers. Customers are a priori not well informed about the quality levels of the product, but they observe product reviews from previous buyers and use a crude form of Bayesian learning to inform purchase decisions. (Throughout the paper, reviews mean numeric ratings and we use the two terms interchangeably.) Buyers may report their ex-post quality perception after purchase, which in turn affects decisions of subsequent customers. This model gives rise to a probability of purchase by the individual customer, which in aggregate form is often referred to as the market *demand function*. As the review process progresses, the demand function evolves dynamically (and stochastically) over time. The seller then uses the aforementioned demand function to seek a pricing policy that maximizes revenue over a finite horizon.

The revenue maximization problem described above is formulated as a stochastic dynamic programming problem, whose solution is difficult to characterize. In particular, in the setting we study demand is affected by the quality of a product; this information is a priori unknown but is revealed by reviews from previous buyers. The review process, in turn, is affected by changes in price over time, which introduces the following intertemporal effect: prices affect the dynamics of the review process, which in turn influence the demand and pricing decisions for subsequent customers. This feedback mechanism is one of the key challenges in solving the dynamic optimization problem and identifying an optimal pricing policy.

1.2. Summary of the Main Results

The paper contributes along three dimensions discussed below.

Demand modeling incorporating review information. We introduce a parsimonious stylized model for customer-level purchase probability that incorporates the "strength" of signal from product reviews at each time instant. The model is sufficiently tractable to enable construction of practically implementable dynamic pricing policies, which is one of the main focal points of the paper. We also present a small-scale illustrative empirical study to validate that the model passes basic "sanity checks," and highlights the practical significance of the proposed pricing policies derived from the model.

Fluid approximation and structural insights. As indicated earlier, the inherent feedback generated by the presence of reviews significantly complicates the study of revenue maximizing policies. To that end, we formulate a deterministic (and continuous) fluid approximation of the underlying revenue maximization problem. The dynamics of reviews are described by ordinary differential equations in this so-called fluid model, and the solution can be expressed in closed form, which enables us to investigate structural properties of optimal pricing strategies in a rigorous mathematical manner. In particular, this allows to characterize the parameter regimes where dynamic pricing yields more significant benefits vis-à-vis fixed-price strategies. As a result, our analysis provides some rough guidelines for practitioners regarding the choice of pricing strategies and in particular assessing whether the benefits of (more complex) dynamic pricing are merited.

Near-optimal pricing strategies. We propose and study a class of pricing policies referred to as Dynamic Fluid-Matching (DFM). We provide a theoretical bound on the performance of the

DFM policy and show via simulation that it yields performance that is close to optimal. Further, we test the DFM policy relative to a model calibrated with actual data in order to illustrate its effectiveness in settings inspired by actual market operating characteristics. One key observation that arises from our analysis is that DFM is highly effective during the "transient" phase where reviews are "most informative." In contrast, when this learning phase stabilizes, the DFM policy presents only modest improvement over fixed price counterparts. Our analysis provides some characterization of the boundary between the two phases within the selling horizon.

The remainder of the paper. The next section concludes the introduction with a review of related work. In Section 2 we develop a demand model that incorporates directly the review platform information and formulate a revenue maximization problem. In Section 3 we analyze a fluid version of the aforementioned revenue maximization problem and discuss structural properties of its optimal solution. In Section 4 we propose pricing policies, whose effectiveness is tested via simulation and counterfactual analysis in Section 5. In Section 6 we examine several extensions to the basic problem: first, we generalize our results to the case where customers' review propensity is time-dependent; and second, the single product results are generalized to a multiproduct setting. Appendix A contains additional numerical results and proofs are collected in Appendix B.

1.3. Related Literature

This paper contributes to various streams of research.

Product reviews and social learning. Product review platforms, or reputation systems in a broad sense, have gained growing attention as an important driver of product sales; see Dellarocas (2003) for a comprehensive overview of reputation systems. Several papers have empirically investigated the effect of product reviews on customers' purchasing behavior. Representatively, Chevalier and Mayzlin (2006) demonstrate that the difference in sales of books between Amazon and Barnes & Nobles is positively correlated with the difference in reviews. Duan et al. (2008) examine movies' daily box office performance to propose the importance of the number of reviews on sales. Li and Hitt (2008) analyze book sales data in Amazon and report that early reviews demonstrate positive bias due to a self-selection effect. Although the last result suggests the possibility of information distortion, a common conclusion from these works is that product reviews have significant effect on customers' purchasing behavior. In a sense, this paper yields a conclusion similar to the preceding papers: our stylized customer purchase model yields a significantly higher predictive power than benchmark models that ignore product reviews.

Product reviews constitute a form of social learning, which has been an active research area in economics. Seminal papers by Banerjee (1992) and Bikhchandani et al. (1992) consider agents who observe a private signal, as well as the entire history of decisions by previous agents, and update their belief in a Bayesian manner. They show that this social learning process may lead to herding behavior on bad decisions, despite the fact that agents are fully rational. Different papers have considered similar issues in the context of product reviews. Studying a market where consumers observe either the full history or some summary statistics of past reviews, Acemoglu et al. (2017) prove that asymptotic learning may fail. Besbes and Scarsini (2018) show that if buyers report perceived quality adjusted by their private signals and only observe the sample mean of past reviews, customers tend to overestimate the underlying quality of a product in the long run. In contrast, this paper assumes that customers are altruistic and thus report unbiased perceived quality for the benefit of subsequent customers, which rules out the possibility of bias in learning. Moreover, while price is not explicitly modeled in the preceding papers, this paper highlights the role of price in the social learning process.

Dynamic pricing with product reviews. Recently, several papers have studied pricing in online markets with uncertain quality, and various approaches to the modeling of product review platforms as information aggregators. Crapis et al. (2016) study conditions under which the true quality of a product is eventually revealed via reviews and propose a two-stage pricing policy that accounts for social learning. They focus on the setting where customers report binary reviews (like/dislike) based on their ex-post utility, which, in turn, depends on the price of the product. In our paper customers report directly the experienced level of quality, which is the most useful information to subsequent customers. The latter is a common assumption in Yu et al. (2015) and Papanastasiou and Savva (2017).

Specifically, Yu et al. (2015) and Papanastasiou and Savva (2017) study two-stage pricing strategies in the presence of strategic customers and product reviews that are modeled as independent and identically distributed random variables. Interestingly, Yu et al. (2015) show that both the firm and customers may be strictly worse off due to product reviews. Papanastasiou and Savva (2017) show that dynamic pricing may alleviates the strategic behavior of customers, and hence, can be preferred to preannounced pricing. This stands in contrast to other findings such as Aviv and Pazgal (2008).

He and Chen (2018) consider the review dynamics as a continuous-time stochastic process and study how information externalities distort optimal pricing strategies. Yang and Zhang (2018) study the joint pricing and inventory management problem in the presence of product reviews and show that the structural properties of optimal pricing strategies are not affected by inventory dynamics under base-stock policies. These studies focus almost exclusively on structural insights. Our paper complements these by explicitly characterizing a near-optimal pricing strategy in closed form, which supports further structural insights and allows the design of pricing strategies that are implementable.

In a more general setting of social learning, customers are subject to various forms of externalities. Bose et al. (2006, 2008) and Ma et al. (2018) study pricing when customers make decisions based on the purchase history of previous customers, and highlight the role of dynamic pricing in balancing immediate and future revenues. Candogan et al. (2012), Bloch and Quérou (2013), Ajorlou et al. (2016) and Makhdoumi et al. (2018) consider optimal pricing strategies in a social network where each customer's action depends directly on the action of their neighbors in the network graph. They characterize the optimal price as a function of the location in the network.

Fluid formulations. Deterministic and continuous (fluid) relaxations have become commonplace in studying various dynamic pricing problems; see Gallego and van Ryzin (1994, 1997) for the inception point of this literature. A prevalent way to implement fluid solutions is via "re-solving" heuristics that reevaluate the fluid policy as a function of the current state and time-to-go. In our paper we propose a policy that is predicated on re-solving a fluid counterpart of the underlying problem, and prove that this type of pricing policy is asymptotically optimal. We also show via numerical examples that it outperforms other benchmark pricing strategies.

2. Problem Formulation

2.1. The Model

Overview of the model. A monopolist sells a single product over a horizon of length T. It will be convenient to think of this planning horizon in terms of customer arrivals. To wit, customers arrive sequentially and are indexed by $t \in \{1,\dots,T\}$, and it will be assumed henceforth that T is known to the seller. Each customer purchases at most one product upon arrival and does not return to the market. The monopolist can influence demand by varying the price it offers to different customers. The demand also depends on the product's quality, which is not known to customers upon their arrival. Reviews reported by previous buyers provide public information about the product, and customers use this information to evaluate the product's quality. In what follows, we formalize the functional relationships among the demand, price, and review outcomes.

Customers' valuation and feedback. We assume that the value of a product to customer t is defined by

$$x_t = Q_t + \beta - \alpha p_t. \tag{1}$$

where $p_t \in [\underline{p}, \bar{p}]$, with $\underline{p} < \bar{p}$, is the price quoted by the monopolist to customer t. In (1), Q_t represents the quality experienced by customer t after purchase, which is taken to be a normal random variable, $Q_t \sim N(M, \Sigma^2)$. The standard deviation Σ captures the degree of heterogeneity in post-purchase quality perceptions. When a purchase happens, then with probability $u \in (0,1]$, the customer reports a numeric rating equal to q_t, where q_t is the projection of Q_t onto the interval $[\underline{q}, \bar{q}]$. We let $\mu = \mathsf{E}[q_t]$ and $\sigma^2 = \mathsf{E}[q_t^2] - \mathsf{E}[q_t]^2$ be the mean and variance of the truncated quality perception, respectively. The parametric structure of the underlying distribution for q_t is known to customers with the exception of its mean μ, which will be the objective of the customer learning process as described in the following paragraphs.

Review mechanism. Upon arrival, customer t observes a summary of reviews from previous buyers $s < t$: the number of reviews n_t and the average rating $r_t = \sum_{i=1}^{n_t} q_{t(i)}/n_t$ for $t \geq 2$, where $t(i)$ is the index of the ith reviewer. For $t = 1$, we let $r_1 = n_1 = 0$. We denote by (n_t, r_t) the state of the review platform observed by customer t. If customer t purchases the product, she reports a rating of q_t with probability $u \in (0,1]$. Then, the state of the review platform is updated as follows:

$$(n_{t+1}, r_{t+1}) = \begin{cases} (n_t + 1, \sum_{s=1}^{n_{t+1}} q_{t(i)}/n_{t+1}) & \text{if customer } t \text{ purchases and reports,} \\ (n_t, r_t) & \text{otherwise.} \end{cases} \tag{2}$$

Belief update mechanism. Customers share a common prior belief, expressed in our model through a normal random variable $\tilde{q}_0 \sim N(\mu_0, \sigma_0^2)$. After observing the state (n_t, r_t) of the review platform, customer t updates her belief according to Bayes' rule. Specifically, the posterior belief over μ is denoted by $\tilde{q}_t \sim N(\mu_t, \sigma_t^2)$, where

$$\mu_t = \frac{\gamma n_t r_t + \mu_0}{\gamma n_t + 1}, \quad \sigma_t^2 = \frac{\sigma_0^2}{\gamma n_t + 1}, \tag{3}$$

and $\gamma = \sigma_0^2/\hat{\sigma}^2$ (see, e.g., DeGroot 2005). The ratio γ measures the degree of ex ante uncertainty relative to the uncertainty in individual product reviews; for example, $\gamma = 0.1$ implies that approximately ten reviews from other customers are as influential to the posterior belief as their own prior belief. Motivated by these observations, throughout the paper we will often refer to γ as the "learning rate" of customers. Finally, customer t's perceived value of the product before purchase is $\tilde{x}_t = \tilde{q}_t + \beta - \alpha p_t$.

Customers' purchase decisions. Faced with the uncertain valuation $\tilde{x}_t$, customers in our model make purchasing decisions by maximizing the *expected* utility (e.g., Roberts and Urban 1988). In particular, we assume that customers are endowed with a linear utility function. (This structural assumption has no significant bearing on our results, but simplifies analysis and exposition.) The expected utility χ_t of a purchase can be written as

$$\chi_t = \mu_t + \beta - \alpha p_t, \tag{4}$$

and the expected utility of the no-purchase option is normalized to zero. We assume that there is an additive error ξ_t associated with χ_t, which arises from measurement, situational factors, and idiosyncratic individual behavior. The error ξ_t is a random variable distributed according to a standard Gumbel distribution. In this manner, the resulting probability of purchase follows the familiar logit model, hereafter referred to as the *demand function*, denoted by $\lambda(p_t, n_t, r_t)$, where

$$\lambda(p_t, n_t, r_t) = \ell(\chi_t) := \frac{\exp(\chi_t)}{1 + \exp(\chi_t)} = \frac{\exp(\mu_t + \beta - \alpha p_t)}{1 + \exp(\mu_t + \beta - \alpha p_t)}. \tag{5}$$

Note that $p_t \in [\underline{p}, \bar{p}]$ and $\mu_t \in [\min(\underline{q}, \mu_0), \max(\bar{q}, \mu_0)]$, and therefore, the demand function values lie in a compact set, i.e., $\lambda(p_t, n_t, r_t) \in [\underline{\lambda}, \bar{\lambda}]$, with $\underline{\lambda} > 0$ and $\bar{\lambda} < 1$. In particular, $\underline{\lambda} > 0$ implies that, even at the lowest possible average rating and the highest possible price, customers will choose to purchase the product with a positive probability. Recalling that a customer posts a rating with positive probability u after purchase, this ensures that the ratings accumulate in the review platform at least at a constant rate so that the unknown quality μ is eventually revealed, i.e., $\mu_t \to \mu$ almost surely as $t \to \infty$.

2.2. The Revenue Maximization Problem

The seller controls the demand function using a non-anticipating pricing policy $\pi = \{p_t \mid t = 1, \ldots, T\}$. We denote by Π the set of all non-anticipating pricing policies such that $p_t \in [\underline{p}, \bar{p}]$ for each t, and let (n_t, r_t) be the *state variables* at the time of arrival of customer t.

The total expected revenue is denoted by

$$J_T^{\pi} = \mathsf{E}^{\pi}\left[\sum_{t=1}^{T} p_t \lambda(p_t, n_t, r_t)\right], \tag{6}$$

where E^{π} is the expectation given that the seller uses policy π. The seller's problem is to find a pricing policy $\pi \in \Pi$ that maximizes the total expected revenue:

$$J_T^{*} = \sup_{\pi \in \Pi} \{J_T^{\pi}\}, \tag{7}$$

and the optimal solution is denoted by π^*.

It is quite difficult, if not impossible, to find an exact solution to the stochastic dynamic programming problem (7) because of the following two major issues. First, there is an intertemporal effect of price: specifically, the current price p_t affects the state variables for all subsequent customers $\{(n_s, r_s)\}_{s=t+1}^{T}$, which in turn influences the demand and pricing decisions in the future. Second, the state variables (n_t, r_t) evolve over $t = 1, \ldots, T$ in a stochastic, non-linear manner.

Numerical illustration of the optimal solution. For problems with a small number of customers, one can approximate the optimal solution by using lattices of the price and state variables and then solving the discretized problem numerically. In Figure 1, the price range $[\underline{p}, \bar{p}] = [0, 40]$ is discretized with interval 0.5 and the range of the average rating, $[1, 5]$, is discretized with interval 0.25. The two panels in Figure 1 illustrate optimal price paths from the discretized problem with $T = 80$, along with those of the average rating over the selling horizon. Figure 1 provides intuitive properties of the optimal solution to the stochastic problem. When the average rating is initially high relative to the true quality (left panel), the seller sets a high initial price to take advantage of the high rating and slowly decreases the price as the average rating converges. On the other hand, when the average rating is initially low (right panel), the optimal price starts from a low level to

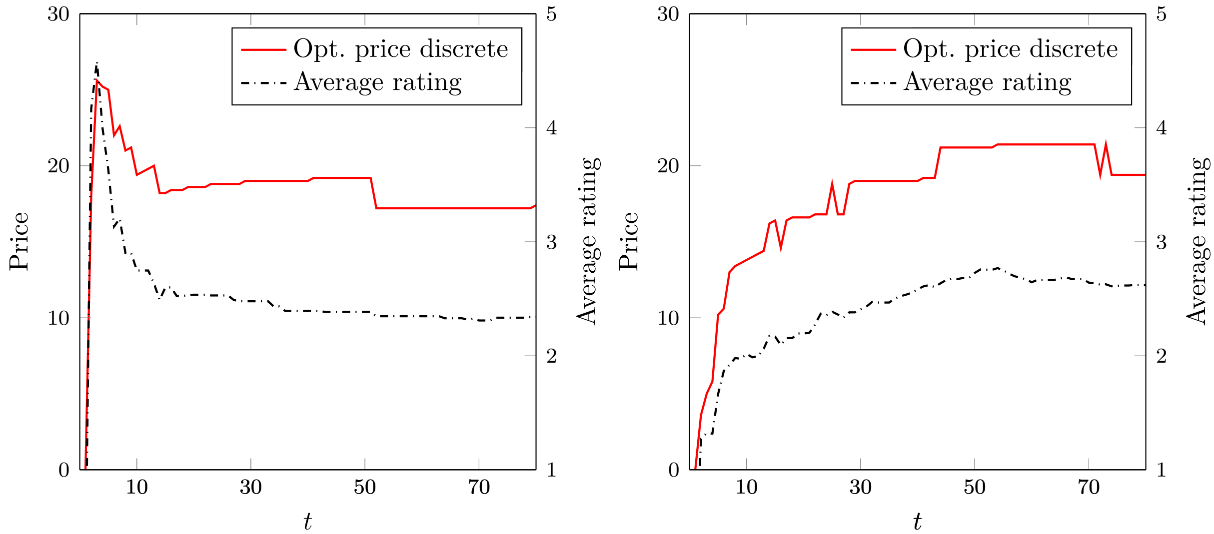

Figure 1 Sample paths of optimal prices for the discretized stochastic problem for $T = 80$ customers. The left panel corresponds to the case with high initial rating, while the right panel corresponds to the case with low initial rating. The parameters for the demand function are $(\alpha, \beta, \gamma) = (0.2, 0, 1)$; the mean prior belief is $\mu_0 = 1.5$; the mean of a rating is $\mu = 3$; and the probability of posting a review after purchase is $u = 1$.

accelerate information collection via the accumulation of customer reviews, and then it gradually increases with the average rating. In both cases, we remark that the optimal price depends not only on the average rating q_t, but also on the number of reviews n_t and time t; Figure 1 only shows the former dependence.

For problems with a large number of customers, however, the solution using discretization poses a significant computational issue. In what follows, we provide alternative methods to construct tractable, well-performing pricing policies that are implementable in many applications with a large number of customers.

2.3. Discussion of Modeling Assumptions

The modeling framework in the previous section is primarily chosen to facilitate delivery and provide reasonable first order approximation to the structure observed in some practical applications. We now briefly touch upon a number of modeling assumptions that can be classified into three categories: asymmetric information; unbiased reviews; time-homogeneity. Later we will present a small scale empirical study that indicates that this modeling framework has some grounding in observed data.

Asymmetric information. We assume asymmetric information between seller and customers. In particular, the seller has a private information about the mean quality μ, which is not available to customers. This is reasonable in the scenario where a manufacturer can retrieve this information via product testing or market research prior to sales. Further, in the scenario where a manufacturer sells indirectly through a retailer, the retailer may not have private information about quality but can truthfully infer its level through the strategic interaction with the manufacturer; for example, the manufacturer's wholesale price may serve as a signal of quality (Milgrom and Roberts 1986, Judd and Riordan 1994). In either scenario, although the seller's price may convey information to customers about quality, we assume that they do not adjust their quality estimate in response to that information. If customers *actively* learn quality through product reviews, which in the language of our model means γ is high, reviews may convey a stronger signal than price, as opposed to the case in traditional brick-and-mortar stores where pricing signals can be relatively strong.

Unbiased reviews. Our model assumes that a buyer reports their quality perception q_t directly to the review system, and not the net utility x_t that takes into account the price they paid. The assumption of price-independent reviews is widely used in the research community (e.g., Papanastasiou and Savva 2017, Yu et al. 2015). This is true if customers are altruistic and report the most useful information to subsequent customers, which is the (truncated) experienced level of quality q_t. Note that this assumption is aligned with the recent development of advanced review systems (e.g., Amazon Vine program), where customers are incentivized through special offers and discounts to post truthful reviews that reveal the quality. Further, if buyers are more likely (respectively, less likely) to appreciate the product's experience than non-buyers, there can be a positive (respectively, negative) bias in the average rating (Li and Hitt 2008). This does not happen in our model since the quality experience Q_t and the idiosyncratic preference ξ_t are not correlated. Our assumption of unbiased reviews helps us elucidate the key features of our problem in a simpler setting, but our theoretical results do not rely on this assumption; all of our theoretical results remain true as long as the average rating converges to any constant.

Time-homogeneity. In our model, customers' purchasing behavior and review propensity are independent of customer index t, i.e., the timing of their arrival along the sales horizon. This assumption is, of course, somewhat restrictive. In particular, it does not account for a phenomenon of self-selection; for example, a product may provide significantly higher value to a group of early adopters than to the broader customer population. It would require a game theoretic formulation to fully capture such a phenomenon. Although our main focus is on the model with time-homogeneous customers, we also discuss a relaxation of such an assumption in Section 6.1 to demonstrate robustness of our main results.

3. A Fluid Approximation and Qualitative Insights

To introduce our main ideas in a simple setting, this section studies a deterministic continuous counterpart of the stochastic discrete problem (7), and leverages it to suggest a simple and implementable pricing policy. We use the argument t in parenthesis to denote variables in continuous time; for example, we use $p(t)$ as a continuous counterpart of the price p_t in the original discrete problem.

3.1. Formulation

Define $\psi = \{p(t) \,|\, t \in [0,T]\}$ as our control variable and let Ψ be the set of all piecewise continuous paths of the control. In this formulation we ignore that each price $p(t)$ is constrained in the interval $[\underline{p}, \bar{p}]$; this constraint will be re-considered in Section 4.2, where we propose a pricing policy for the main problem (7). We denote the state variables by $\{n(t), r(t) \,|\, t \in [0,T]\}$, and let $\dot{n}(t)$ be the derivative with respect to time, i.e., $\dot{n}(t) := dn(t)/dt$. Then, $n(t)$ satisfies the ordinary differential equation:

$$\dot{n}(t) = u\lambda(p(t), n(t), r(t)), \tag{8}$$

with the initial condition $n(0) = 0$. Further, we let $r(t)$ be the counterpart of the average rating, which is simply equal to the population mean μ for all $t \in [0,T]$. The total revenue generated over $[0,T]$ under policy ψ is denoted by

$$\bar{J}_T^{\psi} = \int_0^T p(t)\lambda(p(t), n(t), r(t))dt, \tag{9}$$

and the seller's problem is to maximize the total (fluid) revenue:

$$\begin{aligned} \bar{J}_T^* = \max_{\psi \in \Psi} \quad & \bar{J}_T^{\psi} \\ \text{subject to} \quad & \dot{n}(t) = u\lambda(p(t), n(t), r(t)), \quad t \in [0,T] \\ & r(t) = \mu, \quad t \in [0,T] \\ & n(0) = 0. \end{aligned} \tag{10}$$

In this manner, (9)-(10) are the fluid model counterparts of the original stochastic dynamic problem defined in (6)-(7). Note that the maximum in (10) exists (Theorem 1), and hence the sup in the stochastic formulation (7) can be replaced with a max in (37). We hereafter refer to the solution of (10), ψ^*, as the (fluid) optimal dynamic pricing (ODP) policy. The corresponding state variables are denoted by $\{(n^*(t), r^*(t)) \mid t \in [0,T]\}$.

3.2. Characterization of the Optimal Solution to the Fluid Problem

We begin with Pontryagin's maximum principle to characterize necessary conditions satisfied by the ODP policy. The Hamiltonian function is defined as $H(p,n,r,\theta) = (p + u\theta)\lambda(p,n,r)$, where θ is the shadow price associated with the constraint $\dot{n} = u\lambda(p,n,r)$ in (10). The shadow price represents the net benefit of having the constraint $\dot{n} = u\lambda(p,n,r)$ relaxed by one unit; that is, θ is the marginal value of an additional single review. (In what follows we will eliminate t in the function arguments where there is no confusion, in order to improve clarity.)

The maximum principle states that the optimal solution, if it exists, maximizes the Hamiltonian at each instant t. The optimal price and state variables, (p^*, n^*, r^*), along with the shadow price, θ, satisfy the following system of differential equations:

$$0 = \frac{\partial H(p^*, n^*, r^*, \theta)}{\partial p} \qquad \text{(maximum of } H\text{)} \tag{11a}$$

$$\dot{\theta} = -\frac{\partial H(p^*, n^*, r^*, \theta)}{\partial n}, \ \theta(T) = 0 \qquad \text{(shadow price)} \qquad (11b)$$

$$\dot{n}^* = u\lambda(p^*, n^*, r^*), \ n^*(0) = 0 \qquad \text{(number of reviews)} \qquad (11c)$$

$$r^* = \mu \qquad \text{(average rating)} \qquad (11d)$$

The first condition (11a) implies that

$$p^* = -\frac{\lambda(p^*, n^*, r^*)}{\partial\lambda(p^*, n^*, r^*)/\partial p} - u\theta. \qquad (12)$$

The first term on the right-hand side of the preceding equation resembles the classical static, revenue-maximizing pricing rule, except the subtraction of $u\theta$. Recalling that the shadow price θ represents the marginal value of an additional review, one may interpret (12) as follows: If an additional review tends to increase future demand, i.e., $\theta > 0$, then there is an incentive to sacrifice profits now by lowering price in order to benefit later. On the other hand, if an additional review would decrease future demand, i.e., $\theta < 0$, then the seller would rather realize more instantaneous benefit by raising price, which simultaneously decelerates the accumulation of reviews. This trade-off between price and shadow price will be articulated later in (21).

It is not hard to verify that the following characterization of the optimal and shadow prices, along with the corresponding state variables, satisfies the first order conditions in (11a)-(11d):

$$\begin{aligned} p^*(t) &= \frac{1 + \exp(z^*)}{\alpha} + \frac{\mu - \mu_0}{\alpha}\left(\frac{1}{\gamma n^*(T) + 1} - \frac{1}{\gamma n^*(t) + 1}\right) \\ n^*(t) &= u\ell(z^*)t, \end{aligned} \qquad (13)$$

where the constant z^* is a solution to the following equation:

$$z + \exp(z) = \mu + \beta - 1 + \frac{\mu_0 - \mu}{\gamma u\ell(z)T + 1}. \qquad (14)$$

Note that along the path of the price given by (13), the demand function is constant for $t \in [0, T]$.

It is important to note, however, that (11a)-(11d) only provide necessary conditions for optimality, if an optimal solution exists. Therefore, further analysis is needed to guarantee the existence and optimality of the solution characterized in (13). The following theorem shows that, under suitable conditions, the closed-form expression in (13) is the *unique* optimal solution of (10).

Theorem 1 **(Optimal solution to the fluid problem)**. *An optimal solution to the fluid problem* (10) *exists. Further, if*

$$\frac{\gamma uT(\mu - \mu_0)}{4(1 + \gamma uT)} < 1, \qquad (15)$$

then the price characterized in (13) *is the unique optimal solution.*

Proof Sketch. To show existence, note first that in our model there is a one-to-one correspondence between price and demand (Li and Huh 2011). Therefore, given state variables n and r, the demand function $\lambda(p,n,r)$ has an inverse, denoted $p(\lambda,n,r)$. An equivalent formulation to the fluid problem (10) is

$$\begin{aligned} \max_{\lambda(\cdot)} \quad & \int_0^T p(\lambda(s),n(s),r(s))\lambda(s)ds \\ \text{subject to} \quad & \dot{n}(t)=u\lambda(t), t\in[0,T]. \end{aligned} \tag{16}$$

Note that the instantaneous revenue $p(\lambda(s),n(s),r(s))\lambda(s)$ in (16) is a concave function of $\lambda(s)$. (In the fluid problem (10), the instantaneous revenue $p(s)\lambda(p(s),n(s),r(s))$ is not concave in $p(s)$.) Combining concavitiy along with the compactness of $\lambda(\cdot)$ establishes the conditions required by Theorem 1 of Cesari (1966) for the existence of an optimal solution for (16), which in turn guarantees the existence of an optimal solution for (10). For uniqueness, it is sufficient to show the uniqueness of the solution z^* to equation (14). The condition (15) guarantees that the derivative of the function $g(z)$, defined as

$$g(z)=z+\exp(z)-\mu-\beta+1-\frac{\mu_0-\mu}{\gamma u\ell(z)T+1}, \tag{17}$$

strictly increases with z. Combined with the fact that $g(z)\downarrow -\infty$ as $z\downarrow -\infty$ and $g(z)\uparrow\infty$ as $z\uparrow\infty$, we establish that $g(z)$ crosses zero at a single point. □

Remark 1 (Sufficient conditions for Uniqueness). The condition (15) always holds when customers overestimate the quality a priori (i.e., $\mu_0>\mu$). In the case of an underestimating prior (i.e., $\mu_0<\mu$), it is straightforward to check that (15) holds if $\mu-\mu_0\le 4$, since

$$\frac{\gamma uT(\mu-\mu_0)}{4(1+\gamma uT)}<\frac{\mu-\mu_0}{4}. \tag{18}$$

Remark 2 (Generality of the closed-form solution). From a practical perspective, we remark that the condition (15) is not a significant restriction. Specifically, if (15) is violated, it can be easily shown that there can be at most three solutions to equation (14) since $g'(z)$ changes sign at most at two points over the entire horizon. That is, there can be at most three candidate price paths of the form (13), each of which corresponds to different values of z^*. In this case, one may choose the value of z^* that corresponds to a greater value of the objective function, which can be written in a closed form as

$$\bar{J}_T^*=\frac{\exp(z^*)T}{\alpha}+\frac{(\mu_0-\mu)}{\alpha\gamma u}\left(\log(\gamma u\ell(z^*)T+1)-\frac{\gamma u\ell(z^*)T}{\gamma u\ell(z^*)T+1}\right). \tag{19}$$

Given the optimal solution p^* and the associated state variable n^* specified in (13), the shadow price can be uniquely determined as:

$$\theta(t)=\frac{\mu-\mu_0}{u\alpha}\left(\frac{1}{\gamma n^*(T)+1}-\frac{1}{\gamma n^*(t)+1}\right). \tag{20a}$$

Since $n^*(t)$ is a nondecreasing function, the sign of the shadow price is determined by the difference between the mean prior belief μ_0 and the true mean μ. When customers have low prior belief ($\mu_0 < \mu$), the shadow price is positive (since an additional review increases the quality perception of the following customers) and non-increasing (since later reviews affect less customers). Conversely, when customers have high prior belief ($\mu_0 > \mu$), the shadow price is negative and non-decreasing. Further, combining (13) and (20a), the relationship between the optimal price and the shadow price can be described as follows:

$$p^*(t) + u\theta(t) = \frac{1}{\alpha(1-\ell(z^*))}, \quad t \in [0,T]. \tag{21}$$

This equation implies that the sum of the optimal price (i.e., the immediate benefit from selling a single unit of the product) and the shadow price (i.e., the latent future benefit from having an additional review) for each product is constant over time. For example, when the mean prior belief μ_0 is significantly lower than the true mean μ, additional reviews are highly beneficial for the seller (i.e., high $\theta(t)$) and the seller should decrease the price to accumulate reviews more quickly; that is, the immediate benefit is sacrificed by charging a low price at time t, which is compensated by higher revenue in the future.

3.3. The Value of Learning: Comparison with a Fixed-Price Policy

Theoretical characterization. The dynamic nature of the optimal solution described in the previous section reflects the time-varying demand process driven by customer learning. Despite the presence of time-varying demand functions in many practical settings, it is still fairly common, both in academic studies and in practice, to focus on fixed-price policies. In this section, we highlight the role of dynamic pricing by characterizing the parameter regimes where dynamic pricing leads to a significant improvement in revenue over fixed-price policies.

To this end, recall that the value of γ captures the speed of learning; the higher the value of γ the faster customers' belief converges. For the purpose of the following result we will consider an asymptotic regime where we consider the behavior of γT, separating into three distinctive regimes: (a) The "slow-learning" regime ($\gamma \ll 1/T$), where customers update their beliefs slowly, so that the bias $|\mu(t) - \mu|$ remains significant at the end of the selling horizon T; (b) the "fast-learning" regime ($\gamma \gg 1/T$), where customers quickly recover the true quality from reviews; and (c) the "intermediate" regime ($\gamma \sim 1/T$). In each of the three regimes, we analyze the revenue loss when the seller decides to use a fixed price as opposed to the optimal rating-dependent policy. This will be carried out by varying the values of γT; the dependence of relevant quantities on γ will be suppressed to avoid cluttering the notation.

In the context of a fixed time horizon, which is our main focus in the paper, one can interpret the above parametric regimes in terms of γ converging to zero (slow learning) and diverging to

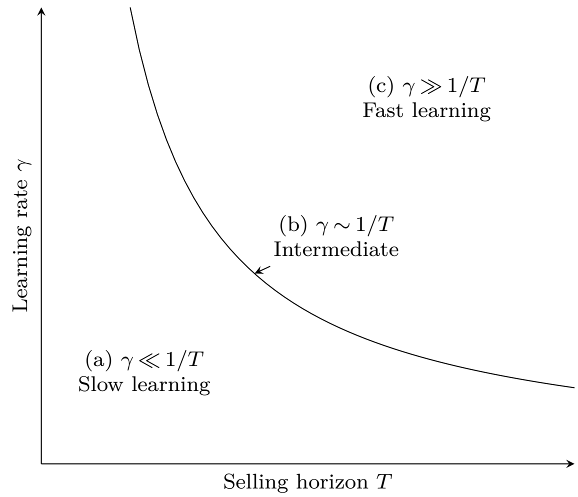

Figure 2 Characterization of the three regimes. Slow-learning and fast-learning regimes correspond to operating regimes represented by the areas below and above the curve, respectively, while the intermediate regime corresponds to the area around the curve.

infinity (fast learning), but further insight is gleaned when one considers this in relevant time scales. As such, the next theorem formalizes the notion that $1/T$ is the critical rate separating the two aforementioned regimes; this is what is referred to above (and below) as the *intermediate learning* regime. To formulate this, let OFP denote the (fluid) optimal fixed-price policy, and denote by p^{OFP} and $\bar{J}_T^{\text{OFP}}$ the price and (fluid) revenue of OFP, respectively. This will be compared with the fluid optimal revenue $\bar{J}_T^*$ which is the optimal value of the fluid problem (10).

THEOREM 2 **(Dynamic versus fixed-price policies)**. *Assume $\mu_0 \neq \mu$ and that the condition in* (15) *is satisfied. Then,*

$$1 - \frac{\bar{J}_T^{OFP}}{\bar{J}_T^*} \to \begin{cases} 0 & \text{as } \gamma T \to 0 \quad \text{[(a) slow learning]} \\ 0 & \text{as } \gamma T \to \infty \quad \text{[(b) fast learning]} \\ \Delta(\kappa) & \text{as } \gamma T \to \kappa \in (0,\infty) \quad \text{[(c) intermediate learning]}, \end{cases} \tag{22}$$

where $\Delta(\kappa) \in (0,1)$ is a constant that depends only on $\kappa \in (0,\infty)$.

The behavior of $\Delta(\kappa)$ is illustrated in Figure 4. Theorem 2 suggests that the value of dynamic pricing relative to fixed-price policies depends on the *learning dynamics* over the time horizon, represented by the product of the learning rate (γ) and the time horizon (T). In particular, when γT is "small", customers do not learn "enough" from product reviews due to a slow learning rate relative to the time horizon; in this regime, corresponding to the region below the curve in Figure 2, dynamic pricing does not exhibit significant gains relative to fixed-price policies. On the other hand, when γT is "large", customers learn significantly from product reviews due to a fast learning rate relative to the time horizon; in this regime, corresponding to the region above the curve in Figure 2, dynamic pricing only has a significant impact over a relatively short interval. Consequently, in these two regimes, the seller can achieve near-optimal revenue using a fixed-price policy. However, a significant loss in revenue can be incurred under OFP in the intermediate values of γT, depicted in the proximity of the curve in Figure 2. It is in this parameter regime that the value of learning is fully realized by the dynamic pricing policy.

To provide a more in-depth intuition behind the result of Theorem 2, we now state an immediate corollary of this theorem. To state the corollary, let $\epsilon^\pi(t)$ be the relative error in mean belief at time t under pricing policy π defined as:

$$\epsilon^\pi(t) = \left|\frac{\mu^\pi(t) - \mu}{\mu_0 - \mu}\right|. \tag{23}$$

That is, $\epsilon^\pi(t) \in [0,1]$ gives the normalized deviation between the customers' quality belief and the true quality at time t, expressed as a fraction of the initial bias $(\mu_0 - \mu)$. For instance, when $\epsilon^\pi(t)$ is close to one, customers have learned almost nothing from product reviews, whereas when $\epsilon^\pi(t)$ is close to zero, customers' quality beliefs are close to the true quality.

Corollary 1 **(Degeneracy in slow- and fast-learning regimes)**. *Assume $\mu_0 \neq \mu$ and that the condition in* (15) *is satisfied. Then, for fixed T, $p^*(t) - p^{OFP} \to 0$ for each $t \in [0,T]$ as $\gamma \to 0$ or ∞. Further,*

$$\epsilon^\pi(t) \to \begin{cases} 1 & \text{if } \gamma \to 0 \text{ (slow learning)} \\ 0 & \text{if } \gamma \to \infty \text{ (fast learning)} \end{cases} \tag{24}$$

for each $t \in (0,T]$ under the ODP and OFP policies.

Discussion of learning in the three regimes. In the case with low initial belief $(\mu_0 < \mu)$, the results of the preceding corollary are illustrated in Figure 3 in (a) slow-learning, (b) fast-learning, and (c) intermediate regimes. We comment on each of the three cases below.

(a) *Slow-learning regime.* In this regime customers are marginally influenced by reviews. In particular, in Figure 3(a), $\mu^*(t)$ increases slowly from μ_0 and customers do not eventually recover the true quality from reviews. Reflecting the "sluggish" changes in the belief process, the optimal price $p^*(t)$ also increases slowly over time. More specifically, from (13) and (14) it can be seen that $p^*(t) \approx p_0 := \arg\max_p \{p\ell(\mu_0 + \beta - \alpha p)\}$, where p_0 represents the optimal fixed price in the absence of reviews, or equivalently, when $\gamma = 0$. Consequently, the optimal price path $p^*(t)$ can be closely approximated by a single price in this regime.

(b) *Fast-learning regime.* In this case customers recover the true quality μ very early in the selling horizon; see the bottom panel of Figure 3(b). Using (13) and (14) once again, we can show that $p^*(t) \approx p_\infty := \arg\max_p \{p\ell(\mu + \beta - \alpha p)\}$, where p_∞ is the optimal fixed price with perfect learning from product reviews, or equivalently, when $\gamma = \infty$. As a result, the optimal price path $p^*(t)$ is in the vicinity of a single price, except for the initial part of the horizon which becomes negligible as $\gamma \to \infty$.

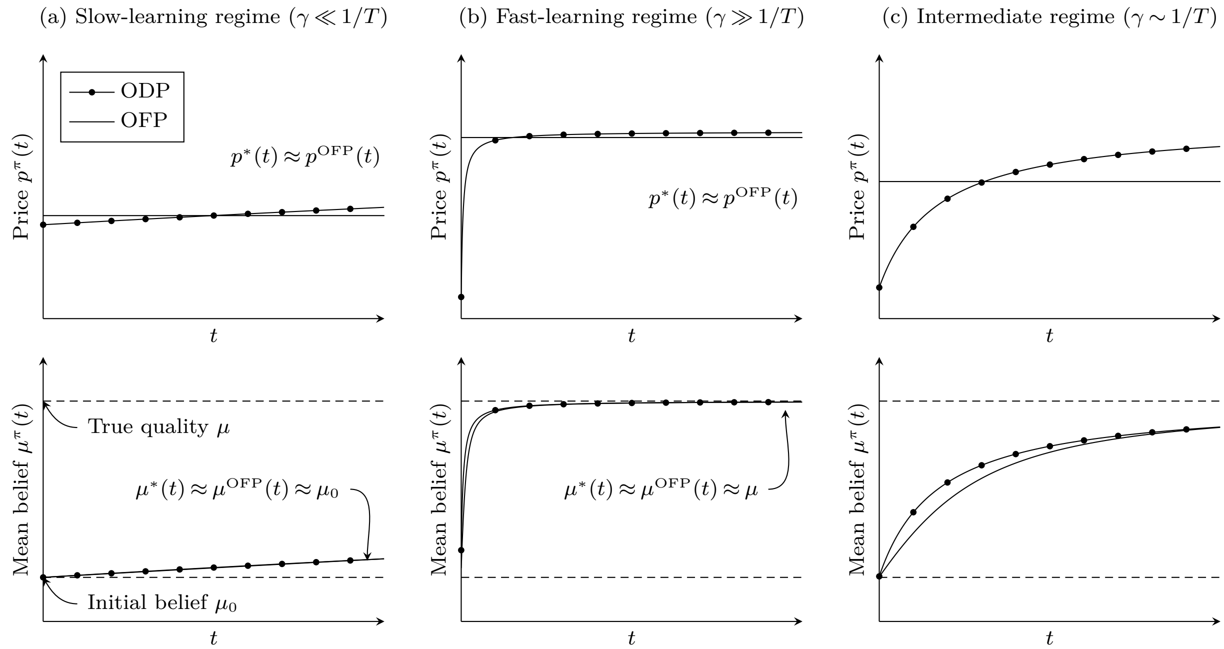

Figure 3 Comparison of the price and mean belief processes under ODP and OFP. In the slow-and fast-learning regimes, the paths of the price and mean belief are almost identical under ODP and OFP, whereas in the intermediate regime, ODP leads to a significantly different price and learning dynamics than OFP. The parameters for the demand function are $(\alpha, \beta, \mu_0, \mu, u, T) = (0.1, -1, 2, 4, 1, 100)$ for all cases and we vary $\gamma \in \{0.001, 0.1, 10\}$.

(c) *Intermediate regime.* In this setting customers are moderately influenced by reviews, and therefore, $\mu^\pi(t)$ does not degenerate to μ_0 or μ under both policies; see the bottom panel of Figure 3(c). In other words, customers are actively learning from reviews over the entire horizon, and hence, the dynamics of the mean belief depends significantly on the price path over time. Specifically, ODP initially sets a low price to encourage learning, at the expense of immediate revenue, which allows to collect the greater revenue in later periods. This intertemporal trade-off is not attainable with a single price—hence the large gap between ODP and OFP.

Lastly, we remark that the revenue loss of OFP stands out when the bias $|\mu_0 - \mu|$ is large. To see this, recall from Theorem 2 that, for $\kappa = \gamma T$, $\Delta(\kappa)$ measures the value of dynamic pricing relative to the fixed one. Figure 4 illustrates that, for fixed κ, the revenue gap between OFP and ODP is high for large bias $|\mu_0 - \mu|$. That is, the seller has more room to improve revenue using ODP when customers' average belief about product quality is significantly different than the truth. Also note from Figure 4 that small and large values of κ correspond to slow- and fast-learning regimes, in which cases $\Delta(\kappa)$ is close to zero.

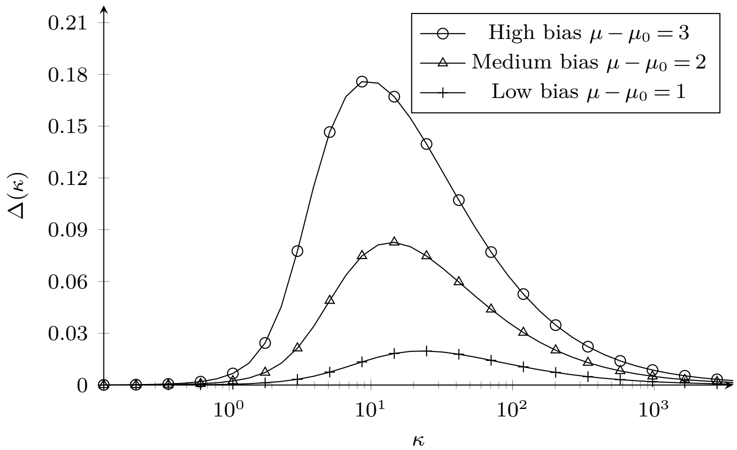

Figure 4 The revenue loss $\Delta(\kappa)$ of fixed-price policies . Each curve displays $\Delta(\kappa)$ in log-linear scale, where $\kappa = \gamma T$ with $T = 100$ and we vary $\gamma \in [0.01, 36]$. The parameters for the demand function are $(\alpha, \beta, \mu_0, u) = (0.1, -1, 0, 1)$, where high, medium, and low biases correspond to $\mu = 3, 2, 1$, respectively.

4. Proposed Pricing Policies: Asymptotic Efficiency and Finite-Time Performance

This section proposes a set of pricing policies for the stochastic revenue-maximization problem (7) using the structural insights from the closed-form characterization of the optimal solution in the fluid formulation (Section 4.1). These policies are shown to be efficient in an asymptotic regime where the number of customers grows to infinity (Section 4.2). With the aim of improving the finite-time performance, we analyze the proposed policies through extensive numerical experiments (Sections 4.3-5.2).

4.1. Proposed Pricing Policies

We discuss four pricing policies for the stochastic revenue-maximization problem.

Optimal Fixed-price (OFPS) policy. This static pricing strategy is a counterpart of OFP for the fluid problem, defined in Section 3.3. This policy sets $p_t = p^{\text{OFPS}} := \arg\max_p \{\mathsf{E}[\sum_{t=1}^{T} p\lambda(p, n_t, r_t)] \mid p \in [\underline{p}, \bar{p}]\}$ for each $t = 1, \ldots, T$, where the p^{OFPS} can be found numerically. This static pricing strategy lacks the capability of corrective action against stochastic fluctuations.

One-step Look-ahead (OSLA) policy. For each stage t, this policy makes a pricing decision as if there is only one more stage to go; in particular, $p_t = \arg\max_p \{p\lambda(p, n_t, r_t) \mid p \in [\underline{p}, \bar{p}]\}$. Although the OSLA policy does not exploit the structural properties of the fluid model and ODP, it provides an adjustment of price along the evolution of the review process and seems of practical interest.

Static Fluid-Matching (SFM) policy. This policy builds on the ODP policy $\psi^* = \{p^*(s) \mid s \in [0, T]\}$. Put $p_t = \max\{\min\{p^*(t), \bar{p}\}, \underline{p}\}$ for each stage $t = 1, \ldots, T$, and note that the resulting price path does not update according to the realization of (stochastic) reviews over time.

Policy 1: Dynamic Fluid-Matching (DFM)

for $t \in \{0, \ldots, T-1\}$ **do**
  Update the number of reviews n_t and the average rating r_t. Calculate μ_t from (3).
  Solve for z_t in (27) to obtain the solution $\psi_t^* = \{p_t^*(s) \mid s \in [t, T]\}$ from (26).
  Offer price $p_t = \max\{\min\{p_t^*(t), \bar{p}\}, \underline{p}\}$ for each product k.
  Let $t = t+1$
end

Dynamic Fluid-Matching (DFM) policy. In each stage t of the DFM pricing policy, we re-solve a fluid problem over the residual selling horizon $[t, T]$, where initial conditions for the state variables are set according to the observed state in that stage:

$$\begin{aligned}
\max_{\psi_t \in \Psi_t} \quad & \int_t^T p(s)\lambda(p(s), n(s), r(s))ds \\
\text{subject to} \quad & \dot{n}(s) = u\lambda(p(s), n(s), r(s)), \quad s \in [t, T] \\
& r(s) = \frac{r_t n_t + \mu(n(s) - n_t)}{n(s)}, \quad s \in [t, T] \\
& n(t) = n_t.
\end{aligned} \tag{25}$$

We let $\psi_t^* = \{p_t^*(s) \mid s \in [t, T]\}$ be the optimal solution to the modified fluid problem, with the subscript t indicating the current stage. Also denote by $\{n_t^*(s), r_t^*(s) \mid s \in [t, T]\}$ the path of the state variables under ψ_t^*. It is straightforward to adopt the characterization given in (13)-(14) to show that an optimal solution to (25) can be characterized as

$$\begin{aligned}
p_t^*(s) &= \frac{1 + \exp(z_t^*)}{\alpha} + \frac{\gamma n_t(r_t - \mu)}{\alpha}\left(\frac{1}{\gamma n_t^*(T) + 1} - \frac{1}{\gamma n_t^*(s) + 1}\right) \\
n_t^*(s) &= n_t + u\ell(z_t^*)(s - t) \\
r_t^*(s) &= \frac{r_t n_t + \mu(n_t^*(s) - n_t)}{n_t^*(s)},
\end{aligned} \tag{26}$$

where z_t^* is a solution z to the equation:

$$z + \exp(z) = \beta + \mu - 1 - \frac{\gamma n_t(r_t - \mu) + (\mu_0 - \mu)}{\gamma(n_t + u(T-t)\ell(z)) + 1}. \tag{27}$$

Based on (26) and (27), we design the DFM policy which iteratively solves the modified fluid problem (25) and use the solution ψ_t^* to make a pricing decision in every stage t. The formal definition of the DFM policy is given in Policy 1. Note that the DFM policy offers the initial value of the optimal solution to the modified fluid problem (25), truncated to the interval $[\underline{p}, \bar{p}]$. The price $p_t = \max\{\min\{p_t^*(t), \bar{p}\}, \underline{p}\}$ in the DFM policy is called a *fluid-matching* price. The proximity between the fluid-matching price and the optimal solution to (7) is illustrated in Figure 6; we delay detailed discussion of the figure until Section 4.3.

4.2. Asymptotic Performance of the Proposed Pricing Policies

This subsection offers an asymptotic characterization of the performance under the proposed pricing policies, formalized in the following theorem. For the remainder of the paper, $g_1(x) = \mathcal{O}(g_2(x))$ indicates that there exist $M < \infty$ and x_0 such that $|g_1(x)| \le M|g_2(x)|$ for all $x \ge x_0$.

THEOREM 3 **(Asymptotic performance).** *For $\pi \in \{OFPS, OSLA, SFM, DFM\}$, $J_T^* - J_T^\pi = \mathcal{O}(\sqrt{T})$.*

The preceding theorem asserts that all of the proposed pricing policies achieve near-optimal performance, verifying their efficacy in an asymptotic regime where the number of customers grows large. We outline the proof of Theorem 3 in two steps. First, we show that the optimal revenue in the fluid formulation $\bar{J}_T^*$ is close to the optimal revenue of the stochastic formulation J_T^* for sufficiently large T (see Lemma 1). Second, we analyze the asymptotic performance J_T^π of the four pricing policies with respect to $\bar{J}_T^*$ (see Lemma 2). A critical aspect of the proof is the convergence of the state variables in the stochastic formulation, which in turn leads to the convergence of optimal prices to their fluid counterparts.

LEMMA 1. *$|J_T^* - \bar{J}_T^*| = \mathcal{O}(\sqrt{T})$.*

REMARK 3 (PROOF SKETCH AND INTUITION). It can be shown that the state variables of the stochastic formulation (7) converge to the fluid counterpart; that is, $n_t \to \infty$ and $r_t \to \mu$ as $t \to \infty$. Using standard concentration inequalities, the convergence rate of r_t to μ can be characterized as $\mathsf{E}^\pi[(r_t - \mu)^2] = \mathcal{O}(1/t)$ for any $\pi \in \Pi$ (Lemma 4). Combined with the fact that the demand function $\lambda(\cdot)$ is Lipschitz continuous with respect to the state variables, the desired result of the lemma follows. $\square$

Note that J_T^* increases at least linearly in T because the expected revenue in each stage is bounded. Hence, the preceding lemma implies that the difference between J_T^* and $\bar{J}_T^*$ becomes negligible relative to J_T^* for sufficiently large T. This, in turn, implies that the revenue loss of a policy π with respect to J_T^* is relatively small if the expected revenue J_T^π is "not too far" from $\bar{J}_T^*$. The latter statement is proved in the following lemma.

LEMMA 2. *$|\bar{J}_T^* - J_T^\pi| = \mathcal{O}(\sqrt{T})$ for $\pi \in \{OFPS, OSLA, SFM, DFM\}$.*

REMARK 4 (PROOF SKETCH AND INTUITION). The main part of the proof is to show that p_t^π for each $\pi \in \{\text{OFPS}, \text{OSLA}, \text{SFM}, \text{DFM}\}$ is close to the fluid counterpart $p^*(t)$ for sufficiently large t. This is true for $\pi = \text{SFM}$ by definition. Define $\lambda_{\lim}(p) := \lim_{t\to\infty} \lambda(p, n_t, r_t)$ and let $p_{\lim} = \arg\max_{p \in [\underline{p}, \bar{p}]} \{p\lambda_{\lim}(p)\}$. Since $n_t \to \infty$ and $r_t \to \mu$ almost surely as $t \to \infty$, we have that $\lambda_{\lim}(p) = \ell(\mu + \beta - \alpha p)$. It can be seen that $p^*(t) \approx p_{\lim}$ for sufficiently large $t \le T$. By construction of each policy $\pi \in \{\text{OFPS}, \text{OSLA}, \text{DFM}\}$, we can show that $p_t^\pi \approx p_\infty$ for sufficiently large t, which in turn implies $p_t^\pi \approx p^*(t)$. To be specific, the convergence rate of $\mathsf{E}^\pi|p_t^\pi - p^*(t)|$ is governed by the expected gap in the state variables, $\mathsf{E}^\pi|r_t^\pi - \mu|$, which is $\mathcal{O}(1/\sqrt{t})$ by standard concentration inequalities. Combined with the Lipschitz continuity of the demand function $\lambda(\cdot)$, the convergence rate of $|\bar{J}_T^* - J_T^\pi|$ is governed by that of $\mathsf{E}^\pi|p_t^\pi - p^*(t)|$, from which we deduce the desired conclusion. $\square$

Theorem 3 follows immediately from Lemmas 1-2 since $J_T^* - J_T^\pi \leq |J_T^\pi - \bar{J}_T^*| + |\bar{J}_T^* - J_T^\pi|$ by the triangle inequality. Theorem 3 provides the effectiveness of the proposed policies in the asymptotic setting. However, it is important to note that all of these policies are not necessarily competitive over a finite horizon. As seen in Section 3.3, dynamic pricing strategies are more appropriate than fixed-price policies for small T. In what follows, we compare the performance of the proposed policies via simulation (Section 4.3) and counterfactual analysis (Section 5.2).

4.3. Simulation for Small-scale Problems

We conduct a simulation for problems with small number of customers $T \leq 160$ to evaluate the performance of the DFM policy with respect to the optimal revenue J_T^* of the stochastic problem. Since it is not numerically tractable to compute the exact value of J_T^*, we report the optimal revenue of the discretized stochastic problem, with the discretization being done in the same manner as in Figure 1.

Figure 5 illustrates the average revenues per-customer under these pricing policies. The performance of the DFM policy is surprisingly good; the revenue generated under the DFM policy is almost identical to that from the optimal policy for the discretized stochastic problem, even with the very small number of customers $T = 20$. Figure 6 shows sample price paths of the optimal pricing policy along with the *fluid-matching* prices that are fabricated from DFM along the path of state variables generated by the optimal pricing policy. The proximity between the two price paths suggests homologous review dynamics between the DFM and the optimal pricing policies—hence the near-optimal performance of DFM. The performance of the SFM policy is relatively poor; the gap between the DFM and SFM policies can be viewed as the loss due to ignoring uncertainties in the review process. The OSLA policy also performs poorly compared to the DFM policy as the former ignores the intertemporal effect of price. The OFPS policy captures neither the uncertainties in the review process nor the intertemporal effect of price, and hence, its performance is relatively poor.

Note that the proposed pricing policies perform differently in the small-scale problems in Figure 5, but they perform comparably in large-scale problems consistent with Theorem 3; see also Figure 7 in Appendix A.2.

5. Numerical Study

5.1. Validating the Demand Model

We show via empirical validation that our model serves as a good representation of online markets operating with product reviews.

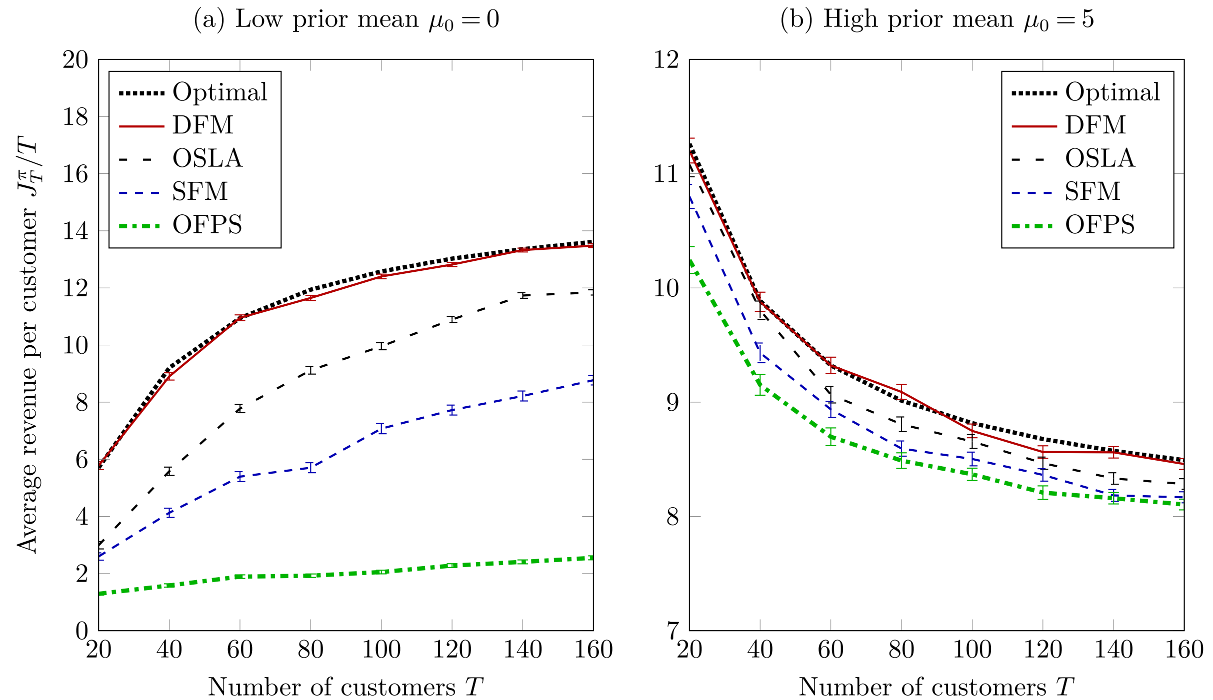

Figure 5 The total revenues under different pricing policies for low and high prior beliefs: (a) $\mu_0 = 0 < 2.5 = \mu$ **and (b)** $\mu_0 = 5 > 2.5 = \mu$**. In both cases, the model parameters are** $(\alpha, \beta, \gamma, u, \sigma_q) = (0.1, -1, 1, 1, 2)$ **with price bounds** $(\underline{p}, \bar{p}) = (0, 40)$**.The revenues of these policies are estimated by taking averages of** 10^3 **simulation trials. The error bars represent standard errors.**

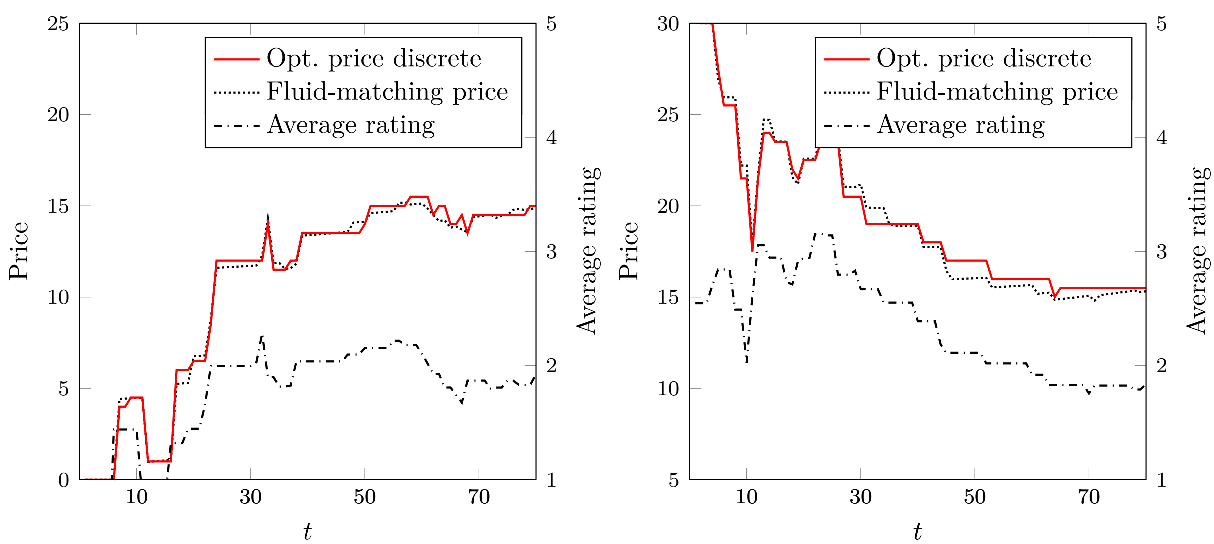

Figure 6 Sample paths of the optimal and fluid-matching prices for low and high prior beliefs: (a) $\mu_0 = 0 < 2.5 = \mu$ **and (b)** $\mu_0 = 5 > 2.5 = \mu$**. In both cases, the model parameters are** $(\alpha, \beta, \gamma, u, \sigma_q) = (0.1, -1, 1, 1, 2)$**.**

5.1.1. Available Data. In order to validate the model developed in Section 2.1, we use data collected from Amazon, one of the largest online retailers in the world. Our dataset includes a

Table 1 Descriptive statistics for the reduced set of our data.

	Mean	Standard deviation	10th Percentile	50th Percentile	90th Percentile
Observation period (days)	471.9	79.1	376.4	461.5	596.0
Sales rank	15790.0	58968.9	181.8	1738.5	25151.5
Price	17.8	7.9	11.18	16.3	25.7
Daily price variation	0.0049	0.0022	0.0025	0.0044	0.0073
Daily price change frequency	0.92	0.51	0.32	0.75	1.70
Average rating	4.54	0.33	4.1	4.6	4.87
Number of reviews	888.2	1045.1	32.1	516.0	2142.8

thousand hardcover books released between January 2015 and December 2016 on Amazon. For each product we collect sales rank, price, number of reviews, and average rating via web scraping on a daily basis. Although we cannot observe sales directly, we can utilize the relationship between sales rank and actual sales found by Schnapp and Allwine (2001) to estimate sales during the relevant period: $\log(\text{ImpliedSales}_m) = 9.61 - 0.78\log(\text{SalesRank}_m)$, where m represents the number of days elapsed after the product release; our qualitative conclusions are not affected by different coefficients of the preceding relationship. We use the implied sales as a proxy to the actual sales in this paper. Reviews posted on Amazon.com have single-valued integer ratings, ranging from 1 to 5. (In this section, variables of course depend on the product but, for the sake of exposition, we suppress that dependence.)

For the purpose of validating our demand model, we restrict our attention to products that satisfy the following criteria: (i) the observation period is greater than 300 days; and (ii) daily price variations for the product's life cycle is greater than 10^{-3}. Daily price variations are measured by

$$\sum_{m=1}^{M-1} \frac{|\text{Price}_{m+1} - \text{Price}_m|}{\text{Price}_m}, \tag{28}$$

where Price_m represents the average daily prices on the mth day for each product. The criterion (i) ensures sufficient amount variation in terms of the number of reviews and the average rating. There are 72 products that meet these criteria in our data set; Table 1 summarizes descriptive statistics of these products.

5.1.2. Estimation Procedure. Our data set does not include information about individual users who visited products' web pages; in particular, purchasing decisions of individual users are not available in our data. Hence, we estimate the demand model using daily sales as follows. First, we estimate the daily number of visitors for a product's web page, denoted NumVisits, using the relationship:

$$\text{NumVisits} = \frac{\sum_{m=1}^{M} \text{ImpliedSales}_m / M}{\text{CR}}, \tag{29}$$

where CR represents the overall visit-to-purchase conversion rate throughout the sales horizon. Note that we assume that NumVisits is constant over time. (This has no significant bearing on our results;

in particular, we draw the same conclusions using time-adjusted NumVisits in Appendix A.) The value of CR cannot be estimated from our data; in light of Monetate (2016) that reports CR is around 2.5-3% on Amazon in 2016, we use five scenarios of CR $\in \{1\%, 2\%, 3\%, 4\%, 5\%\}$ in our estimation to ensure robustness.

Given NumVisits for each product, the implied sales in day m follows binomial distribution as follows:

$$\text{ImpliedSales}_m \sim \text{Binomial}(\text{NumVisits}, \lambda_m), \tag{30}$$

where $\lambda_m := \lambda(\text{Price}_m, \text{NumReviews}_m, \text{AvgRating}_m)$ with Price_m, NumReviews_m, and AvgRating_m representing the average price, number of reviews, and average rating on day m, respectively. We estimate the parameters of the demand function $\lambda(\cdot)$ using maximum likelihood.

5.1.3. Benchmark Models. We also estimate three benchmark models as follows.

1. *Linear model.* Instead of the logistic function in our base model, the demand function is given by

$$\lambda_a(p,n,r) = \begin{cases} 0, & \text{if } \phi < 0, \\ \phi, & \text{if } 0 \le \phi < 1, \\ 1, & \text{if } \phi \ge 1, \end{cases} \tag{31}$$

where $\phi = (\gamma n r + \mu_0)/(\gamma n + 1) + \beta - \alpha p$.

2. *Logistic price-only model.* This model considers price as a single predictor with the logistic function; that is, $\lambda_b(p) = \exp(\beta - \alpha p)/(1 + \exp(\beta - \alpha p))$.
3. *Linear price-only model.* The demand function is linear and the price is a single predictor; that is,

$$\lambda_c(p_t) = \begin{cases} 0, & \text{if } \beta - \alpha p_t < 0, \\ \beta - \alpha p_t, & \text{if } 0 \le \beta - \alpha p_t < 1, \\ 1, & \text{if } \beta - \alpha p_t \ge 1. \end{cases} \tag{32}$$

5.1.4. Out-of-Sample Testing Procedure. We use the parameters estimated from past data to test the ability to predict sales in the future. The out-of-sample testing procedure for each product is as follows.

Out-of-sample testing procedure.

1. Estimate NumVisits using (29).
2. For each day $200 \le m \le M - 1$:
   (a) Calibrate parameters of a demand model using the data for first m days.
   (b) Predict the sales on day $m + 1$ as

$$\text{PredictedSales}_{m+1} = \text{NumVisits} \times \lambda_{m+1}. \tag{33}$$

3. Estimate the correlation between $\{\text{ImpliedSales}_m\}_{m=201}^{M}$ and $\{\text{PredictedSales}_m\}_{m=201}^{M}$.

Table 2 Estimation of predictive power measured by the correlation between the implied and predicted sales over time. The values in the parenthesis represent the standard deviations across 72 products.

	Conversion rate (CR)				
	1%	2%	3%	4%	5%
Base	0.709 (0.274)	0.743 (0.244)	0.747 (0.235)	0.753 (0.230)	0.759 (0.230)
Linear	0.496 (0.308)	0.566 (0.261)	0.623 (0.230)	0.654 (0.244)	0.669 (0.231)
Logistic price-only	0.440 (0.430)	0.443 (0.430)	0.445 (0.430)	0.448 (0.430)	0.450 (0.431)
Linear price-only	0.376 (0.479)	0.394 (0.489)	0.397 (0.493)	0.396 (0.493)	0.396 (0.493)

5.1.5. Predictive Power. The correlation between the implied and predicted sales captures the proportion of variability in sales explained by model. Table 2 summarizes the correlations from the out-of-sample testing procedure. In the scenario with conversion rate of 3%, the correlation is estimated to be about 0.747 on average under our base model, while it is only 0.623 under the linear model. Further, the correlations for the two price-only models are significantly lower than those that involve product reviews. This conforms to analogous findings in previous literature (e.g., Chevalier and Mayzlin 2006, Li and Hitt 2008). These observations are consistent for different scenarios of the conversion rate.

5.1.6. Robustness of the Model. Our base model and the estimation procedure in Section 5.1.2 allows one to demonstrate the predictive power of price and review variables when accounting for the sales in online markets. In Appendix A.1 we consider other variants of the demand model, which take into account: (a) the number of visitors that varies with time; (b) the effect of being listed in New York Times Best Seller list, widely considered as the preeminent list of best-selling books in the US; and (c) the effect of price promotions. Repeating the out-of-sample testing that was performed in this section, we demonstrate that these additional features do not necessarily improve the predictive power of our base model.

5.2. Counterfactual Analysis

To generalize the simulation results in Section 4.3 to settings that are close to the actual markets, we further conduct a simulation study based on the demand model calibrated to our data set. We simulate the revenue of the policy $\pi \in \{\mathrm{OSLA}, \mathrm{SFM}, \mathrm{DFM}\}$ using the following procedure.

1. Set the number of customers T and the probability of posting a review u.
2. For each product $k = 1, \ldots, 72$:
    - Set $t = 0$ with $n_0 = 0$ and $r_0 = 0$.

Table 3 **Relative revenue performance of the OSLA, SFM, and DFM policies with respect to the OFPS policy, measured by $J_T^{\pi}/J_T^{\mathsf{OFPS}}$. For each of the 72 products, the revenue is estimated from 10^3 simulation trials and the standard error is less than 2% of its value.**

	$T=10^3$			$T=10^6$		
	OSLA	SFM	DFM	OSLA	SFM	DFM
Min.	0.226	0.910	0.953	0.690	0.820	0.998
5%	0.757	0.963	0.984	0.872	0.999	0.999
25%	0.946	1.002	1.008	1.000	1.002	1.002
50%	1.001	1.014	1.033	1.006	1.007	1.008
75%	1.024	1.062	1.136	1.012	1.012	1.016
95%	0.129	1.358	1.889	1.030	1.306	1.519
Max.	1.535	1.729	2.173	1.311	1.644	1.871
Average	0.975	1.064	1.144	0.998	1.025	1.069
Std. dev.	0.164	0.148	0.268	0.088	0.145	0.184

- Offer price p_t to customer t, where p_t is determined by π. Customer t purchases the product with probability $e^{\phi_t}/(1+e^{\phi_t})$ with

$$\phi_t = \frac{\gamma n_t \mu + \mu_0}{\gamma n_t + 1} + \beta - \alpha p_t \tag{34}$$

  where α, β, γ, and μ_0 are the calibrated parameters for product k and μ is the mean quality for product k, estimated by the average rating on the last day of the sales horizon.
- In case customer t purchases the product, with probability u, she reports a review equal to q_t, where q_t is randomly drawn from the set of actual ratings for product k.
- Let $t=t+1$, update state variables, and repeat until $t \le T$.

3. Calculate total revenue collected from T customers.

The above procedure is repeated for the DFM policy as well as for three benchmark policies given in Section 4.3. We assume that u is common across the products and set $u = 6\%$, which is estimated by dividing the total number of reviews by the total implied sales from our data set. We consider small and large market sizes with $T = 10^3$ and $T = 10^6$, respectively.

Table 3 summarizes the simulation results. For each product and for each pricing policy, we estimate revenue with the average over 10^3 simulation trials. In this table, the percentage standard error of the estimated revenues is less than 2%, and it is thus omitted. We take the optimal fixed-price policy as a baseline and estimate $J_T^{\pi}/J_T^{\mathrm{OFPS}}$ for $\pi \in \{\mathrm{OSLA}, \mathrm{SFM}, \mathrm{DFM}\}$. Overall, the same qualitative conclusions from the small-scale problems in Section 4.3 hold in our counterfactual analysis: The DFM policy significantly outperforms the benchmark policies, while the SFM and OSLA policies only exhibit mild improvement over the optimal fixed-price policy.

Specifically, when $T = 10^3$, the median ratio is 1.033 under the DFM policy, that is, the DFM policy improves the revenue by 3.3% from that of the optimal fixed-price policy, while the SFM and OSLA policies only provide 1.4% and 0.1% of improvement, respectively. The results are more extreme in terms of lower and higher percentiles. The large gap between the DFM and the OSLA policies suggests that the seller may incur a significant revenue loss if the intertemporal effect of price is not properly accounted for. Further, the gap between the DFM and the SFM policies can be considered as the value of taking into account the stochastic evolution of reviews.

When the market size is larger, i.e., $T = 10^6$, the DFM policy still outperforms other benchmarks, but the improvement is not as substantial as in the case with $T = 10^3$. On the three right columns of Figure 3, observe that the median ratio under the DFM policy is about 1.008, smaller than the ratio of 1.033 in the case with $T = 10^3$. This is expected because, when the number of customers is sufficiently large, a fixed price is close to optimal as seen in Section 3.3.

The preceding observations allude to an important aspect of the DFM policy: the most significant benefit of the DFM policy comes from the transient portion of the product's life cycle, a portion from launch to stabilization of the average rating, which is consistent with our observation in Section 3.3. In particular, if T is small and the fluctuation of μ_t is significant, there can be significant revenue loss corresponding to a fixed-price policy. Hence, the DFM policy is more suitable for products with short life cycles, such as fashion items, than for those with long life cycles, such as durable goods.

Lastly, we remark that, if our model differs significantly (i.e., is misspecified) relative to the underlying demand, the revenue performance of the DFM and benchmark policies in Table 3 could be quite different, since the sales are generated using simulation based on our ground truth demand model. Further, in practice sellers may not have full information on the parameters of the demand model, in which case they cannot implement the DFM policy exactly. Hence, the performance of the DFM policy in Table 3 must be viewed as a best-possible performance benchmark.

6. Extensions to the Basic Problem

6.1. Time-Varying Demand

In this section we assume that the demand function depends on time since the start of the selling season. Assume further that dependence in time is given as a positive multiplicative factor $g(t)$, so the demand function is given as

$$\tilde{\lambda}(p,n,r,t) = g(t)\lambda(p,n,r) \quad t \in [0,T]. \tag{35}$$

For example, $g(t)$ may be a convex, decreasing function in the scenario with early adopters that are more likely to purchase the product than the broader population. Also, $g(t)$ can be cyclical when demand exhibits a periodic behavior; e.g., textbooks. In this case, a simple method introduced in Gallego and van Ryzin (1994) can be used to transform the problem into one in which demand is time homogeneous. Define

$$G(t) = \int_0^t g(s)ds, \quad t \in [0,T] \tag{36}$$

Then it is not difficult to see that the fluid problem (10) is equivalent to

$$\begin{aligned} \max_{\psi\in\Psi} \quad & \int_0^{G(T)} \tilde{p}(\tau)\lambda(\tilde{p}(\tau),\tilde{n}(\tau),\tilde{r}(\tau))d\tau \\ \text{subject to} \quad & \dot{\tilde{n}}(\tau) = u\lambda(\tilde{p}(t),\tilde{n}(t),\tilde{r}(t)), \quad \tau\in[0,G(T)] \\ & \tilde{r}(t)=\mu, \quad \tau\in[0,G(T)] \\ & n(0)=0. \end{aligned} \tag{37}$$

The optimal solution with resepct to the original time scale can be recovered by $p^*(t)=\tilde{p}^*(G(t))$. Further, it is straightforward to construct a variant of DFM, called DFM-t, which basically re-solves the problem (37) for each t. Although the structural properties of the DFM-t pricing policy remain unchanged from the case with constant review propensity, the performance can change significantly as the time-varying review propensity affects how fast the average rating converges to true quality. In the fluid formulation (37), the (stochastic) average rating r_t is replaced with μ. Hence, the convergence rate of r_t to μ influences the proximity between the fluid problem (37) and its stochastic counterparts. This intuition is formalized in the following corollary.

Corollary 2 **(Performance of DFM-t)**. *Suppose that the demand function is given as* (35)*. Then for $\pi = DFM\text{-}t$,*

$$1-\frac{J_T^\pi}{J_T^*} = \mathcal{O}\left(1\Big/\sqrt{G(T)}\right). \tag{38}$$

Note that if $g(t)\to 0$ as $t\to\infty$, customers with large t are less likely to purchase and post reviews. Specifically, if $g(t)\approx c/t^v$ with $v>1$, then the average rating r_t does not converge to the true quality μ with positive probability. (See Lemma 4.) In this case, the bound in (38) does not converge to zero. Conversely, if $g(t)\approx c/t^v$ with $v\in[0,1]$, then the average rating converges to the true quality with probability one and the optimality gap converges to zero at the rate of $T^{(v-1)/2}$.

6.2. Multiproduct Case with Substitutable Demand

In this subsection we consider a seller that sells K distinct products for T customers. We adopt the same notation as in the single-product case, but use the subscript k to denote variables particular to product k. Denote by $P_t=(p_{1t},\dots,p_{Kt})$ the prices of the K products for each t and let $N_t = (n_{1t},\dots,n_{Kt})$ and $R_t=(r_{1t},\dots,r_{Kt})$ be the vectors of the number of reviews and the average rating for customer t, respectively. The value of product k for costumer t is $\beta_k+\hat{q}_{kt}-\alpha_k p_{kt}$, where $\hat{q}_{kt}\sim N(\mu_k,\sigma_k^2)$ the quality of product k as perceived by costumer t. Customers are aware of the distribution of $\hat{q}_{kt}$ except for the value of the mean μ_k and share a common prior belief over μ_k denoted by $\hat{q}_{k0}\sim N(\mu_{k0},\sigma_{k0}^2)$. After observing R_t and N_t customer t has a posterior belief $\hat{q}_{kt}\sim N(\mu_{kt},\sigma_{kt}^2)$ where

$$\mu_{kt}=\frac{\gamma_k n_{kt} r_{kt}+\mu_{k0}}{\gamma_k n_{kt}+1}, \quad \sigma_{kt}^2=\frac{\sigma_{k0}^2}{\gamma_k n_{kt}+1}, \tag{39}$$

and $\gamma_k = \sigma_{k0}^2/\sigma_k^2$.

Each customer purchases at most one product. The vector of purchasing probabilities, or the multiproduct demand function, is denoted by $\Lambda(P_t, N_t, R_t) = (\lambda_1(P_t, N_t, R_t), \ldots, \lambda_K(P_t, N_t, R_t))$, where

$$\lambda_k(P_t, N_t, R_t) = \frac{\exp(\chi_{kt})}{1 + \sum_{i=1}^K \exp(\chi_{it})} \tag{40}$$

and $\chi_{kt} = \beta_k + \mu_{kt} - \alpha_k p_{kt}$.

The seller chooses a non-anticipating pricing policy $\pi = \{P_t \in [\underline{p}, \bar{p}]^K \mid t = 0, \ldots, T-1\}$ to maximize the total expected revenue given by

$$J_T^\pi = \mathsf{E}^\pi \left[\sum_{t=0}^{T-1} P_t \cdot \Lambda(P_t, N_t, R_t) \right], \tag{41}$$

where $A \cdot B = \sum_{k=1}^K a_i b_i$ denotes the inner product of the vectors $A = (a_1, \ldots, a_K)$ and $B = (b_1, \ldots, b_K)$. A fluid approximation of the multiproduct revenue maximization problem can be formulated as

$$\begin{aligned} \bar{J}_T^* = \max_{\psi \in \Psi} \quad & \int_0^T P(t) \cdot \Lambda(P(t), N(t), R(t)) dt \\ \text{subject to} \quad & \dot{N}(t) = u \Lambda(P(t), N(t), R(t)), \quad t \in [0, T] \\ & R(t) = (\mu_1, \ldots, \mu_K), \quad t \in [0, T] \\ & N(0) = (0, \ldots, 0), \end{aligned} \tag{42}$$

where $P(t)$, $N(t)$, and $R(t)$ are counterparts of the discrete variables P_t, N_t, and R_t, respectively. We hereafter refer to (42) as the multiproduct fluid problem.

Applying the maximum principle to the multiproduct fluid problem, in Theorem 4 we show that an optimal solution $P^*(t) = (p_1^*(t), \ldots, p_K^*(t))$, provided it exists, is characterized by

$$\begin{aligned} p_k^*(t) &= \frac{1}{\alpha_k} + \sum_{i=1}^K \frac{\exp(z_i^*)}{\alpha_i} + \frac{\mu_k - \mu_{k0}}{\alpha_k} \left(\frac{1}{\gamma_k n_k^*(T) + 1} - \frac{1}{\gamma_k n_k^*(t) + 1} \right) \\ n_k^*(t) &= u \lambda_k(P^*(t), N^*(t), R^*(t)) = u \ell_k(Z^*) t, \end{aligned} \tag{43}$$

where $N^*(t) := (n_1^*(t), \ldots, n_K^*(t))$ and $R^*(t) := (\mu_1, \ldots, \mu_K)$ are, respectively, the trajectories of the vectors of the numbers of reviews and of the average ratings induced by $P^*(t)$, and where $\ell_k(X) := \exp(x_k)/(1 + \sum_{j=1}^K \exp(x_j))$. In the above relations, the vector of constants $Z^* = (z_1^*, \ldots, z_K^*)$ satisfies

$$z_k^* + \alpha_k \sum_{i=1}^K \frac{\exp(z_i^*)}{\alpha_i} = \beta_k + \mu_k - 1 - \frac{\mu_k - \mu_{k0}}{\gamma_k u \ell_k(Z^*) T + 1}, \quad k = 1, \ldots, K. \tag{44}$$

Notice that, similarly to the single-product case, the demand function for product k evaluated over the optimal trajectory is constant and is equal to $\ell_k(Z^*)$. In particular, the constants $z_1^*, \ldots, z_K^*$ determine the product demand levels over the optimal price trajectory $P^*(t)$.

Our candidate solution $P^*(t)$ is the unique optimal solution to (42), provided that Z^* is the unique solution to (44). The following lemma provides the condition that guarantees the uniqueness of Z^*.

LEMMA 3. *Let Z^* be any solution to* (44)*. Then, for any k, there exist $\underline{l}_k$ and $\bar{l}_k$ such that $0 < \underline{l}_k \leq l_k(Z^*) \leq \bar{l}_k < 1$. Moreover, assume that the following condition is satisfied*

$$\max_{j=1,\ldots,K}\left[\frac{\exp(\beta_j - 1 + \max(\mu_{j0},\mu_j))}{\alpha_j/\sum_{i=1}^K \alpha_i} + \sum_{j=1,j\neq k}^{K}\frac{|\mu_j-\mu_{j0}|}{1+\gamma_j u\underline{l}_j T} + \frac{|\mu_k-\mu_{k0}|}{4}\frac{\gamma_k uT}{1+\gamma_k uT}\right] < 1. \tag{45}$$

Then Z^* is the unique solution of (44)*.*

Building on the result of Lemma 3, we can now state the main result of this section.

THEOREM 4 **(Optimal solution of the fluid multiproduct problem)**. *An optimal solution to the multiproduct fluid problem* (42) *exists and can be characterized by* (43)*. Further, if condition in* (45) *is satisfied, then the candidate solution $P^*(t)$, defined in* (43) *is the unique optimal solution.*

The result of Theorem 4 shows that optimal price paths can be characterized in closed form even in the presence of multiple alternatives in the market. This robustness property also guarantees that the main properties of the optimal price solution in the single product case carry over to the multi-product case. In particular, by looking at (43), we notice that the optimal price path for product k is strictly increasing over time if $\mu_{k0} < \mu_k$, and it is strictly decreasing when $\mu_{k0} > \mu_k$. This observation highlights the presence of the same structural interdependence between the optimal pricing policy and the review-driven information diffusion process, mirroring the one described in the single-product case in Section 3.2.

The proof of Theorem 4 follows a natural generalization of the proof of Theorem 1 (the single product case). The main difficulty in establishing this result is to provide the conditions under which (44) has a unique solution, which, in turn, determines whether our candidate solution $P^*(t)$ is the unique optimal solution to (42). Indeed, the monotonicity properties exploited in (14) of the single product case fail to apply in the multi-product case, and, as a result, we need a more technically involved analysis, which finally results in the condition stated in (45). This condition may seem too restrictive and difficult to interpret. However, as the following proposition establishes, we can provide an approximation for Z^* in the case of large values of T. This approximation intuitively relates Z^* with the solution $Z^\infty := (z_1^\infty,\ldots,z_K^\infty)$ of (44) in the limit of $T \to \infty$, namely,

$$z_k^\infty + \alpha_k \sum_{i=1}^{K}\frac{\exp(z_i^\infty)}{\alpha_i} = \beta_k + \mu_k - 1, \quad k = 1,\ldots,K. \tag{46}$$

PROPOSITION 1 **(Rate of convergence to steady-state)**. *Let Z^* be any solution of* (44)*. Then Z^* satisfies $Z^* = Z^\infty + O(T^{-2})$, where Z^∞ is the unique solution of* (46)*.*

The preceding proposition suggests a heuristic method to find a near-optimal solution to the multi-product problem by replacing Z^* by Z^∞ in (43). The corresponding price must be close to the optimal price, since $p_k^*(t)$ is a continuous function of Z^*, and Z^∞ is close to Z^* for sufficiently large T by Proposition 1.

7. Concluding Remarks

We have shown how the revenue maximization problem with product reviews can be analyzed using a fluid approximation. By analyzing the fluid version of the basic problem, we were able to obtain structural insights into near-optimal policies, which are leveraged to design a well-performing pricing policy. We have found that the proposed policy is not only asymptotically near-optimal with respect to the underlying revenue maximization problem, but also practically implementable with almost negligible computational cost, thanks to the fact that price under this policy can be expressed in a closed form.

Most importantly, our results highlight the value of dynamic pricing: simple fixed-price policies are strictly revenue-suboptimal in certain parameter regimes, while the proposed dynamic policy achieves near-optimal revenue by judiciously balancing actual earnings from customers and benefits due to information from product reviews. This is encouraging since ever-growing e-commerce systems make the logistics of dynamic pricing much easier. In particular, pricing managers can now collect valuable information, including product reviews, and process it in real time, which in turn allows them to implement dynamic pricing strategies more easily and effectively.

Appendix A: Additional Results

A.1. Robustness of the Model

In this section we demonstrate the robustness of our model validation in Section 5.1. For this purpose we use two variables that might be relevant to sales on Amazon: (a) $\text{NYT}_m = 1$ if a product is listed in the New York Times Best Seller list on day m and zero otherwise; (b) ThirdPartyPrice_m that represents the lowest price from third-party sellers. As suggested by Sorensen (2004), the majority of book buyers use the New York Times Best Seller list as a signal of what is worth reading. In our data, 11 out of 72 products had appeared in the New York Times Best Seller list at least once between January 2015 and December 2016. Also, third parties can sell the same book at their own prices and the sales through those sellers are counted toward sales rank. In addition to the benchmark models in Section 5.1.3, we consider an extension of our base model as follows:

$$\lambda_d(p,n,r,X) = \ell\left(\frac{\gamma nr + \mu_0}{\gamma n + 1} + \beta - \alpha p + \Gamma X\right), \tag{47}$$

where X denotes the vector of NYT_m and ThirdPartyPrice_m.

Note that in Section 5.1 we assumed that the number of visits (NumVisits) is constant over time. In this section we estimate NumVisits_m as a function of the number of days since release (m) in order to ensure that we are not confounding our review measures with a simple time trend. To be specific, we first estimate the coefficients (ρ_1, ρ_2) of the equation:

$$\text{ImpliedSales}_m = \rho_1 \exp(-\rho_2 m) + \epsilon_m \tag{48}$$

using the least squares method. In (48), ϵ_m represents the portion of sales that does not explicitly depend on time. Then we estimate NumVisits_m as

$$\text{NumVisits}_m = \frac{\rho_1 \exp(-\rho_2 m)}{\text{CR}}, \tag{49}$$

Table 4 Predictive power (correlation) with time-dependent estimation model. The values in the parenthesis represent the standard deviations across 72 products.

	Conversion rate (CR)				
	1%	2%	3%	4%	5%
Base+NYT+3rd price	0.510 (0.380)	0.560 (0.361)	0.591 (0.359)	0.616 (0.332)	0.639 (0.313)
Base+NYT	0.674 (0.353)	0.697 (0.342)	0.704 (0.338)	0.703 (0.354)	0.700 (0.366)
Base	0.682 (0.350)	0.694 (0.343)	0.704 (0.338)	0.702 (0.357)	0.698 (0.365)
Linear	0.547 (0.320)	0.603 (0.288)	0.632 (0.288)	0.649 (0.295)	0.682 (0.287)
Logistic price-only	0.431 (0.498)	0.430 (0.497)	0.431 (0.497)	0.433 (0.498)	0.434 (0.498)
Linear price-only	0.382 (0.546)	0.388 (0.550)	0.388 (0.551)	0.386 (0.550)	0.388 (0.551)

where we vary the overall conversion rate $\text{CR} \in \{1\%, \ldots, 5\%\}$. Then, given ImpliedSales_m and NumVisits_m, we use the maximum likelihood method to estimate the parameters for demand λ_m:

$$\text{ImpliedSales}_m \sim \text{Binomial}(\text{NumVisits}_m, \lambda_m). \tag{50}$$

Table 4 illustrates out-of-sample correlations using the time-dependent estimation model described above. As in Table 2, the base model outperforms the three benchmark models discussed in Section 5.1.3. The base model with the additional covariate NYT_m does not improve the predictive power of the base model. Further, the predictive power is significantly reduced under the base model with NYT_m and ThirdPartyPrice_m, which might be due to overfitting. This demonstrates that enriching our model with more information does not necessarily improve the predictive power with respect to what established in Section 5.1.

A.2. Simulation for Large-scale Problems

In small-scale problems, the DFM policy significantly outperforms the other proposed policies, SFM, OSLA, and OFPS, which is illustrated in Figure 5. However, all of these policies have an optimality gap of order $\sqrt{T}$ asymptotically as $T \to \infty$ (Theorem 3). To illustrate the asymptotic performance of these policies, we perform numerical experiments in the same settings as those in Figure 5, but with a longer sales horizon $T \in [20, 1280]$. Note that if $J_T^\pi / T \approx b$ for some $b > 0$, then the revenue is approximately linear in T with slope b. Figure 7 illustrates that J_T^π / T approaches a constant b under the proposed policies. Each point in this figure is obtained by averaging the revenues over 10^3 simulation trials. In this figure, (percentage) standard errors of estimated revenues are less than 3%.

Appendix B: Proofs

B.1. Proofs of the Main Results

Proof of Theorem 1. To show existence, note first that there is one-to-one relationship between the price and the demand function for any state variables. In particular, fix the state variables $n(t) = n$ and $r(t) = \mu$.

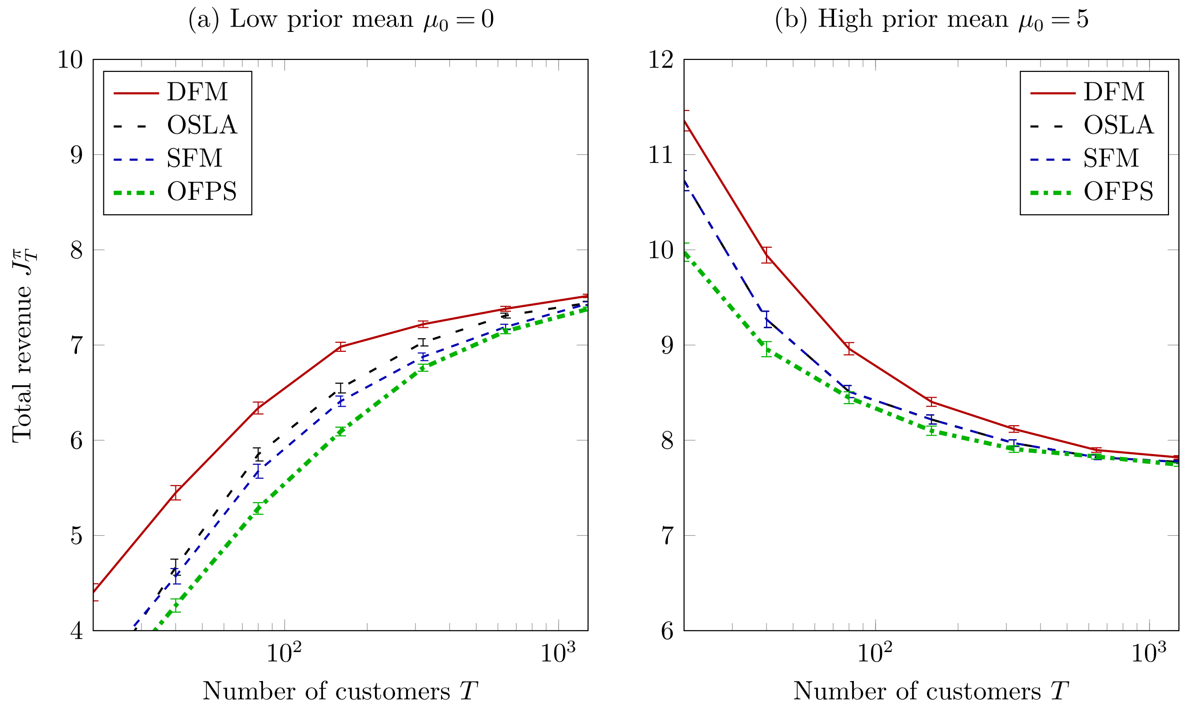

Figure 7 The total revenues under different pricing policies for low and high prior beliefs: (a) $\mu_0 = 0 < 2.5 = \mu$ and (b) $\mu_0 = 5 > 2.5 = \mu$. In both cases, $(\alpha, \beta, \gamma, u, \sigma_q) = (0.1, -1, 1, 1, 2)$. The revenues of these policies are estimated by averaging over 10^3 simulation trials. The error bars represent standard errors.

Given a price $p(t) = p$, the demand is determined by $\lambda(p, n, r) = \ell(\beta + \mu - \alpha p)$. Conversely, we can construct a mapping from demand λ to price $p(\lambda, n, r)$, where

$$p(\lambda, n, r) = \frac{\beta + \mu + \log((1-\lambda)/\lambda)}{\alpha}. \tag{51}$$

Therefore, the fluid optimization problem in (10) is equivalent to

$$\begin{aligned} \max_{\lambda(\cdot)} \quad & \int_0^T p(\lambda(s), n(s), r(s))\lambda(s) ds \\ \text{subject to} \quad & \dot{n}(t) = u\lambda(t), t \in [0, T]. \end{aligned} \tag{52}$$

In the reformulated fluid problem (52), we can replace $\max_{\lambda(\cdot)}$ with $\max_{\lambda(\cdot) \in [\underline{\lambda}, \bar{\lambda}]}$, for which the corresponding price satisfies $p(\lambda(\cdot), n(\cdot), r(\cdot)) \in [\underline{p}, \bar{p}]$ for any admissible values of $n(\cdot)$ and $r(\cdot)$. Obviously, $[\underline{\lambda}, \bar{\lambda}]$ is independent of state variables and t, and therefore, is an upper semicontinuous function of $(t, n(t), r(t))$. Further, let

$$U(t, n, r) = \{(z^0, z) \mid z^0 \geq \lambda p(\lambda, n, r), z = \lambda, \lambda \in [\underline{\lambda}, \bar{\lambda}]\}. \tag{53}$$

Note that $\lambda p(\lambda, n, r)$ is a concave function of λ, and therefore, $U(t, n, r)$ is a convex set for each (t, n, r). By Theorem 1 of Cesari (1966), there exists an optimal solution to the refomulated fluid problem (52). This in turn implies the existence of the optimal solution to the fluid problem (10) thanks to the one-to-one mapping between the price and demand.

To show uniqueness, it suffices to show that the solution to (14) is unique under the condition (15). Rewrite (14) as

$$g(z) := z + \exp(z) - \mu - \beta + 1 - \frac{\mu_0 - \mu}{\gamma u \ell(z) T + 1}. \tag{54}$$

First, if $\mu_0 \geq \mu$, it is straightforward to check that $g'(z) > 0$ for all z, so the solution to the preceding equation is unique. In the remaining part of the proof we focus on the case with $\mu_0 < \mu$. Observe that

$$g'(z) = 1 + \exp(z) - \frac{(\mu - \mu_0)(1 - \ell(z))\gamma u \ell(z) T}{(\gamma u \ell(z) T + 1)^2}. \tag{55}$$

We now show that, if (15) is satisfied, then $g'(z) > 0$ for all z. Since $1 + \exp(z) \geq 1$ for all z, it suffices to show that (15) ensures

$$h(z) := \frac{(\mu - \mu_0)(1 - \ell(z))\gamma u \ell(z) T}{(\gamma u \ell(z) T + 1)^2} < 1. \tag{56}$$

It is easy to check that $h(z)$ has a global maximum for $z = \tilde{z}$ such that $\ell(\tilde{z}) = 1/(2 + \gamma u T)$, from which we deduce that

$$h(z) \leq \frac{\gamma u T(\mu - \mu_0)}{4(1 + \gamma u T)}. \tag{57}$$

Hence, (15) guarantees $g'(z) > 0$. Combined with the fact that $g(z) \to -\infty$ as $z \to -\infty$ and $g(z) \to \infty$ as $z \to \infty$, we establish that there is a unique solution to (54). $\square$

Proof for Theorem 2. We first show the result for the slow-learning regime where $\gamma T \to 0$. Without loss of generality, we fix T and vary γ. Denote by p^{OFP} the optimal fixed price which depends on γ and T. Note from (13) that as $\gamma \to 0$,

$$p^*(t) \to \frac{1 + \exp(z_0^*)}{\alpha}, \tag{58}$$

where the constant z_0^* satisfies

$$z_0^* + \exp(z_0^*) = \mu_0 + \beta - 1. \tag{59}$$

Note that the limit in (58) is independent of t. Also, as established in Theorem 1, this is the unique optimal solution, and hence, p^{OFP} should satisfy $\lim_{\gamma \to 0} p^{\mathrm{OFP}} = (1 + \exp(z_0^*))/\alpha$ for $t \in [0, T]$. Therefore, we conclude that $1 - \bar{J}_T^{\mathrm{OFP}}/\bar{J}_T^* \to 0$ as $\gamma \to 0$. Using the same logical steps we can show that $1 - \bar{J}_T^{\mathrm{OFP}}/\bar{J}_T^* \to 0$ as $\gamma \to \infty$, which will be omitted.

Next, we show that $1 - \bar{J}_T^{\mathrm{OFP}}/\bar{J}_T^* \to \Delta(\kappa)$ as $\gamma T \to \kappa \in (0, \infty)$. Observe that as $\gamma T \to \kappa$, the equation in (14) becomes

$$z + \exp(z) = \mu + \beta - 1 + \frac{\mu_0 - \mu}{\kappa u \ell(z) + 1}, \tag{60}$$

which depends on γ and T only through κ. Let z_κ be the constant that satisfies the preceding equation, which is unique by Theorem 1. Note from (19) that

$$\frac{\bar{J}_T^*}{T} \to \frac{\exp(z_\kappa)}{\alpha} + \frac{\mu_0 - \mu}{\alpha u \kappa}\left(\log(u\kappa\ell(z_\kappa) + 1) - \frac{u\kappa\ell(z_\kappa)}{u\kappa\ell(z_\kappa) + 1}\right) \quad \text{as } \gamma T \to \kappa. \tag{61}$$

Note that the limit in (61) is invariant to γ and T. Hence, it remains to show that the limit of $\bar{J}_T^{\text{OFP}}/T$ as $\gamma T \to \kappa$ depends on κ but invariant to γ and T. Consider a fixed-price policy π with $p^\pi(t) = p$. Observe that

$$\begin{aligned} n^\pi(T) &= \int_0^T \dot{n}^\pi(t)dt \\ &= u\int_0^T \frac{\exp\left(\beta - \alpha p + \mu - \frac{\mu-\mu_0}{1+\gamma n^\pi(t)}\right)}{1+\exp\left(\beta - \alpha p + \mu - \frac{\mu-\mu_0}{1+\gamma n^\pi(t)}\right)}dt \\ &= uT - \int_0^T \frac{1}{1+\exp\left(\beta - \alpha p + \mu - \frac{\mu-\mu_0}{1+\gamma n^\pi(t)}\right)}dt \\ &= uT - \int_0^T \dot{n}^\pi(t)\exp\left(-\beta + \alpha p - \mu + \frac{\mu-\mu_0}{1+\gamma n^\pi(t)}\right)dt \\ &= uT - \exp(-\beta + \alpha p - \mu)\int_0^{n^\pi} \exp\left(\frac{\mu-\mu_0}{1+\gamma x}\right)dx, \end{aligned} \tag{62}$$

where the second and fourth equations follow from (11c), while the last equality is obtained from a straightforward change of variable in the integration. Hence, $n^\pi(T)$ is solution n to the following equation:

$$n - uT + \exp(-\beta + \alpha p - \mu)\int_0^n \exp\left(\frac{\mu-\mu_0}{1+\gamma x}\right)dx = 0. \tag{63}$$

Let $\tilde{n}^\pi(T) = \gamma n^\pi(T)$ and $\tilde{n} = \gamma n$. Then, it can be easily seen that as $\gamma T \to \kappa$, $\tilde{n}^\pi(T)$ approaches the solution $\tilde{n}$ of the following equation:

$$\tilde{n} - u\kappa + \exp(-\beta + \alpha p - \mu)\int_0^{\tilde{n}} \exp\left(\frac{\mu-\mu_0}{1+x}\right)dx = 0. \tag{64}$$

Since the preceding equation is invariant to γ and T, we conclude that $\tilde{n}^\pi(T)$ is invariant to γ and T. Further, for any fixed-price policy π with price p, we have that

$$\begin{aligned} \frac{\bar{J}_T^\pi}{T} &= \frac{pn^\pi(T)}{uT} \\ &= \frac{p\gamma n^\pi(T)}{\gamma uT} \\ &\to \frac{p\Delta'(\kappa)}{u\kappa} \text{ as } \gamma T \to \kappa, \end{aligned} \tag{65}$$

where the first equation follows from the fact that the total number of buyers is equal to the number of reviews divided by u, the probability of posting a review after a purchase. Instead, the last equation follows from the previous observation that $\tilde{n}^\pi(T) = \Delta'(\kappa)$, sincewe know that the constant $\Delta'(\kappa) = \gamma n(T)$ depends on κ but is invariant to γ and T. Since these arguments hold for any fixed-price policy, it also holds for OFP. Combining (61) and (65), the desired result in the theorem follows. □

We need the following lemmas to prove Lemma 1 and Theorem 3. The proofs for auxiliary lemmas are proved in Appendix B.2.

LEMMA 4. *Suppose that customer t posts a review with probability $f(n_t)$, where $f(x) = ux^{-v}$. For any $\pi \in \Pi$ and $t \geq 1$, there exists positive constants $M_1, M_2 < \infty$ such that*

$$\mathsf{E}^\pi\left[(r_t - \mu)^2\right] \begin{cases} \leq \frac{M_1}{t^{1-v}+1} & \text{if } v \in [0,1] \\ \geq M_2 & \text{if } v \in (1,\infty), \end{cases} \tag{66}$$

where r_t is the average rating in stage t induced by the policy π.

LEMMA 5. *For any $\pi \in \Pi$ and $t \geq 1$,*

$$\mathsf{E}^{\pi}|\lambda(p, n^*(t), r^*(t)) - \lambda(p, n^{\pi}_{\lfloor t \rfloor}, n_{\lfloor t \rfloor})| = \mathcal{O}\left(\frac{1}{\sqrt{t}}\right) \tag{67}$$

for any $p \in [\underline{p}, \bar{p}]$, where $n^*(t)$ and $r^*(t)$ are state variables under the optimal solution ψ^*.

Proof of Lemma 1. Recall the definitions $\pi^* = \{p^*_t \,|\, t = 0, \ldots, T-1\}$ and $\psi^* = \{p^*(t) \,|\, t \in [0, T]\}$, the maximizers for J^{π}_T and $\bar{J}^{\psi}_T$, respectively. Also, define $\hat{\pi} = \{\hat{p}_t = p^*(t) \,|\, t = 0, \ldots, T-1\} \in \Pi$ as the discrete pricing policy induced by the ODP policy ψ^*. Likewise, define $\hat{\psi} = \{\hat{p}(t) = p^*_{\lfloor t \rfloor} \,|\, t \in [0, T]\}$ as the continuous pricing policy induced by π^*, where $\lfloor t \rfloor$ is the largest integer less than or equal to t. We define $L > 0$ as the Lipschitz constant for the logistic function $\ell(x) = \exp(x)/(1 + \exp(x))$. Lastly, for any $\pi \in \Pi$, it will be useful to note that

$$J^{\pi}_T = \sum_{t=1}^{T} \mathsf{E}^{\pi}[p_t \lambda(p_t, n_t, r_t)] = \int_0^T \mathsf{E}^{\pi}[p_{\lfloor t \rfloor} \lambda(p_{\lfloor t \rfloor}, n_{\lfloor t \rfloor}, r_{\lfloor t \rfloor})] dt. \tag{68}$$

Since $\bar{J}^*_T \geq \mathsf{E}^{\hat{\psi}}[\bar{J}^{\hat{\psi}}_T]$ and $J^*_T \geq J^{\hat{\pi}}_T$, observe that

$$|\bar{J}^*_T - J^*_T| \leq \max\{|\bar{J}^*_T - J^{\hat{\pi}}_T|, |\mathsf{E}^{\hat{\psi}}[\bar{J}^{\hat{\psi}}_T] - J^*_T|\}. \tag{69}$$

Therefore, it suffices to show that the two terms on the right-hand side of (69) are $\mathcal{O}(\sqrt{T})$.

Step 1. To bound the first term of (69), $|\bar{J}^*_T - J^{\hat{\pi}}_T|$, let $\{(\hat{n}_t, \hat{r}_t) \,|\, t = 0, \ldots, T-1\}$ be the path of the state variables driven by $\hat{\pi}$ and observe that

$$\begin{aligned}
|\bar{J}^*_T - J^{\hat{\pi}}_T| &= \left| \int_0^T p^*(t) \lambda(p^*(t), n^*(t), r^*(t)) dt - \int_0^T \mathsf{E}^{\hat{\pi}} \left[\hat{p}_{\lfloor t \rfloor} \lambda(\hat{p}_{\lfloor t \rfloor}, \hat{n}_{\lfloor t \rfloor}, \hat{r}_{\lfloor t \rfloor}) \right] dt \right| \\
&\leq \int_0^T \mathsf{E}^{\hat{\pi}} \left[|p^*(t) \lambda(p^*(t), n^*(t), r^*(t)) - \hat{p}_{\lfloor t \rfloor} \lambda(\hat{p}_{\lfloor t \rfloor}, n^*(t), r^*(t))| \right] dt \\
&\quad + \int_0^T \mathsf{E}^{\hat{\pi}} \left[|\hat{p}_{\lfloor t \rfloor} \lambda(\hat{p}_{\lfloor t \rfloor}, n^*(t), r^*(t)) - \hat{p}_{\lfloor t \rfloor} \lambda(\hat{p}_{\lfloor t \rfloor}, \hat{n}_{\lfloor t \rfloor}, \hat{r}_{\lfloor t \rfloor})| \right] dt \\
&\doteq U_T + V_T
\end{aligned} \tag{70}$$

Observe from (13) that $n^*(t)$ is an affine function of $t \in [\underline{\tau}, \bar{\tau})$, and hence, the reciprocal of $p^*(t)$ in (13) is affine with respect to t. Therefore, it is not difficult to see that there exists $C_1 > 0$ such that, for any $t \in [0, T]$,

$$|p^*(t) - \hat{p}_{\lfloor t \rfloor}| \leq |p^*(\lfloor t \rfloor) - p^*(\lfloor t+1 \rfloor)| \leq \frac{C_1}{(t+1)^2}. \tag{71}$$

A triangular inequality, combined with the fact that $\lambda(p, n, s)$ is Lipschitz continuous in p with modulus L, gives

$$\left| p^*(t) \lambda(p^*(t), n^*(t), r^*(t)) - \hat{p}_{\lfloor t \rfloor} \lambda(\hat{p}_{\lfloor t \rfloor}, n^*(t), r^*(t)) \right| \leq \frac{LC_1}{(t+1)^2}. \tag{72}$$

Since $\int_0^\infty 1/(t+1)^2 dt < \infty$, we establish that U_T is bounded by a constant that does not depend on T. Further, $V_T = \mathcal{O}(\sqrt{T})$ follows from Lemma 5. Combining these observations, we establish that $|\bar{J}^*_T - J^{\hat{\pi}}_T| = \mathcal{O}(\sqrt{T})$.

Step 2. Next, to bound the second term of (69), $|\mathsf{E}[\bar{J}^{\hat{\psi}}_T] - J^*_T|$, recall that $\hat{\psi} = \{\hat{p}(t) = p^*_{\lfloor t \rfloor} \,|\, t \in [0, T]\}$ and $\hat{p}(t)$ is adopted to $\mathcal{F}_{\lfloor t \rfloor}$, the filtration generated by the past observations along the optimal price π^* and corresponding state variables up to stage $\lfloor t \rfloor$. Define $\hat{n}(t)$ and $\hat{r}(t)$ as the state variables for the fluid problem induced by $\hat{\psi}$, which are also adopted to $\mathcal{F}_{\lfloor t \rfloor}$. From the fact that $\hat{p}(t) = p^*_{\lfloor t \rfloor}$, one can write

$$\left| \mathsf{E}^{\hat{\psi}}[\bar{J}^{\hat{\psi}}_T] - J^*_T \right| \leq \int_0^T \mathsf{E}^{\pi^*} \left[p^*_{\lfloor t \rfloor} \left| \lambda(p^*_{\lfloor t \rfloor}, \hat{n}(t), \hat{r}(t)) - \lambda(p^*_{\lfloor t \rfloor}, n^*_{\lfloor t \rfloor}, r^*_{\lfloor t \rfloor}) \right| \right] dt. \tag{73}$$

From Lemma 5 the integrand on the right-hand side of the preceding inequality is $\mathcal{O}(1/\sqrt{t})$, from which we conclude that $|\mathsf{E}^{\hat{\psi}}[\bar{J}^{\hat{\psi}}_T] - J^*_T| = \mathcal{O}(\sqrt{T})$. $\square$

Proof of Lemma 2. Using the triangular inequality, It is straightforward to check that $|\bar{J}_T^* - J_T^\pi| \le \Delta_T^p + \Delta_T^s$, where

$$\begin{aligned} \Delta_T^p &= \int_0^T \mathsf{E}\left|p^*(t)\lambda(p^*(t),n^*(t),r^*(t)) - p_{[t]}^\pi\lambda(p_{[t]}^\pi,n^*(t),r^*(t))\right| dt \\ \Delta_T^s &= \int_0^T \mathsf{E}\left|p_{\lfloor t\rfloor}^\pi\lambda(p_{\lfloor t\rfloor}^\pi,n^*(t),r^*(t)) - p_{\lfloor t\rfloor}^\pi\lambda(p_{\lfloor t\rfloor}^\pi,n_{\lfloor t\rfloor}^\pi,r_{\lfloor t\rfloor}^\pi)\right| dt, \end{aligned} \tag{74}$$

where $\lfloor t\rfloor$ is the largest integer less than or equal to t. It is immediate to see that $\Delta_T^s = \mathcal{O}(\sqrt{T})$ from Lemma 5. Hence, it remains to show that $\Delta_T^p = \mathcal{O}(\sqrt{T})$. To this end, observe that the integrand of Δ_T^p is bounded by $L\mathsf{E}^\pi|p^*(t) - p_{[t]}^\pi|$ by the triangular inequality and the fact that $\lambda(p,n,q)$ is Lipschitz continuous with respect to p with modulus L. Hence, the proof of the theorem is complete if we show that $\mathsf{E}^\pi|p^*(t) - p_{\lfloor t\rfloor}^\pi| = \mathcal{O}(1/\sqrt{t})$ for $\pi \in \{\text{OFPS}, \text{OSLA}, \text{SFM}, \text{DFM}\}$. For clarity of exposition, in the reminder of the proof we assume that t is an integer value.

For $\pi = \text{OFP}$, observe that for any $p \in [\underline{p}, \bar{p}]$

$$\mathsf{E}^\pi|\lambda(p,n_t^\pi,r_t^\pi) - \lambda(p,n^*(t),r^*(t))| = \mathcal{O}\left(\frac{1}{\sqrt{t}}\right) \tag{75}$$

by Lemma 5. This implies

$$\mathsf{E}^\pi\left|\frac{1}{T}\sum_{t=1}^T \lambda(p,n_t^\pi,r_t^\pi) - \lambda(p,n^*(t),r^*(t))\right| = \mathcal{O}\left(\frac{1}{\sqrt{T}}\right). \tag{76}$$

Since $|\lambda(p,n^*(t),r^*(t)) - \lambda(p,\infty,\mu)| = \mathcal{O}(1/t)$, the preceding relation implies

$$\mathsf{E}^\pi\left|\frac{1}{T}\sum_{t=1}^T \lambda(p,n_t^\pi,r_t^\pi) - \lambda(p,\infty,\mu)\right| = \mathcal{O}\left(\frac{1}{\sqrt{T}}\right). \tag{77}$$

Applying proposition 6.1 of Bonnans and Shapiro (1998), we establish that the fixed price $\hat{p}_T = \arg\max_p\{\mathsf{E}[\sum_{t=0}^{T-1} p\lambda(p,n_t,r_t)] \mid p \in [\underline{p},\bar{p}]\}$ satisfies $|\hat{p} - p_\infty| = \mathcal{O}(1/\sqrt{T})$, where $p_\infty := \arg\max_p\{p\lambda(p,\infty,\mu) \mid p \in [\underline{p},\bar{p}]\}$. Further, note that $\theta(T) = 0$, and hence, $p^*(T) = \arg\max_p\{p\lambda(p,n^*(T),r^*(T))\}$ from (11a). Observing that $|\lambda(p,\infty,\mu) - \lambda(p,n^*(T),r^*(T))| = \mathcal{O}(1/T)$, and by Proposition 6.1 of Bonnans and Shapiro (1998), we have that $|p_\infty - p^*(T)| = \mathcal{O}(1/T)$. Lastly, from (13) we have that

$$|p^*(T) - p^*(t)| = \frac{|\mu - \mu_0|}{\alpha}\left(\frac{1}{\gamma n^*(T)+1} - \frac{1}{\gamma n^*(t)+1}\right) = \mathcal{O}\left(\frac{1}{t}\right). \tag{78}$$

Combining these results with the triangle inequality, we establish that $|\hat{p} - p^*(t)| \le |\hat{p} - p_\infty| + |p_\infty - p^*(T)| + |p^*(T) - p^*(t)| = \mathcal{O}(1/\sqrt{t})$.

For $\pi = \text{OSLA}$, observe that

$$\begin{aligned} |\lambda(p,n_t^\pi,r_t^\pi) - \lambda(p,\infty,\mu)| &\le L\left|\frac{\gamma\hat{n}_{\lfloor t\rfloor}\hat{r}_{\lfloor t\rfloor} + \mu_0}{\gamma\hat{n}_{\lfloor t\rfloor}+1} - \mu\right| \\ &\le L\left(|r_t^\pi - \mu| + \left|\frac{\mu_0 - \mu}{\gamma n_t^\pi + 1}\right|\right). \end{aligned} \tag{79}$$

From proposition 6.1 of Bonnans and Shapiro (1998), we have that

$$|p_t^\pi - p_\infty| = \mathcal{O}\left(|r_t^\pi - \mu| + \left|\frac{\mu_0 - \mu}{\gamma n_t^\pi + 1}\right|\right). \tag{80}$$

Since $\mathsf{E}^\pi|r_t^\pi - \mu| = \mathcal{O}(1/\sqrt{t})$ by Lemma 4 and $\mathsf{E}^\pi|(\mu_0 - \mu)/(\gamma n_t^\pi + 1)| = \mathcal{O}(1/t)$, we establish that $\mathsf{E}^\pi|p_t^\pi - p_\infty| = \mathcal{O}(1/\sqrt{t})$. Further, as shown in the case with $\pi = \text{OFPS}$, we have that $|p_\infty - p^*(t)| = \mathcal{O}(1/t)$, from which we conclude that $\mathsf{E}^\pi|p_t^\pi - p^*(t)| = \mathcal{O}(1/\sqrt{t})$.

For $\pi = \text{SFM}$, we have $p_t^\pi = p^*(t)$ by definition. Now, fix $\pi = \text{DFM}$. Recalling the ODP policy in (13)-(14), and after some straightforward algebra, it can be seen that, for any integer t,

$$\mathsf{E}^\pi|p^*(t) - p_t^\pi| \le \frac{1}{\alpha}\mathsf{E}^\pi\left|\exp(z^*) - \exp(z_t^*)\right| + \frac{1}{\alpha}\mathsf{E}^\pi\left|\frac{\gamma n^*(t)\mu + \mu_0}{\gamma n^*(t) + 1} - \frac{\gamma n_t^\pi r_t^\pi + \mu_0}{\gamma n_t^\pi + 1}\right|, \tag{81}$$

where z^* and z_t^* are defined in (14) and (27), respectively. Observe that

$$\mathsf{E}^\pi\left|\frac{\gamma n^*(t)\mu + \mu_0}{\gamma n^*(t) + 1} - \frac{\gamma n_t^\pi r_t^\pi + \mu_0}{\gamma n_t^\pi + 1}\right| \le \left|\frac{\mu_0 - \mu}{\gamma n^*(t) + 1}\right| + \mathsf{E}^\pi\left|\frac{\gamma n_t^\pi r_t^\pi + \mu_0}{\gamma n_t^\pi + 1} - \mu\right| = \mathcal{O}\left(\frac{1}{\sqrt{t}}\right), \tag{82}$$

where the last equation follows from (114) and (115). Further, observe from (14) and (27) that

$$\begin{aligned}
\mathsf{E}^\pi|\exp(z^*) - \exp(z_t^*)| &\le \mathsf{E}^\pi|z^* + \exp(z^*) - z_t^* - \exp(z_t^*)| \\
&\le \mathsf{E}^\pi\left|\frac{\mu_0 - \mu}{\gamma u\ell(z^*)T + 1} - \frac{\gamma n_t^\pi(r_t^\pi - \mu) + \mu_0 - \mu}{\gamma(n_t^\pi + u\ell(z_t^*)(T - t)) + 1}\right| \\
&\le \left|\frac{\mu_0 - \mu}{\gamma u\ell(z^*)T + 1}\right| + \mathsf{E}^\pi|r_t^\pi - \mu| + \mathsf{E}^\pi\left|\frac{\mu_0 - \mu}{\gamma(n_t^\pi + u\ell(z_t^*)(T - t)) + 1}\right|,
\end{aligned} \tag{83}$$

where the first term is $\mathcal{O}(1/T)$, the second term is $\mathcal{O}(1/\sqrt{t})$ by Lemma 4, and the last term is $\mathcal{O}(1/t)$ by theorem 1 of Chao and Strawderman (1972). Therefore, we establish that $\mathsf{E}^\pi|p_t^\pi - p^*(t)| = \mathcal{O}(1/\sqrt{t})$. This completes the proof of the theorem. $\square$

Proof of Corollary 2. Observe that the objective function of the stochastic problem (7) can be written as

$$\begin{aligned}
J_T^\pi &= \sum_{t=0}^{T-1}\mathsf{E}^\pi[p_t\lambda(p_t, n_t, r_t)] \\
&= \int_0^T \mathsf{E}^\pi[p_{\lfloor t\rfloor}\lambda(p_{\lfloor t\rfloor}, n_{\lfloor t\rfloor}, r_{\lfloor t\rfloor})]dt \\
&= \int_0^{G(T)} \mathsf{E}^\pi[p_{\lfloor \tau\rfloor}\lambda(p_{\lfloor \tau\rfloor}, n_{\lfloor \tau\rfloor}, r_{\lfloor \tau\rfloor})]d\tau.
\end{aligned} \tag{84}$$

The remaining steps of the proof are exactly identical to the proof of Theorem 1, except for the minor modifications due to the different time scale. The full proof will thus be omitted. $\square$

Proof of Lemma 3. The proof of the lemma is based on two steps. In *Step 1* we prove the bounds for $l_k(Z^*)$. In *Step 2* we use the Banach fixed-point Theorem to establish the uniqueness of Z^* under the condition (45).

Step 1. Notice that, since $\min(0, \mu_k - \mu_{k0}) < \frac{\mu_k - \mu_{k0}}{1 + c_k T l_k(Z^*)} < \max(0, \mu_k - \mu_{k0})$ for any $k = 1, \ldots, K$, any component z_k^* of Z^* can be bounded by above by $z_k^* \le \overline{z}_k := \beta_k - 1 + \max(\mu_{k,0}, \mu_k)$. Moreover, observe that $\sum_{i=1}^K \frac{1}{\alpha_i}\exp(z_i^*) > 0$ and that $\mu_k - \max(0, \mu_k - \mu_{k0}) = \max(\mu_{k,0}, \mu_k)$. Using this result, and using $\mu_k - \min(0, \mu_k - \mu_{k0}) = \min(\mu_{k,0}, \mu_k)$, it is easy to see that $z_k^* \ge \underline{z}_k := \beta_k - 1 - \alpha_k\sum_{i=1}^K \frac{1}{\alpha_i}\exp(\overline{z}_i) + \min(\mu_{k,0}, \mu_k)$. Finally, the bounds for $l_k(Z^*)$ follow from noticing that $l_k(Z)$ is strictly increasing in z_k and strictly decreasing in z_i for $i \ne k$.

Step 2. For this step of the proof, we rewrite (44) as $Z = F(Z)$, where $F : \mathcal{R}^K \to \mathcal{R}^K$ is defined by $F(Z) := (F_1(Z), \ldots, F_K(Z))$ and by

$$F_k(Z) := \beta_k + \mu_k - 1 - \alpha_k\sum_{i=1}^K \frac{\exp(z_i)}{\alpha_i} - \frac{\mu_k - \mu_{k0}}{\gamma_k u\ell_k(Z)T + 1}. \tag{85}$$

As a result, a solution Z^* of (44) is a fixed point of the vector function F. Observe that we can use the bounds for Z^* proved in the first step of the present proof to establish that $F_k(Z) \in C^\infty$ over the compact convex set $\mathcal{Z}_K := [\underline{z}_1, \overline{z}_1] \times [\underline{z}_2, \overline{z}_2] \times \ldots [\underline{z}_K, \overline{z}_K]$ for all k. By the Banach fixed point theorem, this ensures that F has at least a fixed point. Hence, in the remainder of this step of the proof we assume that (45) is satisfied, and we show that, under this assumption F is a contraction mapping under the ℓ_1 norm, which, in turn, implies that F has a unique fixed point.

By the mean value theorem we know that, for any $Z, Z' \in \mathcal{Z}_K$,

$$||F(Z) - F(Z')||_1 \leq \max_{\delta \in [0,1]} ||\mathbf{J}_F(Z' + \delta(Z - Z'))||_1 \cdot ||Z - Z'||_1, \tag{86}$$

where $\mathbf{J}_F(Z' + \delta(Z - Z'))$ represents the Jacobian matrix of F evaluated in the point $Z' + \delta(Z - Z')$. Observe that, for $j \neq k$, we have

$$\begin{aligned} \frac{\partial F_j(Z)}{\partial z_k} &= -\frac{\alpha_j}{\alpha_k}\exp(z_k) - \frac{(\mu_j - \mu_{j0})\gamma_j\, uTl_j(Z)l_k(Z)}{(1 + \gamma_j\, ul_j(Z)T)^2} \\ \frac{\partial F_k(Z)}{\partial z_k} &= -\exp(z_k) + \frac{(\mu_k - \mu_{k0})\gamma_k\, uTl_k(Z)(1 - l_k(Z))}{(1 + \gamma_k\, ul_k(Z)T)^2} \end{aligned} \tag{87}$$

Hence, using the triangular inequality, we find that

$$\begin{aligned} \sum_{j=1}^{K} \left| \frac{\partial F_j(Z)}{\partial z_k} \right| \leq \frac{\exp(z_k)}{\alpha_k / \sum_{i=1}^{K} \alpha_i} + \sum_{j=1, j\neq k}^{K} |\mu_j - \mu_{j0}| \frac{\gamma_j\, uTl_j(Z)l_k(Z)}{(1 + \gamma_j\, uTl_j(Z))^2} \\ + |\mu_k - \mu_{k0}| \frac{\gamma_k\, uTl_k(Z)(1 - l_k(Z))}{(1 + \gamma_k\, uTl_k(Z))^2}. \end{aligned}$$

for all $k = 1, \ldots, K$. Using the above inequality, we can show that under (45) we have $\sum_{j=1}^{K} \left| \frac{\partial F_j(Z)}{\partial z_k} \right| < 1 - a$ for some $a > 0$ and for all $Z \in \mathcal{Z}_K$. In fact, it is easy to verify that, for $x \in [0,1]$, the function $g_k(x) := \frac{|\mu_k - \mu_{k0}|\gamma_k\, uTx(1-x)}{(1+\gamma_k\, uTx)^2}$ is strictly increasing for $0 \leq x < x^* := 1/(2 + \gamma_k\, uT)$ and strictly decreasing for $x^* < x \leq 1$, which implies that $g_k(x) \leq g_k(x^*) = \frac{|\mu_k - \mu_{k0}|}{4} \frac{\gamma_k\, uT}{1+\gamma_k\, uT}$. Moreover, since we have that

$$\frac{\gamma_j uTl_j(Z)}{(1 + \gamma_j\, uTl_j(Z))} < 1 \tag{88}$$

and $0 < \underline{l}_k < l_k(Z) < \bar{l}_k < 1$ for all j, we obtain

$$\sum_{j=1, j\neq k}^{K} \frac{|\mu_j - \mu_{j0}|\gamma_j\, uTl_j(Z)l_k(Z)}{(1 + \gamma_j\, uTl_j(Z))^2} < \sum_{j=1, j\neq k}^{K} \frac{|\mu_j - \mu_{j0}|}{1 + \gamma_j uT\underline{l}_j}. \tag{89}$$

This shows that $||\mathbf{J}_F(Z)||_1 < 1 - a$ for some $a > 0$ and for all $Z \in \mathcal{Z}_K$, which in turn implies that, under the assumption (45), F is a contraction mapping. This guarantees that, if (45) is satisfied, then F has a unique fixed point Z^* which is the unique solution of (44) and concludes the proof. $\square$

Proof of Theorem 4. We begin the proof with Pontryagin maximum principle to characterize the first order necessary conditions for the multi-product optimal pricing policy. To simplify notation, throughout the proof, we omit the dependence from t where there is no ambiguity. Moreover, given that in the fluid approximation the average rating vector R is constant over time, we also omit the dependence from R. The Hamiltonian function is defined $H(P, N, \Theta) := (P + u\Theta)\Lambda(P, N)$, where $\Theta := (\theta_1, \ldots, \theta_K)$ is the shadow price vector associated to the constraint $\dot{N} = u\Lambda(P,N,R)$. The Pontryagin maximum principle states that the optimal solution (P^*, N^*, Θ^*), if it exists, must satisfy

$$0 = \frac{\partial H(P^*, N^*, \Theta^*)}{\partial p_k} \qquad \text{(Maximum of } H\text{)} \tag{90a}$$

$$\dot{\theta}_k = -\frac{\partial H(P^*, N^*, \Theta^*)}{\partial n_k}, \;\; \theta_k(T) = 0 \qquad \text{(Shadow price)} \tag{90b}$$

$$\dot{n}_k^* = u\lambda_k(P^*, N^*), \;\; n_k^*(0) = 0 \qquad \text{(Number of reviews)} \tag{90c}$$

It is straightforward to verify that the candidate solution provided in (43) satisfies (90a)-(90c). Hence, to prove the theorem it suffices to show existence and uniqueness.

To show existence, we first establish that there is a one-to-one correspondence between the price and the demand function vector. In particular, we can adapt the proof in Li and Huh (2011) to show that, given a demand function vector $\Lambda = (\lambda_1, \dots, \lambda_K)$, prices can be obtained from

$$p_k(\Lambda, N) = \frac{1}{\alpha_k}\left[\beta_k + \frac{\gamma_k \mu_k n_k + \mu_{k0}}{\gamma_k n_k + 1} + \log(1 - \sum_{j=1}^{K} \lambda_k) - \log \lambda_k\right]. \tag{91}$$

Similar to the proof of Theorem 1, we can reformulate (42) as

$$\begin{aligned} \max_{\Lambda(\cdot)} \quad & \int_0^T P(\Lambda(s), N(s)) \cdot \Lambda(s) ds \\ \text{subject to} \quad & \dot{N}(t) = u\Lambda(t), \quad t \in [0, T]. \end{aligned} \tag{92}$$

Invoking Li and Huh (2011), we establish that $P(\Lambda(t), N(t)) \cdot \Lambda(t)$ is always concave in Λ, which guarantees that an optimal solution to the problem (92) always exists. Using (91), this also establishes existence of an optimal solution to the original problem (42).

To show uniqueness, we show that the candidate optimal solution in (43) is the unique solution to the FOC in (90a)-(90c). Using (90a)-(90c), we can prove that

$$\dot{\theta}_k(t)^* = \frac{\gamma_k(\mu_k - \mu_{k0})}{u\alpha_k} \frac{\dot{n}_k^*(t)}{(\gamma_k n_k^*(t) + 1)^2}. \tag{93}$$

By the transversality condition $\theta_k^*(T) = 0$, we have that

$$\begin{aligned} \theta_k^*(t) &= -\int_t^T \dot{\theta}_k^*(s) \mathrm{d}s = -\frac{\gamma_k(\mu_k - \mu_{k0})}{u\alpha_k} \int_t^T \frac{\dot{n}_k^*(s)}{(\gamma_k n_k^*(s) + 1)^2} \mathrm{d}s \\ &= \frac{\mu_k - \mu_{k0}}{u\alpha_k}\left(\frac{1}{\gamma_k n_k^*(t) + 1} - \frac{1}{\gamma_k n_k^*(T) + 1}\right) \end{aligned} \tag{94}$$

Moreover, (90a) can be rewritten as

$$1 = \alpha_k(p_k^*(t) + u\theta_k^*(t)) + \alpha_k \sum_{i=1}^{K} (p_i^*(t) + u\theta_i^*(t))\lambda_i(P^*(t), N^*(t)). \tag{95}$$

The summation in the right-hand side of the preceding equation is the Hamiltionian function, which is constant over the optimal trajectory, i.e. $H(P^*(t), N^*(t), \Theta^*(t)) = h$ for all $t \in [0, T]$. Hence, we have that $1/\alpha_k + h = p_k^*(t) + u\theta_k^*(t)$ for all k. Combining this result with (94), we obtain that

$$p_k^*(t) = \frac{1}{\alpha_k} + h - \frac{\mu_k - \mu_{k0}}{\alpha_k}\left(\frac{1}{\gamma_k n_k^*(t) + 1} - \frac{1}{\gamma_k n_k^*(T) + 1}\right). \tag{96}$$

Moreover, the preceding relation can be used to write

$$\beta_k + \frac{\gamma_k \mu_k n_k^*(t) + \mu_{k0}}{\gamma_k n_k^*(t) + 1} - \alpha_k p_k^*(t) = \beta_k + \mu_k - 1 - \alpha_k h - \frac{\mu_k - \mu_{k0}}{\gamma_k u n_k^*(T) + 1} =: z_k, \tag{97}$$

which, plugged into (40), establishes that demand functions are constant over the optimal solution, i.e., $\lambda_k(P^*(t), N^*(t)) = l_k(Z)$ where $Z = (z_1, \ldots, z_K)$, and hence that $n_k^*(t) = u l_k(Z) t$ for $t \in [0, T]$. Moreover, since $p_i(T) = 1/\alpha_k + h$, combining the fact that the Hamiltonian function is constant over time with the transversality condition $\theta_k^*(T) = 0$ we can find that

$$h = H(P^*(T), N^*(T), \Theta^*(T)) = \sum_{i=1}^K p_i^*(T) \lambda_i(P^*(T), N^*(T)) = \sum_{i=1}^K \left(\frac{1}{\alpha_i} + h\right) l_i(Z), \tag{98}$$

or, equivalently,

$$h = \frac{\sum_{i=1}^K \frac{1}{\alpha_i} l_i(Z)}{1 - \sum_{i=1}^K l_i(Z)} = \sum_{i=1}^K \frac{1}{\alpha_i} \exp(z_i). \tag{99}$$

Plugging this result in the definition of z_k establishes that z_k satisfies (44), which proves that any optimal solution must be characterized by (43) and therefore concludes the proof of the first part of the statement of the theorem. Since the second part of the theorem follows directly from Lemma 3, this concludes the proof of the theorem. □

Proof of Proposition 1. We first establish that Z^∞ is unique. Observe that from (46) we have for all $i \neq k$

$$\frac{z_k^\infty - \beta_k - \mu_k + 1}{\alpha_k} = \frac{z_i^\infty - \beta_i - \mu_i + 1}{\alpha_i} = -\sum_{j=1}^K \frac{1}{\alpha_j} \exp(z_j^\infty), \qquad k = 1, 2, \ldots, K. \tag{100}$$

Using the above result, (46) can be rewritten as

$$z_k - \beta_k - \mu_k + 1 + \sum_{i=1}^K \frac{\alpha_k}{\alpha_i} \exp\left[\beta_i + \mu_i - 1 - \frac{\alpha_i}{\alpha_k}(\beta_k + \mu_k - 1) + \frac{\alpha_i}{\alpha_k} z_k\right] = 0, \tag{101}$$

which is independent of z_i for $i \neq k$, i.e., (46) can be reformulated as a system of K independent equations. Moreover, notice that the real function

$$H_k(x) := x - \beta_k - \mu_k + 1 + \sum_{i=1}^K \frac{\alpha_k}{\alpha_i} \exp\left[\beta_i + \mu_i - 1 - \frac{\alpha_i}{\alpha_k}(\beta_k + \mu_k - 1) + \frac{\alpha_i}{\alpha_k} x\right] \tag{102}$$

is strictly increasing in x and that it satisfies the limits $\lim_{x \to -\infty} H_k(x) = -\infty$ and $\lim_{x \to \infty} H_k(x) = \infty$. This implies that, for each k, there exists a unique point z_k^∞ such that $H_k(z_k^\infty) = 0$, and, therefore, that Z^∞ is unique.

Next, to prove that $Z^* = Z^\infty + O(T^{-2})$ we provide a Taylor expansion of (44) for large values of T. Let $\varepsilon_1, \ldots, \varepsilon_K$ be constants to be determined, and let the "perturbed solution" $Z^* := (z_1^*, \ldots, z_K^*)$ be defined by $z_k^* = z_k^\infty + \frac{\varepsilon_k}{T}$. When evaluated in Z^*, (44) gives

$$z_k^\infty + \frac{\varepsilon_k}{T} = \beta_k + \mu_k - 1 - \sum_{i=1}^K \frac{\alpha_k}{\alpha_i} \exp\left(z_i^\infty + \frac{\varepsilon_i}{T}\right) - \frac{\mu_k - \mu_{k0}}{1 + \gamma_k u l_k(Z^\infty + \frac{\varepsilon}{T}) T}, \tag{103}$$

where $Z^\infty + \frac{\varepsilon}{T} := (z_1^\infty + \frac{\varepsilon_1}{T}, \ldots, z_K^\infty + \frac{\varepsilon_K}{T})$. The Taylor expansion of the above relation gives:

$$z_k^\infty - \beta_k - \mu_k + 1 + \sum_{i=1}^K \frac{\alpha_k}{\alpha_i} \exp(z_i^\infty) + \frac{\alpha_k}{T}\left[\frac{\varepsilon_k}{\alpha_k} + \sum_{i=1}^K \frac{\varepsilon_i}{\alpha_i} \exp(z_i^\infty) - \frac{\mu_k - \mu_{k0}}{\gamma_k u l_k(Z^\infty)}\right] + \varphi_k(T^{-2}) = 0, \tag{104}$$

where $\varphi_k(T^{-2}) = O(T^{-2})$ for all k. The term in the first square brackets of the above relation is zero because of (46). Hence, if we choose $\varepsilon_1, \ldots, \varepsilon_K$ such that

$$\frac{\varepsilon_k}{\alpha_k} + \sum_{i=1}^{K} \frac{\varepsilon_i}{\alpha_i} \exp(\bar{z}_i) - \frac{\mu_k - \mu_{k0}}{\gamma_k u l_k(\bar{Z})} = 0, \qquad k = 1, 2, \ldots, K, \tag{105}$$

when T is large enough, then the perturbed solution Z^* approximates the solution of (46) within a $O(T^{-2})$-error. Finally, it is not difficult to see that there exist a unique vector $(\varepsilon_1^*, \ldots, \varepsilon_K^*)$ that solves (105). Specifically, after some straightforward algebra, we can find that

$$\varepsilon_k^* = \alpha_k \sum_{i=1}^{K} \frac{\mu_i - \mu_{i0}}{\alpha_i \gamma_i u} - \frac{\mu_k - \mu_{k0}}{\gamma_k u l_k(\bar{Z})}, \qquad k = 1, 2, \ldots, K, \tag{106}$$

which concludes the proof. $\square$

B.2. Proofs for Auxiliary Results

Proof for Lemma 4. We fix π and let $\mathsf{E}[\cdot] = \mathsf{E}^{\pi}[\cdot]$. Note that $n_t = 0$ implies $\mathsf{E}[(r_t - \mu)^2] = \mu^2$. Hence, hereafter we only consider the case with $n_t \geq 1$. Observe that

$$\mathsf{E}\left[(r_t - \mu)^2 \,|n_t \geq 1\right] = \mathsf{E}\left[\frac{\sigma^2}{n_t} \,\middle|\, n_t \geq 1\right]. \tag{107}$$

First, consider the case with $v \in [0,1]$. To bound the above quantity, let $\underline{\lambda} > 0$ be the minimum of the demand function $\lambda(p, n, r)$ over all admissible values of the price p and the state variables n and r. Define by n_t' a binomial random variable with $(t-1)$ number of trials and success probability $\underline{\lambda} f(t) > 0$. Since $n_{t+1} = n_t + 1$ with probability $f(n_t)\lambda(p_t, n_t, r_t) \geq f(t)\underline{\lambda}$ and $n_{t+1} = n_t$ otherwise, we have that $\mathsf{E}[1/n_t|\ n_t \geq 1] \leq \mathsf{E}[1/(n_t' + 1)]$. Further, applying Theorem 1 of Chao and Strawderman (1972) for n_t' we have

$$\mathsf{E}\left[\frac{\sigma^2}{n_t' + 1}\right] \leq \frac{1 - (1 - \underline{\lambda} f(t))^{t+1}}{\underline{\lambda}(t+1) f(t)} = \frac{M_1'}{t f(t)} \leq \frac{2M_1'}{t f(t) + 1} \tag{108}$$

for $t \geq 1$ and for some constants $M_1' < \infty$. Combining these inequalities we get

$$\mathsf{E}\left[(r_t - \mu)^2 \,|n_t \geq 1\right] \leq \frac{2M_1}{t f(t) + 1}, \tag{109}$$

and the desired result follows by letting $M_1 = \max(2M_1', \mu^2)$.

Next, consider the case with $v \in (1, \infty)$. It can be easily seen from Jensen's inequality that

$$\mathsf{E}\left[\frac{\sigma^2}{n_t} \,\middle|\, n_t \geq 1\right] \geq \frac{\sigma^2}{\mathsf{E}[n_t|\ n_t \geq 1]} \geq \frac{\sigma^2}{\mathsf{E}[n_t] + 1}. \tag{110}$$

Define $n_\infty := \lim_{t \to \infty} n_t$ and $\tau_m := \inf\{t : n_t = m\}$. Observe that $\mathsf{E}[n_t] \leq \mathsf{E}[n_\infty]$ from the monotone convergence theorem and that

$$\begin{aligned}
\mathsf{E}[n_\infty] &\leq \mathsf{E}\left[\sum_{t=1}^{\infty} \mathbf{I}\{\text{Customer } t \text{ post a review}\}\right] \\
&= \mathsf{E}\left[\sum_{m=1}^{\infty} \mathbf{I}\{\text{Customer } \tau_m \text{ post a review}\}\right] \\
&= \sum_{m=1}^{\infty} \frac{1}{u m^v} \\
&< \infty,
\end{aligned} \tag{111}$$

where $\mathbf{I}\{A\}$ is one if A is true and zero otherwise. The proof is therefore concluded, as we can always find a positive constant M_2 such that

$$\mathsf{E}\left[\frac{\sigma^2}{n_t}\,\middle|\, n_t \geq 1\right] \geq M_2. \quad \square \tag{112}$$

Proof for Lemma 5. Observe that

$$\begin{aligned}
&\mathsf{E}^\pi\left[\left|\hat{p}_{\lfloor t\rfloor}\lambda(\hat{p}_{\lfloor t\rfloor}, n^*(t), r^*(t)) - \hat{p}_{\lfloor t\rfloor}\lambda(\hat{p}_{\lfloor t\rfloor}, n_{\lfloor t\rfloor}, r_{\lfloor t\rfloor})\right|\right] \\
&\leq L\mathsf{E}^\pi\left|\frac{\gamma n^*(t)r^*(t)+\mu_0}{\gamma n^*(t)+1} - \frac{\gamma \hat{n}_{\lfloor t\rfloor}\hat{r}_{\lfloor t\rfloor}+\mu_0}{\gamma\hat{n}_{\lfloor t\rfloor}+1}\right| \\
&\leq L\left(\left|\frac{\gamma n^*(t)r^*(t)+\mu_0}{\gamma n^*(t)+1} - \mu\right| + \mathsf{E}^\pi\left|\frac{\gamma \hat{n}_{\lfloor t\rfloor}\hat{r}_{\lfloor t\rfloor}+\mu_0}{\gamma\hat{n}_{\lfloor t\rfloor}+1} - \mu\right|\right),
\end{aligned} \tag{113}$$

where the first inequality follows from the Lipschitz continuity of λ, while the second follows from the triangular inequality. Consider the first term on the right-hand side of the preceding inequality and observe that $r^*(t) = \mu$ and $n^*(t) = u\ell(z)t$. Therefore,

$$\left|\frac{\gamma n^*(t)r^*(t)+\mu_0}{\gamma n^*(t)+1} - \mu\right| = \left|\frac{\mu_0-\mu}{\gamma n^*(t)+1}\right| = \mathcal{O}\left(\frac{1}{t}\right). \tag{114}$$

To bound the second term on the right-hand side of (113), observe that

$$\mathsf{E}^\pi\left|\frac{\gamma \hat{n}_{\lfloor t\rfloor}\hat{r}_{\lfloor t\rfloor}+\mu_0}{\gamma\hat{n}_{\lfloor t\rfloor}+1} - \mu\right| \leq \mathsf{E}^\pi\left|\hat{r}_{\lfloor t\rfloor} - \mu\right| + \mathsf{E}^\pi\left|\frac{\mu_0-\mu}{\gamma\hat{n}_{\lfloor t\rfloor}+1}\right| = \mathcal{O}\left(\frac{1}{\sqrt{t}}\right), \tag{115}$$

where the last equality follows from Lemma 4 with $v = 0$ and from the fact that $\mathsf{E}|X| \leq (\mathsf{E}|X|^2)^{0.5}$. This completes the proof of the lemma. $\square$

References

Acemoglu D, Makhdoumi A, Malekian A, Ozdaglar A (2017) Fast and slow learning from reviews. Technical Report 24046, National Bureau of Economic Research, URL `http://dx.doi.org/10.3386/w24046`.

Ajorlou A, Jadbabaie A, Kakhbod A (2016) Dynamic pricing in social networks: The word-of-mouth effect. *Management Sci.* 64(2):971–979.

Aviv Y, Pazgal A (2008) Optimal pricing of seasonal products in the presence of forward-looking consumers. *Manufacturing & Service Oper. Management* 10(3):339–359.

Banerjee AV (1992) A simple model of herd behavior. *The Quarterly J. Econom.* 797–817.

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.* 992–1026.

Bloch F, Quérou N (2013) Pricing in social networks. *Games and Economic Behavior* 80:243–261.

Bonnans J, Shapiro A (1998) Optimization problems with perturbations: A guided tour. *SIAM Review* 40(2):228–264.

Bose S, Orosel G, Ottaviani M, Vesterlund L (2006) Dynamic monopoly pricing and herding. *RAND J. Econom.* 37(4):910–928.

Bose S, Orosel G, Ottaviani M, Vesterlund L (2008) Monopoly pricing in the binary herding model. *Economic Theory* 37(2):203–241.

Candogan O, Bimpikis K, Ozdaglar A (2012) Optimal pricing in networks with externalities. *Oper. Res.* 60(4):883–905.

Cesari L (1966) Existence theorems for weak and usual optimal solutions in lagrange problems with unilateral constraints. i. *Trans. Amer. Math. Soc.* 124(3):369–412.

Chao M, Strawderman WE (1972) Negative moments of positive random variables. *J. Amer. Stat. Assoc.* 67(338):429–431.

Chevalier JA, Mayzlin D (2006) The effect of word of mouth on sales: Online book reviews. *J. Marketing res.* 43(3):345–354.

Crapis D, Ifrach B, Maglaras C, Scarsini M (2016) Monopoly pricing in the presence of social learning. *Management Sci.* 63(11):3586–3608.

DeGroot MH (2005) *Optimal statistical decisions* (John Wiley & Sons).

Dellarocas C (2003) The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Sci.* 49(10):1407–1424.

Duan W, Gu B, Whinston AB (2008) Do online reviews matter?—an empirical investigation of panel data. *Decision Support Systems* 45(4):1007–1016.

Gallego G, van Ryzin G (1994) Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Sci.* 40(8):999–1020.

Gallego G, van Ryzin G (1997) A multiproduct dynamic pricing problem and its applications to network yield management. *Oper. Res.* 45(1):24–41.

He QC, Chen YJ (2018) Dynamic pricing of electronic products with consumer reviews. *Omega* 80:123–134.

Judd KL, Riordan MH (1994) Price and quality in a new product monopoly. *The Review of Economic Studies* 61(4):773–789.

Li H, Huh WT (2011) Pricing multiple products with the multinomial logit and nested logit models: Concavity and implications. *Manufacturing & Service Oper. Management* 13(4):549–563.

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

Ma Q, Shou B, Huang J, Başar T (2018) Dynamic pricing in the presence of participation-dependent social learning. *Proc. 18th ACM Internat. Symp. on Mobile Ad Hoc Networking and Computing*, 101–110 (ACM).

Makhdoumi A, Malekian A, Ozdaglar AE (2018) Strategic dynamic pricing with network externalities. Technical report, Rotman School of Management Working Paper No. 2980109, URL `https://ssrn.com/abstract=2980109`.

Milgrom P, Roberts J (1986) Price and advertising signals of product quality. *J. Political Economy* 94(4):796–821.

Monetate (2016) Ecommerce quarterly report. URL `http://www.monetate.com/resources/research/`.

Papanastasiou Y, Savva N (2017) Dynamic pricing in the presence of social learning and strategic consumers. *Management Sci.* 63(4):919–939.

Roberts JH, Urban GL (1988) Modeling multiattribute utility, risk, and belief dynamics for new consumer durable brand choice. *Management Sci.* 34(2):167–185.

Schnapp M, Allwine T (2001) Mining of book data from amazon.com. URL `http://www2.sims.berkeley.edu/resources/affiliates/workshops/webmining/Slides/ORA.ppt`.

Sorensen AT (2004) Bestseller list and product variety: the case of book sales. *Stanford GSB Research Paper No. 1878:1-31* .

Yang N, Zhang R (2018) Dynamic pricing and inventory management in the presence of online reviews. *Available at SSRN.* URL `https://ssrn.com/abstract=2571705`.

Yu M, Debo L, Kapuscinski R (2015) Strategic waiting for consumer-generated quality information: Dynamic pricing of new experience goods. *Management Sci.* 62(2):410–435.