Newsvendor Decisions with Two-Sided Learning

Yiangos Papanastasiou
Haas School of Business · University of California, Berkeley
yiangos@haas.berkeley.edu

A substantial collection of work in operations management considers settings where a firm faces uncertain demand that depends on parameters that are ex ante unknown, but can be learned by observing historical data. This work typically assumes that future demand is unaffected by the firm's learning process. However, in the new era of social media, it is increasingly the case that the information used by the firm to gauge future demand (e.g., sales and stockouts, consumer feedback) is now also observable to the consumers and may influence their purchase decisions. In this paper, we consider a newsvendor model where a product of ex ante unknown value is sold in an environment where learning is "two-sided," in that both the firm and the consumers learn the product's value over time by observing the same information. The analysis establishes a consequential insight: when learning is two-sided, the value of information is often *negative* for the firm; as a result, the firm's optimal stocking quantity is often lower than that under one-sided learning. Moreover, under certain conditions we identify, we show that the optimal stocking quantity can be even lower than the critical-fractile policy, in stark contrast to the recurring prescription found in existing literature of "stocking more" in the presence of learning. Further results and numerical experiments suggest that the loss for the firm from failing to account for the two-sidedness of the learning process can be significant.

Key words: operations management, newsvendor model, two-sided learning, social learning

---

1. Introduction

Choosing how many units to stock in the presence of demand uncertainty is one of the most significant challenges faced by firms selling new products. The associated tradeoff is epitomized by the famous "newsvendor problem": stock too many units and the firm runs the risk of leftover inventory that is costly to store and hard to move; stock too few units and the firm risks being stocked out and leaving money on the table. Although the parameters that govern demand can be unknown to the firm to begin with, it is often the case that these can be learned over time, for example, by observing past sales and stockouts or consumer feedback.

Demand uncertainty is in many cases associated with various demand-side characteristics, such as the total size of the market, the consumers' tastes and assessment of product quality, the distribution of consumer preferences, etc. A substantial collection of work in operations management has focused on developing insights on how firms can adapt their quantity decisions so as to be in a better position to learn these characteristics and match supply with demand more effectively in the future (e.g., Besbes and Muharremoglu 2013, Ding et al. 2002, Lariviere and Porteus 1999, Lu et al. 2005); from this work, a recurring structural result is the prescription to "stock more" than the quantity derived through the well-known critical-fractile formula. When demand uncertainty stems from such demand-side characteristics, the associated learning process is referred to in this paper as "one-sided," in that the information generated and the learning that occurs is only relevant to the firm's actions moving forward, and future demand is exogenous to this learning process.

But demand can also be difficult to predict because of uncertainty associated with the product itself, in particular when the product is new and innovative. In such cases, the product's value is a function of the match between the consumers' tastes and the product's underlying characteristics, and can therefore be difficult to assess ex ante for both the firm and the consumers: while the firm is uncertain about the consumers' tastes and assessments, the consumers are uncertain about the product's unobservable attributes (e.g., performance, usefulness). A key difference in these settings is that learning is "two-sided," in that both the firm and the consumers are simultaneously learning the product's value and making decisions accordingly. With the advent of social media platforms where consumers share their actions and experiences, two-sided learning has received significant attention in the recent literature, which highlights the new challenges faced by the firm when internalizing the impact of its actions not only on its own learning process, but also on that of the consumers (e.g., Crapis et al. 2016, Feldman et al. 2018, Papanastasiou and Savva 2016, Yu et al. 2015).

Despite this recent attention, there exists no work to date that addresses newsvendor decisions in environments where learning is two-sided. As a result, it remains unclear whether and to what extent the insights and prescriptions that apply to one-sided (firm) learning extend to settings with two-sided (firm-and-consumer) learning. In this paper, we develop a model that allows us to perform a direct comparison between optimal newsvendor decisions under these two learning regimes and to investigate their associated implications.

We consider a two-period newsvendor model of a monopolist firm selling a new product of ex ante unknown value. The consumer population is divided into "innovators," who enter the market in the first period, and "followers," who enter the market in the second period. Innovators are assumed to be expert consumers, in the sense that they can accurately predict the product's value upon inspection and before making their purchase decision. Followers can be either experts or non-experts, depending on the learning regime being considered. In the case where followers are experts, the model is one of one-sided learning, because while the firm learns the product's value for the consumers (and therefore future demand) from the outcomes of its first-period actions, the consumers' decisions in the second period are unaffected by the firm's learning process. By contrast, in the case where followers are non-experts, the model is one of two-sided learning, because both the firm and the followers learn the product's value simultaneously from the outcomes of the firm's first-period actions. Consistent with the existing literature, we consider learning based on past sales observations, but we demonstrate that our model insights extend to other forms of learning as well, such as learning from product reviews.

We solve the firm's two-period inventory problem via backward induction, and use the distinction between scenarios of low and high value realizations (which result in excess first-period inventory and demand, respectively), to separate the impact of the firm's first-period quantity decision on second-period profit into two distinct channels. The first is the inventory channel, which refers to the cost-savings associated with satisfying second-period demand with leftover units from the first period rather than producing new units. The second is the information channel, which refers to the informational gains (with respect to learning product value) associated with the events of the first period. The analysis builds on the structural properties of the two channels in order to develop intuition on how they each affect the firm's optimal policy and profits.

We focus first on the implications of one- versus two-sided learning for the firm's optimal quantity decision in the first period. When learning is one-sided, both the inventory and the information channels call for an increase in the firm's first-period quantity decision. As a result, we retrieve the prescription commonly encountered in the existing literature that the firm should "stock more" than the myopic-optimal quantity (i.e., the single-period critical-fractile quantity), because this enhances the firm's ability to learn by reducing demand censoring, while at the same time the second period serves as a form of salvage market for excess first-period inventory.

When learning is two-sided, we show that the effect of the inventory channel is unchanged; however, the information channel now acts in the opposite direction: if being better-informed means that the consumers will also be better-informed, the firm prefers to know less; in other words, when learning is two-sided, the value of information is *negative* for the firm. To explain this phenomenon, we highlight the key difference in the informational role of the quantity decision under one- and two-sided learning: under one-sided learning, the firm's quantity decision should be viewed as a tool for *information extraction* (from the innovators); on the other hand, under two-sided learning, the quantity decision should be viewed as a tool for *information transmission* (to the followers). While information extraction is always accompanied by a "more-is-better" property, the same does not necessarily apply in the case of information transmission; indeed, we show that in our model the corresponding property is of a "less-is-better" nature. Building on this observation, we establish that (i) the optimal quantity under two-sided learning is lower than that under one-sided learning, and (ii) if the effect of the inventory channel is weak, the optimal quantity under two-sided learning is also lower than the myopic-optimal quantity; one sufficient condition for the latter is that the proportion of innovators in the population is not too high—thus, when learning is two-sided, we find that the firm should in fact "stock less."

We then consider the implications of the two alternative learning regimes for the firm's expected profit. Here, we identify another interesting phenomenon. Although our preceding results paint the picture of a firm that "likes" and takes actions to encourage learning when this is one-sided, but "dislikes" and takes actions to decrease learning when it is two-sided, we nevertheless show that a firm operating in a two-sided environment is ex ante better off compared to one that operates under one-sided learning. One interpretation of this result is that the firm, before knowing the product's realized value, prefers to face customers that are non-experts, because this allows it to exercise a form of ex ante control on the consumers' knowledge, through its choice of the first-period stocking level. At the same time, our analysis underscores the need for the firm to account for the correct learning regime when making its quantity decisions; in particular, we show that adjusting the quantity decision to account for the wrong type of learning regime is typically worse than ignoring learning altogether (i.e., employing the myopic-optimal policy). Through numerical experiments, we further demonstrate that policies which do not internalize the two-sidedness of the learning process can lead to significant profit losses for the firm, reaching up to 36% in our experiments.

2. Literature Overview

This paper bridges the gap between two existing streams of research in operations management. The first is the relatively well-developed line of work that incorporates demand learning into newsvendor decisions. The second is a more recent area of research that investigates the implications of two-sided learning for firm-level operational decisions. In this section, we provide an overview of each stream along with some of their main takeaways.

The research that focuses on how learning should be taken into account in newsvendor decisions has focused to date on what we refer to in this paper as "one-sided" learning. The firm faces demand that depends on one or more parameters whose values are ex ante unknown but can be learned over time by observing historical data. Learning in these models is assumed to be one-sided, in the sense that the firm's learning process does not affect the parameters that govern future demand realizations.[1] This assumption is reasonable in cases where demand uncertainty stems from demand-side characteristics which are irrelevant to the purchase decision of any individual

[1] Throughout this paper, "one-sided" learning refers to learning of demand by the firm. We do not consider one-sided learning of product characteristics by the consumers, which typically focuses on the firm's ability to signal its private knowledge; for such work, see Stock and Balachander (2005) and references therein.

consumer (e.g., the total size of the market, or the distribution of the consumers' willingness-to-pay), or where the source of demand uncertainty could potentially affect individual purchase decisions, but consumers do not have access to the data observed by the firm. Earlier work in this area focused on one-sided learning from past demand observations; for examples of such work, see Azoury (1985), Lovejoy (1990), and Scarf (1959).

More recent papers consider the more realistic case where learning occurs in the presence of "demand censoring": in the event of a stockout, the firm observes only a censored signal of demand. This is the mode of learning considered in our model as well. Examples of work that incorporate demand censoring include, among others, Bensoussan and Sethi (2009), Besbes and Muharremoglu (2013), Ding et al. (2002) (see also Lu et al. 2005), Lariviere and Porteus (1999), Petruzzi and Dada (2002); see also Chen and Mersereau (2015) and Mersereau (2015) for a more comprehensive discussion of this work. Demand censoring gives rise to an exploration-exploitation tradeoff, in that the firm must decide how to balance current against future profits: being in a better position (from an informational perspective) in the future invariably requires decisions that result in suboptimal current profits. From this work, a recurring structural result is that the firm should "stock more" in the presence of learning; that is, the firm should increase production relative to the myopic-optimal policy (given by the single-period critical-fractile formula). This paper is the first to introduce two-sided learning in a newsvendor setting and to perform a comparison between the two learning regimes of one- and two-sided learning. When learning is one-sided, our model is much in the mould of those mentioned above, and for this case we also retrieve the "stock more" prescription. By contrast, when learning is two-sided, we show that this prescription is often reversed, while even when it is not, failing to account for the two-sidedness of learning by appropriately adjusting quantity decisions can result in significant profit losses.

Two-sided learning is a central feature in the growing stream of research on the implications of social learning for operational decisions. In contrast to one-sided learning, two-sided learning refers to environments where both the firm and the consumers are simultaneously learning their parameters of interest by observing the same information. In such settings, the firm's exploration-exploitation problem often becomes more complex, because the firm must now additionally account for the impact of learning on the consumers' future actions. In the economics literature, Bergemann and Välimäki (1997) study the diffusion of a new product in a duopolistic market with two-sided learning, while Liu and Schiraldi (2012) study a monopolist's launch sequence as a tool to induce or prevent herding behavior among potential product adopters. In operations management, Swinney (2011) investigates the value of quick response production practices when the firm and the consumers simultaneously learn the size of the market and their valuations for the product, respectively. Closer to our paper is the work that focuses on two-sided learning in the context of experience goods, where the firm and the consumers simultaneously learn the product's value/quality (i.e., the match between consumer tastes and product characteristics). Crapis et al. (2016), Papanastasiou and Savva (2016) and Yu et al. (2015) investigate the implications of two-sided learning for a monopolist firm's dynamic pricing policy, and Feldman et al. (2018) focus on its implications for product design and quality. The aforementioned papers consider various policy implications of two-sided learning in make-to-order systems; our paper contributes to this area of research by studying the implications of two-sided learning for inventory decisions in a make-to-stock system.

This work also adds to a long list of papers that build on the classic newsvendor problem to investigate the performance of various operational strategies. Among such work, Cachon and Swinney (2009) study the value of price commitment and quick response inventory practices; Su and Zhang (2009) focus on the value of availability guarantees when facing strategic consumers; Bimpikis and Markakis (2015) discuss the benefits of inventory pooling under different types of demand distributions; Hu et al. (2015) propose operations and marketing policies that enable the firm to leverage social influence. This paper considers how a firm should adjust early production in the presence of two-sided learning, and can therefore be viewed as an investigation into the efficacy of mass versus limited release strategies for new experience goods.

Finally, our result on the optimality of stocking less than the critical-fractile quantity implies that stock outs occur more frequently under two-sided learning. In this respect, our work relates to a collection of papers that highlight the use of "strategic stockouts" as a way of generating increased overall demand for the product, in settings other than the newsvendor. DeGraba (1995) and Courty and Nasiry (2016) demonstrate that a firm may intentionally undersupply the market to induce buying frenzies when consumers are uncertain about their valuations for the product; Stock and Balachander (2005) show that stockouts can be an effective signal of quality when an informed firm faces uninformed consumers; Debo and Van Ryzin (2009) find that an asymmetric allocation of inventory to ex ante identical retailers can lead to higher overall demand when consumers interpret stockouts as a signal of high quality; and Papanastasiou et al. (2016) highlight that deliberate product scarcity can be used to leverage consumer self-selection in order to generate more favorable product reviews.

3. Model Description

We consider a two-period version of the classic newsvendor problem (see Arrow et al. 1951). There is a single firm selling a single product over two periods, $t \in \{1,2\}$. In both periods, the price and per unit production cost of the product are exogenous and denoted by p and c respectively, where $0 < c < p$.[2] The production quantity in period t is denoted by Q_t. The product is non-perishable;

[2] In Appendix B we demonstrate that endogenous pricing does not affect our model insights.

leftover units in the first period are carried over to the second period, while leftover units at the end of the second period have zero salvage value. For simplicity, we assume that excess demand in either period is lost forever, and that inventory holding costs are negligible; relaxing either of these assumptions would only strengthen our main insights.

The product's value is ex ante unknown and denoted by v. Ahead of the selling season, v is viewed as a random variable with positive support on the closed interval $[v_l, v_h]$ for some $v_h > v_l \geq 0$; let $f(v)$ denote the corresponding probability density function (pdf), $F(v)$ the cumulative distribution function (cdf), and define $\bar{F}(v) := 1 - F(v)$. The uncertainty in the product's value in our model represents the uncertainty associated with the match between the product's characteristics on one side and the consumers' tastes on the other. In turn, this can depend on various uncertainties on either side of the match, including the outcome of the manufacturing process, the performance of the product's individual components and overall design, the usefulness of the product's functions for the consumers, the consumers' assessment of how innovative the product is, etc.

Demand for the product is composed of a population of infinitesimally small consumers with total mass normalized to one. An individual consumer's willingness to pay for a product of value v is given by $u_i = x_i v - p$, where the consumers' value sensitivity or "type" x_i is a random draw from a distribution with pdf denoted by $g(x)$, cdf denoted by $G(x)$, and positive support on $x \in [x_l, x_h]$ for some $x_h > x_l \geq 0$. To simplify exposition, we present our results for the special case where $g(x)$ is the standard uniform pdf, and comment as necessary wherever results are sensitive to this assumption. Accordingly, to avoid trivial cases we also require that $v_h > p$ (i.e., there are at least some realizations of product value v at which some consumers are willing to purchase the product) and we set $v_l = 0$ (i.e., there are at least some realizations of product value v at which no consumer is willing to purchase).

We assume that there are two categories of consumers, "innovators" who arrive in the first period, and "followers" who arrive in the second period. The proportion of innovators in the population is $\gamma \in (0, 1]$. Consistent with a large body of work in marketing and the diffusion of innovations (see Peter et al. (1999) and Rogers (2010)), innovators are assumed to be expert consumers who are able to assess the product's value before making their purchase decision.[3] Depending on the learning regime under investigation, followers can be either experts or non-experts; more specifically:

(i) When the followers are experts, learning is "one-sided": all consumers are informed about the product's value and only the firm is learning from the events of the first period.

(ii) When the followers are non-experts, learning is "two-sided": both the firm and the followers are simultaneously learning the product's value from the events of the first period.

[3] Our main analysis assumes that expert consumers are perfect in their assessments of product value; in §8.2 we relax this assumption and demonstrate that the analysis does not change significantly.

It is important to note that either type of learning regime is motivated by a distinct type of product. When the product is a "search good," it is sensible to assume that all consumers are experts, since by definition there is no significant uncertainty associated with product attributes that are unobservable before purchase. When the product is an "experience good," this uncertainty is significant and only consumers with specialized knowledge (the innovators) are likely to be in a position to judge the product before purchase. The firm, on the other hand, faces uncertainty with regards to the consumers' tastes irrespective of the type of product. To be consistent with the existing literature on newsvendor decisions with learning, our analysis focuses on learning from observed sales and stockouts (see §4.2); in §8.1, we consider an alternative mode of learning based on consumer reviews, and show that our main results are qualitatively unchanged.

Both the firm and the consumers are assumed to be risk-neutral utility maximizers, with outside options normalized to zero. We first analyze the drivers underlying the firm's optimal quantity policy when the firm operates under either of the two learning regimes. We then consider the associated implications for the firm's ex ante profit. We conclude the analysis by demonstrating robustness of our main insights under alternative model specifications.

4. Preliminaries

4.1. Myopic-Optimal Quantity Decisions

The main point of interest in our analysis is the change in the firm's optimal first-period quantity decision when the firm internalizes the impact of the resulting (either one- or two-sided) learning process on its second-period profit. It is therefore useful to first establish a benchmark where the firm disregards the future in the first period, employing a so-called "myopic-optimal" policy.

When choosing the first-period quantity, the firm is uncertain about the product's realized value. In the first period, a mass γ of expert consumers inspect the product and decide whether to purchase. The firm's expected profit as a function of the quantity decision Q_1 is given by

$$\pi_1(Q_1) = pE\left[\min\left\{D_1(v), Q_1\right\}\right] - cQ_1,$$

where the first-period demand at value v is $D_1(v) = \gamma\left[1 - \frac{p}{v}\right]^+$ and $[d]^+ := \max\{0, d\}$. The optimal policy is specified by the well-known "critical-fractile formula," as described in Proposition 1.

Proposition 1. *The unique myopic-optimal first-period quantity decision is*

$$Q^*_{1,m} = \gamma\left[1 - \frac{p}{v_m}\right]^+, \text{ where } v_m = F^{-1}\left[\frac{p-c}{p}\right]. \tag{1}$$

All proofs are provided in Appendix C. The quantity that maximizes the firm's expected profit in the first period is determined by optimizing over the critical value v_m, such that if the product's value is higher than v_m the firm sells out, while if it is lower than v_m the firm has leftover units. As in the standard newsvendor model, v_m is chosen by the firm according to the familiar recipe which optimally balances overage against underage costs.

To avoid obvious discussions regarding corner solutions, the analysis that follows will assume that $Q^*_{1,m} > 0$. Moreover, it will be useful to define $Q_{\max} := \gamma\left[1 - \frac{p}{v_h}\right]$, which is the maximum possible demand in the first period (i.e., $Q_{\max}$ is the quantity demanded by the experts in the first period when the product's realized value is v_h).

4.2. Learning From Sales and Stockouts

While the product's realized value is unobservable, the firm and the consumers can observe the events of the first period and use this information to update their beliefs. In particular, we assume that the firm and the consumers can observe (i) the volume of sales, and (ii) whether or not the product sold out in the first period (in §8.3, we extend the model to the case where the consumers observe this information only with some probability).[4] A potential obstacle in learning the product's value is the "demand-censoring" effect, which refers to the inability to observe demand precisely in the event of a stockout (see Chen and Mersereau 2015). We discuss here the process of learning based on sales in the context of our model, and highlight how demand censoring affects this process.

Consider first the case where S_1 sales occurred in the first period and the product did not stock out. Since the innovators are by assumption informed about product value, upon observing a volume $S_1 > 0$ of first-period sales, the product's value can be inferred by equating first-period demand with first-period sales, $\gamma\left(1 - \frac{p}{v}\right) = S_1$, to get $v = \frac{p}{1-\frac{S_1}{\gamma}}$. By contrast, if S_1 sales occurred and the product stocked out in the first period, the product's value can only be inferred up to the inequality $\gamma\left(1 - \frac{p}{v}\right) \geq S_1$, that is, up to the fact that demand was greater than the observed sales. The updated expectation of product value in the event of a stockout is then $E\left[v \mid v \geq \frac{p}{1-\frac{S_1}{\gamma}}\right]$.

Now, suppose the firm chooses the first-period stocking quantity $Q_1 \in [0, Q_{\max}]$.[5] If a stockout occurs in the first period, the updated value expectation will be

$$\hat{v}(Q_1) := E\left[v \;\middle|\; v \geq \frac{p}{1 - \frac{Q_1}{\gamma}}\right], \tag{2}$$

since the observed volume of sales will be $S_1 = Q_1$. Using the last expression, we may define

$$v_c(Q_1) := \frac{p}{1 - \frac{Q_1}{\gamma}} \tag{3}$$

[4] In practice, consumers have access to information on the sales performance of products through several channels, including popular press articles (e.g., Cnet.com 2018), firms' performance reports (e.g., Apple Inc. 2017), investment analyst reports (Bloomberg 2018), direct in-store observations, purchases among members of their immediate social network, word-of-mouth communications, social media posts (e.g., Facebook likes, comments and shares), etc.

[5] Note that it is not necessary in our model for consumers to observe Q_1.

as the critical product value above which (below which) a stockout occurs (does not occur) in the first period when the firm stocks Q_1 units. To summarize the above discussion,

LEMMA 1. *When the first-period quantity is Q_1, learning takes one of two possible forms:*

(i) If the realized product value satisfies $v < v_c(Q_1)$, then a stockout does not occur in the first period and the product's value is perfectly revealed.

(ii) If the realized product value satisfies $v \geq v_c(Q_1)$, then a stockout occurs in the first period and value is only imperfectly revealed. The updated value expectation is given by $\hat{v}(Q_1) = E[v \mid v \geq v_c(Q_1)]$.

5. One-Sided Learning

We begin our analysis with the case of one-sided learning. The key feature of this setting is that the firm learns the product's value for the consumers over time, and the information generated during this learning process does not affect future demand for the product. To capture one-sided learning within our model setup, we assume that consumers in both periods are experts, in the sense that they can accurately assess the match between the product's characteristics and their own tastes before purchase (e.g., as is the case when the product is closer to a search good). The analysis focuses on the properties of the optimal first-period quantity $Q^*_{1,1s}$ (throughout our analysis, subscripts "$1s$" and "$2s$" denote the "one-sided" and "two-sided" learning regimes, respectively).

The first-period quantity Q_1 is chosen by the firm so as to maximizing its total expected profit,

$$\pi_{t,1s}(Q_1) = \pi_1(Q_1) + \pi^*_{2,1s}(Q_1), \tag{4}$$

where $\pi_1(Q_1)$ is the firm's expected first-period profit and $\pi^*_{2,1s}(Q_1)$ is the firm's optimal expected second-period profit (i.e., the firm's expected profit assuming the second-period quantity is chosen optimally following the events of the first period). The firm's first-period expected profit $\pi_1(Q_1)$ is the function considered in the benchmark case of §4.1; recall that $\pi_1(Q_1)$ has a unique maximum at the myopic-optimal quantity $Q^*_{1,m}$ described in Proposition 1. In the two-period problem, any deviation from the quantity decision $Q^*_{1,m}$ is attributed to the firm's future-profit considerations. Our approach will be to establish the properties $\pi^*_{2,1s}(Q_1)$ with respect to Q_1, which will in turn determine the firm's policy adjustment relative to $Q^*_{1,m}$.

In doing so, it is useful to observe that there are two distinct channels through which the first-period quantity decision Q_1 affects the firm's second-period profit: the first is the *inventory channel*, which refers to leftover inventory carried forward from the first to the second period; the second is the *information channel*, which refers to the informational gains achieved in the first period (with regards to product value). It is possible to isolate these two channels by expressing the firm's second-period profit (see Appendix A.1 for a detailed analysis of the firm's second-period problem) as $\pi^*_{2,1s}(Q_1) = \pi^*_{2C,1s}(Q_1) + \pi^*_{2I,1s}(Q_1)$, where

$$\pi^*_{2C,1s}(Q_1) := \int_{v_l}^{v_c(Q_1)} c \min\left\{(1-\gamma)\left[1-\frac{p}{v}\right]^+, Q_L(Q_1,v)\right\} f(v)dv, \text{ and}$$

$$\begin{aligned}\pi^*_{2I,1s}(Q_1) := &\int_{v_l}^{v_c(Q_1)} (p-c)(1-\gamma)\left[1-\frac{p}{v}\right]^+ f(v)dv \\ &+ \int_{v_c(Q_1)}^{v_h} \left[p \min\left\{(1-\gamma)\left(1-\frac{p}{v}\right), (1-\gamma)\left(1-\frac{p}{v_{2c}(Q_1)}\right)\right\}\right] f(v)dv \\ &- \int_{v_c(Q_1)}^{v_h} \left[c(1-\gamma)\left(1-\frac{p}{v_{2c}(Q_1)}\right)\right] f(v)dv.\end{aligned}$$

In words, $\pi^*_{2I,1s}(Q_1)$ is the firm's expected second-period profit assuming that all units sold in the second-period will be newly-produced, while $\pi^*_{2C,1s}(Q_1)$ is the expected second-period cost savings associated with Q_1, which come from the fact that the firm may have leftover inventory from the first period (thus avoiding some production costs). The impact of the inventory channel is captured in $\pi^*_{2C,1s}(Q_1)$ through the nonnegative leftover inventory $Q_L(Q_1,v)$, which increases in Q_1 under any value realization v. On the other hand, the impact of the information channel is captured in $\pi^*_{2I,1s}(Q_1)$ through the stockout threshold $v_c(Q_1)$, which is increasing in Q_1—the higher the firm's first-period quantity decision the larger the range of value scenarios in which the firm is able to learn v with perfect accuracy.

Using the above expressions, we next investigate the relationship between the firm's first-period quantity decision Q_1 and the contributions of the two channels to the firm's second-period profit. Starting with the inventory channel, we observe that

Proposition 2. *When learning is one-sided, the marginal expected value of cost savings* $\frac{\partial \pi^*_{2C,1s}}{\partial Q_1}$ *is strictly positive for all* $Q_1 \in [0, Q_{\max}]$.

The result is intuitive: Under any realization of product value, the larger the number of units the firm produces in the first period, the larger the number of leftover units, and therefore the lower the production costs the firm will have to incur in the second period.[6]

The next result establishes an analogous property for the information channel.

Proposition 3. *When learning is one-sided, the marginal expected value of information* $\frac{\partial \pi^*_{2I,1s}}{\partial Q_1}$ *is strictly positive for all* $Q_1 \in [0, Q_{\max}]$.

That is, the value of gaining additional information by increasing the first-period quantity is positive across all values of $Q_1 \in [0, Q_{\max}]$. The higher the first-period quantity, the higher the firm's ability

[6] Note that policies involving $Q_1 > Q_{\max}$ (i.e., producing more than the maximum first-period demand) are strictly dominated by the policy $Q_1 = Q_{\max}$, so that such policies do not require attention.

to learn product value accurately (through the higher stockout threshold $v_c(Q_1)$). In turn, knowing more about product value means that the firm is in a better position to match supply with demand in the second period (and therefore avoid overage and underage costs) in a larger range of value scenarios. As a result, the firm's expected second-period profit increases. Note that Proposition 3 does not depend on the distribution of product value F or on the consumers' type distribution G.

Returning now to the firm's total profit function, we rewrite (4) as

$$\pi_{t,1s}(Q_1) = \pi_1(Q_1) + \pi^*_{2C,1s}(Q_1) + \pi^*_{2I,1s}(Q_1). \tag{5}$$

Recall that from the benchmark analysis of §4.1, the first-period profit $\pi_1(Q_1)$ is maximized at $Q^*_{1,m}$. Moreover, Propositions 2 and 3 suggest that $\pi^*_{2C,1s}(Q_1)$ and $\pi^*_{2I,1s}(Q_1)$ are strictly increasing in $Q_1 \in [0, Q_{\max}]$; Proposition 4 follows immediately.

Proposition 4. *When learning is one-sided, the optimal first-period quantity is strictly higher than the myopic-optimal first-period quantity; that is,* $Q^*_{1,1s} > Q^*_{1,m}$.

Increasing the quantity decision with respect to $Q^*_{1,m}$ results in second-period benefits both from an inventory and from an informational perspective. Several existing papers offer analogous structural results (e.g., Bensoussan and Sethi 2009, Ding et al. 2002, Lariviere and Porteus 1999, Lu et al. 2005), albeit typically under the assumption of perishable inventory (in that case, the inventory channel discussed here vanishes and the information channel is the only driver behind the increase in the firm's optimal quantity decision). These results form the basis of the well-known "stock more" prescription, which refers to the optimality of increasing supply relative to the critical-fractile quantity when the firm operates in the presence of learning.

Before proceeding to the analysis of the two-sided learning regime, it will be convenient to assume that the distribution F is sufficiently well-behaved so that

Assumption 1. *The function* $\pi_{t,1s}(Q_1)$ *is unimodal in* $Q_1 \in [0, Q_{\max}]$.

6. Two-Sided Learning

We now consider the case of two-sided learning. In this case, the firm and the consumers learn the product's value simultaneously over time, so that, unlike the case of one-sided learning, the information generated in the first period now affects the second-period demand. We capture two-sided learning in our model by assuming that the second-period consumers are non-experts who cannot accurately assess the product's value before purchase (e.g., as is the case when the product is closer to an experience good), and update their value beliefs along with the firm after observing the first-period sales. The analysis of this section follows the same approach as that in §5; wherever possible, we refer to the preceding analysis to shorten the exposition.

Focusing again on the first-period quantity decision, the firm chooses Q_1 to maximize

$$\pi^*_{t,2s}(Q_1) = \pi_1(Q_1) + \pi^*_{2,2s}(Q_1).$$

Using the same approach as in the case of one-sided learning (see Appendix A.2 for a detailed analysis of the firm's second-period problem), the firm's optimal expected second-period profit can again be expressed as the sum of profit contributions from the inventory and the information channels. In this case, we may write $\pi^*_{2,2s}(Q_1) = \pi^*_{2C,2s}(Q_1) + \pi^*_{2I,2s}(Q_1)$, where

$$\pi^*_{2C,2s}(Q_1) := \int_{v_l}^{v_c(Q_1)} c \min\left\{(1-\gamma)\left[1-\frac{p}{v}\right]^+, Q_L(Q_1,v)\right\} f(v)dv, \text{ and}$$
$$\pi^*_{2I,2s}(Q_1) := \int_{v_l}^{v_c(Q_1)} (p-c)(1-\gamma)\left[1-\frac{p}{v}\right]^+ f(v)dv + \int_{v_c(Q_1)}^{v_h} (p-c)(1-\gamma)\left[1-\frac{p}{\hat{v}(Q_1)}\right]^+ f(v)dv.$$

Note first that for any given Q_1, the scenarios in which the firm has leftover inventory are identical under one- and two-sided learning so that $\pi^*_{2C,2s}(Q_1) = \pi^*_{2C,1s}(Q_1)$. It follows from §5 that the marginal expected value of cost savings remains strictly positive for all $Q_1 \in [0, Q_{\max}]$ under two-sided learning (i.e., Proposition 2 holds unchanged).

By contrast,

Proposition 5. *When learning is two-sided, the marginal expected value of information* $\frac{\partial \pi^*_{2I,2s}}{\partial Q_1}$ *is strictly negative for all* $Q_1 \in [0, Q_{\max}]$.

Recall that in the case of one-sided learning, the firm is always willing to pay in exchange for more accurate information (see Proposition 3). Interestingly, under two-sided learning we find that the firm is always willing to pay for *less* accurate information. Put differently, if being better-informed means that the consumers will also be better-informed, the firm prefers to know less about the product's value. We note that the result of Proposition 5 holds beyond our specific model assumptions, under the sufficient condition that demand for the product is concave in the product's value (see the proof of Proposition 5); furthermore, in §8.1 we demonstrate that the result also holds under an alternative mode of learning, based on product reviews.

The key intuition underlying the difference between Propositions 3 and 5 has to do with the difference in the informational role of the quantity decision under the two learning regimes. Under one-sided learning, the quantity decision Q_1 should be viewed as a tool for *information extraction* from the innovators. Figuratively speaking, we might think of a situation where nature has disclosed all information relevant to the firm-consumer interaction to the consumer population, leaving the firm at an informational disadvantage. To improve its position in the second period, the firm must extract this information from the first-period consumers; in turn, the higher the first-period quantity decision, the more precise the information the firm extracts from the innovators (because the firm extracts perfect information in a larger range of value scenarios). In the case of information extraction, more is always better.

By contrast, under two-sided learning, the quantity decision Q_1 should be viewed mainly as a tool for *information transmission* to the followers. Here, we think of an alternative situation where nature has left both the firm and the second-period consumers uninformed. The firm can again attempt to extract information from the innovators, but now whatever is learned by the firm is also learned by the consumers. Thus, the firm's first-period quantity decision now determines not only the information extracted from the innovators, but also, and more importantly, the information transmitted from the innovators to the followers. Generally speaking, information transmission as a process to be optimized is significantly more complex, and the simple "more-is-better" rationale often does not apply (see, for example, Alizamir et al. (2018), Papanastasiou et al. (2017)).

In the specific case of our model, the quantity decision Q_1 is positively related to the threshold $v_c(Q_1)$, below which (above which) learning of product value is perfect (imperfect), as described in Lemma 1. Therefore, the firm's choice of Q_1 can be thought of as a decision that controls both the firm's and the consumers' ability to learn product value, with higher quantities resulting in increased learning. In the case of one-sided learning, more information creates value for the firm by putting it in a better position to match supply with demand in the second period. By contrast, in the case of two-sided learning, more information destroys value for the firm by decreasing its ability to extract surplus from the now-better-informed consumers. As a result, the "more-is-better" property associated with one-sided learning is replaced by a "less-is-better" property when learning is two-sided.

Turning to the firm's optimal first-period quantity decision, Proposition 5 suggests that from an informational perspective, the firm is better off by stocking fewer units in the first period, so as to reduce the information transmission from the innovators to the followers. This observation leads to our main result.

Theorem 1. *When learning is two-sided:*

(i) The optimal first-period quantity is strictly lower than the optimal first-period quantity under one-sided learning; that is, $Q^*_{1,2s} < Q^*_{1,1s}$.

(ii) If the proportion of innovators is sufficiently small, the optimal first-period quantity is also strictly lower than the myopic-optimal first-period quantity; that is, there exists $\bar{\gamma} \in (0,1)$ *such that if* $\gamma < \bar{\gamma}$ *then* $Q^*_{1,2s} < Q^*_{1,m}$.

Both statements of the theorem are a direct consequence of Proposition 5. To understand the first statement, recall from the analysis of §5 that when learning is one-sided, future considerations cause the firm to increase its quantity decision relative to the myopic optimal quantity $Q^*_{1,m}$, owing to the combined effects of the inventory and information channels, both of which provide incentives for increasing the first-period production quantity (see Propositions 2 and 3). When learning is two-sided, the inventory channel still "pushes" the firm to increase its quantity relative to $Q^*_{1,m}$, but now the information channel "pulls" in the opposite direction (see Proposition 5); the end result is $Q^*_{1,2s} < Q^*_{1,1s}$.

Consider next the second statement of Theorem 1. To understand this result, it is helpful to compare the preferred policy of a firm operating under one-sided learning against that of a firm operating under two-sided learning, as we scale down the proportion of innovators γ. Observe that as γ approaches zero, the firm's expected first-period profit as well as the effects of the inventory channel vanish, leaving only the information channel to determine how the firm chooses its first-period quantity. From Propositions 3 and 5 it then follows that

$$\lim_{\gamma \to 0} Q^*_{1,1s} = Q_{\max}, \text{ while } \lim_{\gamma \to 0} Q^*_{1,2s} = 0.$$

In either case, the firm chooses the first-period quantity to optimize the learning process. For a firm operating under one-sided learning, this implies choosing a policy that enables perfect learning in all possible realizations of product value v. By contrast, for a firm operating under two-sided learning, optimizing the learning process in fact implies choosing a policy that prevents *any* learning from occurring.

More generally, for the firm to choose a first-period quantity higher than $Q^*_{1,m}$, it must be the case that the impact of the inventory channel dominates that of the information channel; in other words, if the inventory channel is sufficiently "weak," then the optimal quantity satisfies $Q^*_{1,2s} < Q^*_{1,m}$. One sufficient condition for the latter is that γ is sufficiently low, because the optimal first-period quantity decision is upper bounded by γ, and therefore the total inventory cost-savings cannot exceed γc (at the same time, the impact of the information channel increases in the size of the follower market, $1-\gamma$). Although not part of our model, obvious alternative sufficient conditions include high inventory holding costs and inventory perishability.

To complement the result of Theorem 1, we present the numerical experiments of Figure 1. We model product value as a Beta-distributed random variable with parameters (a_v, b_v), where $a_v \in [1.5, 3]$ and $b_v = 4.5 - a_v$, thus capturing uncertainty distributions ranging from right- to left-skewed. We plot the optimal first-period quantity under one- and two-sided learning $Q^*_{1,1s}$ and $Q^*_{1,2s}$, as well as the myopic-optimal quantity $Q^*_{1,m}$ for four combinations of our model parameters with relatively low and high production costs c, and low and high proportion of innovators γ.

From these plots, we note that (i) as suggested by Proposition 4, the optimal quantity under one-sided learning $Q^*_{1,1s}$ is always higher than the myopic-optimal quantity $Q^*_{1,m}$, (ii) as suggested by Theorem 1, the optimal quantity under two-sided learning $Q^*_{1,2s}$ is always lower than $Q^*_{1,1s}$, and (iii) as also suggested by Theorem 1, $Q^*_{1,2s}$ is often lower than both $Q^*_{1,1s}$ and $Q^*_{1,m}$, and in these cases the firm typically chooses to forego essentially all potential first-period profit.[7] By comparing the plots of Figure 1 in the vertical direction, we observe that the lower the proportion of innovators γ, the smaller the firm's first-period quantity decision relative to the size of the first-period market; this is consistent with Theorem 1, which suggests a positive relationship between γ and $Q^*_{1,2s}$. By comparing the plots of Figure 1 in the horizontal direction, we find that the optimal quantity decision is also lower when the production cost c is higher, consistent with standard newsvendor intuition. The above comparisons also lead to the conclusion that the firm tends to operate under the second regime of Theorem 1 (where $Q^*_{1,1s} < Q^*_{1,m}$) when the market is not too optimistic about the product's value and when production costs are relatively high.

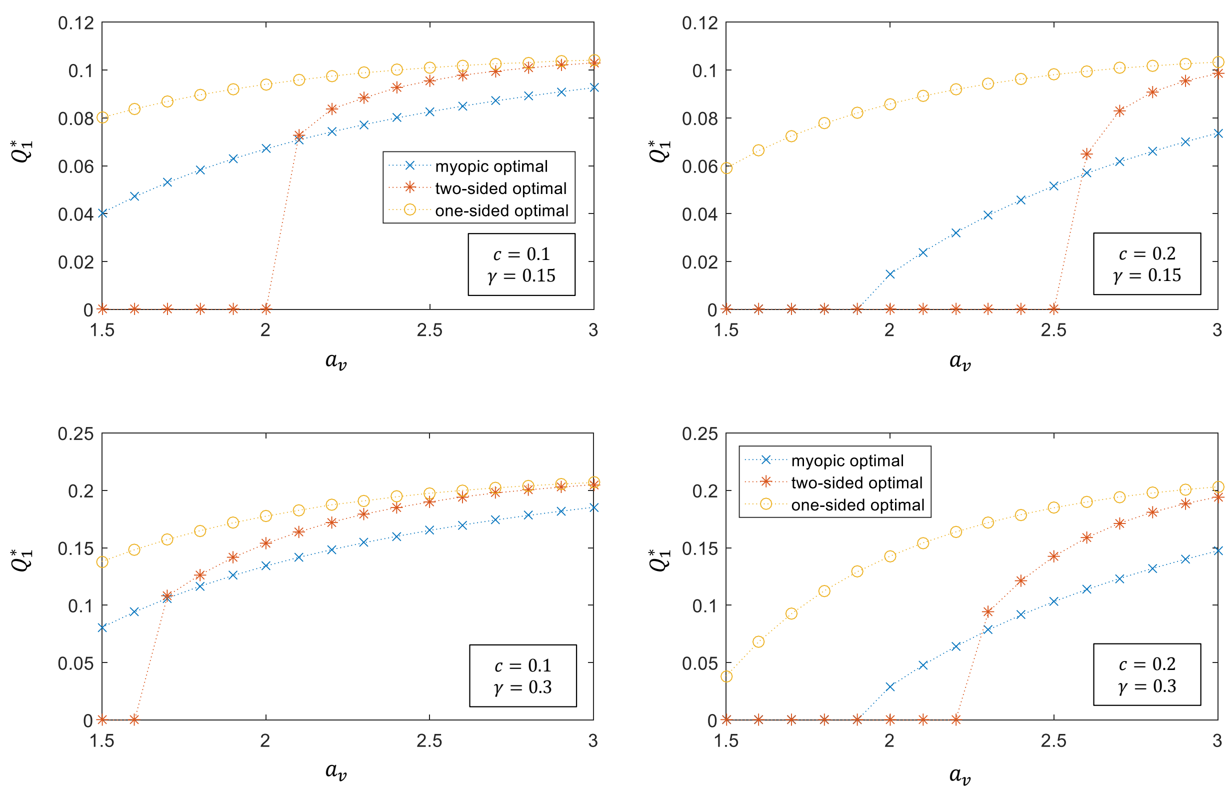

Figure 1 **Optimal first-period quantities under one- and two-sided learning. Parameter values:** $p = 0.3$**,** F **Beta-distributed with** $b_v = 4.5 - a_v$**.**

[7] Rather than refusing to serve the innovators completely, this observation is in fact more consistent with a limited release; in particular, it can be shown that a policy with Q_1 small but positive dominates the policy $Q_1 = 0$.

The experiments of Figure 2 explore in more detail the value of the threshold $\bar{\gamma}$ used in Theorem 1(ii) to describe the critical proportion of innovators below which the firm's optimal quantity choice $Q^*_{1,2s}$ is lower than the myopic-optimal quantity $Q^*_{1,m}$ given in Proposition 1. The plots suggest that when production costs are relatively low, the critical proportion ranges from approximately 10% to 40% with lower values corresponding to environments where the market is more optimistic about the product's value. When costs are relatively high, the critical proportion range is approximately 10% to 100%, while we further observe that when the value distribution is right-skewed, the optimal policy involves a limited release irrespective of the innovator proportion γ.

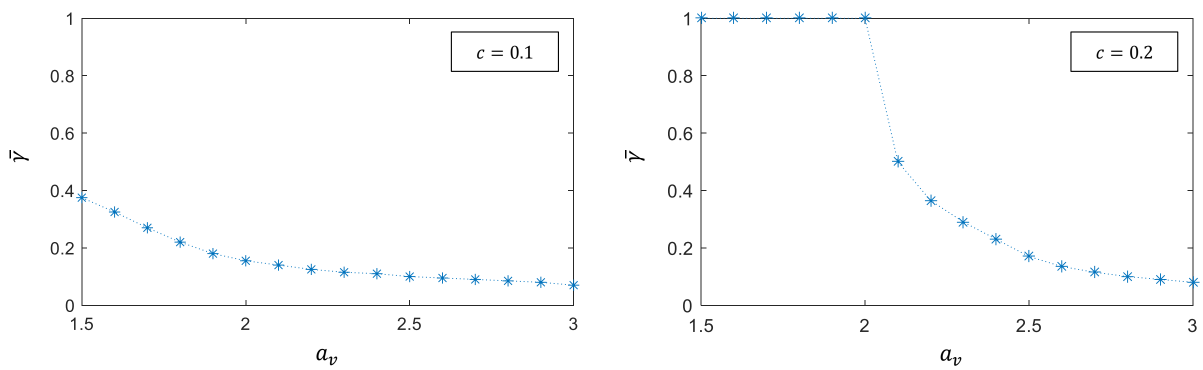

Figure 2 **Innovator threshold $\bar{\gamma}$ described in Theorem 1(ii). Parameter values: $p = 0.3$, F Beta-distributed with $b_v = 4.5 - a_v$.**

To conclude this section, it is important to reiterate that the result of Theorem 1, and the second statement of the result in particular, stands in contrast to existing work that supports the optimality of an increase in the newsvendor quantity in the presence of learning. The source of the difference is the two-sided nature of the learning process in our model, and the associated negativeness of the value of information for the firm (in §8, we demonstrate the robustness of this result under alternative model specifications). As a whole, these findings serve as a caution that not all learning environments should be approached by firms in the same fashion, and underline the importance of understanding the implications of operational decisions for learning on the demand side before choosing the policy to be implemented.

7. Profit Implications

The preceding analysis demonstrates that the firm's optimal quantity decisions in a newsvendor context can differ significantly depending on the learning regime in which the firm operates. In this section, we investigate the implications of two-sided learning for firm profit. We focus on two relevant questions. First, under which learning regime is the firm better off ex ante? This question may have implications for the firm's efforts to inform the consumers about unobservable product features ahead of the selling season. Second, by how much does firm suffer, in terms of profit, if it fails to appropriately internalize the learning regime in which it operates?

To answer the first question, we perform a direct comparison of expected firm profit under the two learning regimes. Recall that Proposition 4 (Theorem 1) suggests that when learning is one-sided (two-sided), the firm takes actions in order to increase (decrease) learning. However, it is not immediately obvious under which learning regime the firm achieves higher profit; our next result addresses this question.

PROPOSITION 6. *The firm's total expected profit is strictly higher when learning is two-sided.*

It is interesting to observe that while the firm in some sense "dislikes" learning in the two-sided regime and "likes" learning in the one-sided regime, it still prefers to operate in a two-sided learning environment. To see why this is the case, consider the following thought experiment. Suppose that the firm stocks the same first-period quantity in two separate environments that differ only in the learning regime. Under either regime, the firm's first-period profit is identical, and if no stockout occurs in the first period, then the same is true for the firm's second-period profit (since both the firm and the consumers become fully informed). Now consider scenarios where a stockout occurs in the first period. Under one-sided learning, the consumers know v but the firm does not, and it follows from Proposition 3 that the firm would be strictly better off if it was also informed about v (because the value of one-sided information is positive). However, even if this were the case, from Proposition 5 it follows that the firm would be strictly better off if neither the consumers or itself were informed (because the value of two-sided information is negative); the latter is exactly the case in the event of a stockout under two-sided learning.[8]

Having established that two-sided learning represents a more profitable environment for the firm, we next consider to what extent the firm's profit suffers if it does not appropriately account for the two-sidedness of the learning process. The following result is a qualitative one that compares the performance of policies that either ignore learning altogether, or (erroneously) assume that learning is one-sided.

[8] For the consumers, two-sided learning typically introduces surplus losses (relative to one-sided learning) for both the innovators and the followers. For the innovators, the two learning regimes are equivalent informationally, but the firm's decreased quantity decision under two-sided learning results in decreased access to the product in the first period. For the followers, the imperfect learning outcomes that occur under high value scenarios lead to either erroneous purchase decisions (when the updated value expectation is higher than the product's true value, $\hat{v}(Q_1) > v$) or to erroneous non-purchase decisions (when the updated value expectation is lower than the product's true value, $\hat{v}(Q_1) < v$); these losses typically outweigh the surplus losses that occur under one-sided learning in the highest value scenarios where the second-period demand is undersupplied by the firm.

PROPOSITION 7. *When learning is two-sided, if the proportion of innovators γ is sufficiently small, then the myopic-optimal policy $Q^*_{1,m}$ results in strictly higher expected profit than the one-sided-optimal policy $Q^*_{1,1s}$; that is, $\pi_{t,2s}(Q^*_{1,1s}) < \pi_{t,2s}(Q^*_{1,m}) < \pi_{t,2s}(Q^*_{1,2s})$.*

Proposition 7 suggests that when learning is two-sided and the proportion of innovators is small, the myopic policy outperforms the one-sided optimal policy—interestingly, adjusting the quantity decision to account for the wrong type of learning regime is worse than ignoring learning altogether. The numerical experiments of Figure 3 suggest that the result often holds even when the proportion of innovators is relatively large. Moreover, from these experiments we also observe that failing to account appropriately for learning can lead to substantial profit losses for the firm, reaching up to 36% in Figure 3; in particular, the largest profit losses occur when the optimal two-sided policy differs substantially from the myopic and one-sided optimal policies, which tends to occur for value distributions ranging from symmetric to right-skewed.

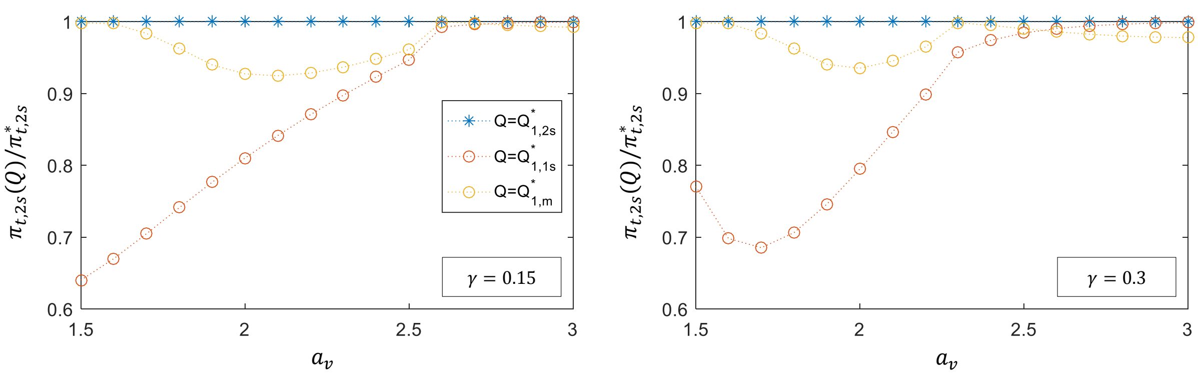

Figure 3 **Fraction of expected profit achieved in a two-sided learning environment under policies $Q^*_{1,2s}$, $Q^*_{1,1s}$, $Q^*_{1,m}$. Parameter values: $p = 0.3$, $c = 0.2$, F Beta-distributed with $b_v = 4.5 - a_v$.**

8. Model Extensions

In this section we present three extensions of our main model, with the goal of establishing the robustness of our insights with respect to some of our model assumptions. In each extension we focus on establishing riobustness of the basic observation that the marginal value of two-sided information is negative (i.e., Proposition 5 in our main analysis), which in turn provides an incentive for the firm to decrease the first-period quantity.

8.1. Learning From Reviews

We consider first an extension of our model where consumer learning occurs on the basis of product reviews (as opposed to product sales). To model learning from product reviews, we use the modeling approach proposed in Bergemann and Välimäki (1997) and used more recently in Yu et al. (2015). We assume that there are two possible states of the world $\theta \in \{0,1\}$, such that when the state is $\theta = 0$ the product's value is $\underline{v}$, while when the state is $\theta = 1$ the product's value is $\bar{v} > \underline{v}$. The firm and all the consumers share a common prior belief over the state $P_0(\theta = 1) =: b_1$, and we assume $\underline{v} > p$ so that for any prior belief $b_1 \in (0,1)$ there is at least some positive demand for the product in the first period. After purchasing the product, the first-period buyers generate product reviews which are aggregated into a market signal $\zeta \sim N(n\theta, \sigma\sqrt{n})$, where n is the volume of product reviews (assumed to be equal to the volume of first-period purchases), and σ is a parameter that captures review informativeness (for more details on the described modeling approach, see §2 in Bergemann and Välimäki (1997) and §2.2 in Yu et al. (2015)). The second-period consumers observe the review signal and update their beliefs via Bayes' rule before making their own purchase decision.

In this model, notice first that demand for the product in the first period is given by $D_1(b_1) = \gamma\left[1 - \frac{p}{b_1\underline{v} + (1-b_1)\bar{v}}\right]$, so that the sales achieved by the firm in the first period are equal to $S_1 = \min\{Q_1, Q_{\max}\}$, where Q_1 is the firm's first-period quantity decision and $Q_{\max} = D_1(b_1)$. The following proposition establishes that our basic result on the marginal expected value of information persists in the model with learning from product reviews.

PROPOSITION 8. *When learning is two-sided and occurs on the basis of product reviews, the marginal expected value of information* $\frac{\partial \pi_{2I}^*}{\partial Q_1}$ *is strictly negative for all* $Q_1 \in [0, Q_{\max}]$.

The firm's quantity decision Q_1 determines the volume of sales S_1 achieved in the first period, which in turn is equal to the volume of product reviews n made available to consumers in the second period. Thus, the result of Proposition 8 suggests that all else being equal, the firm prefers fewer product reviews or, equivalently, less learning. The proof of the result relies on a different approach to that of Proposition 5, but the core idea remains that of optimally controlling (i.e., restricting) the information transmission from the innovators to the followers, as described in §6. This demonstrates that the informational considerations associated with the first-period quantity decision are qualitatively similar to those in our main model, pushing the firm to decrease its first-period quantity decision. Intuitively, these considerations become stronger as product reviews become more informative and therefore more influential in shaping consumer beliefs.

8.2. Imperfect Experts

We consider next an extension of our model where the expert consumers are imperfect in their ex ante assessment of the product. Specifically, we assume that before making a purchase decision, each expert consumer i receives an informative signal $s_i = v + \epsilon_i$, where v is the product's underlying value and $\epsilon_i \sim N(0, \sigma_i)$ (our main model thus corresponds to the limit case $\sigma_i \to 0$). That is, in this model the assessments of the expert consumers are correct on average, although the assessments of individual experts may differ. When this is the case, the volume of sales that occur in the first period serves as a market signal that aggregates the information contained in the individual assessments of the expert consumers. The following proposition suggests that our model insights are qualitatively robust with respect to imperfections in the experts' product assessments.

Proposition 9. *When learning is two-sided and the experts are imperfect in their assessments of product value, the marginal expected value of information* $\frac{\partial \pi_{2I}^*}{\partial Q_1}$ *is strictly negative for all* $Q_1 \in [0, Q_{\max}]$.

After receiving the signal s_i, each first-period consumer has an updated value expectation $\mu(s_i)$ which is strictly increasing in s_i (this follows from the strict monotone likelihood property satisfied by the signal s_i; see Milgrom (1981)). Thus, demand for the product in the first period $D_1(v)$ is strictly increasing in the product's value v. The learning process described in §4.2 remains qualitatively intact, with the only difference being that the threshold function $v_c(Q_1)$ given in (3) must now be replaced by an alternative function $v_{c2}(Q_1)$. To establish that the value of two-sided information remains negative for the firm, it then suffices to point out that the function $v_{c2}(Q_1)$ retains strict monotonicity with respect to Q_1, so that reducing the first-period quantity results in decreased two-sided learning and increased second-period profits π_{2I}^* associated with the information channel, as described in §6.

8.3. Imperfect Sales Feedback

We now consider the case where sales feedback from the first period does not always become observable to the consumers in the second. Specifically, we assume that consumers in the second period become aware of the events that occurred in the first period with probability ζ (our main model corresponds to the limit case $\zeta = 1$). When $\zeta < 1$, the dependence of the second-period consumers' purchase decisions on the firm's first-period quantity decision is weaker. Nevertheless,

Proposition 10. *When learning is two-sided and the followers observe first-period sales with probability* $\zeta > 0$*, the marginal expected value of information* $\frac{\partial \pi_{2I}^*}{\partial Q_1}$ *is strictly negative for all* $Q_1 \in [0, Q_{\max}]$.

That is, as long as the events of the first period are observed by consumers in the second period with some positive probability, the firm still has an incentive to reduce its first-period quantity decision, although we note that the strength of this incentive clearly increases with ζ. Consider first the extreme case where $\zeta = 0$. In this case, the followers remain uninformed in the second period, irrespective of the firm's actions in the first period. Thus, the informational channel has no bearing on the firm's decisions, and the firm's first-period quantity is higher than the myopic quantity $Q^*_{1,m}$ as a result of the effects of the inventory channel. As ζ increases from zero, the effect of the information channel grows stronger, and the firm's incentive to reduce its early quantity decision increases.

9. Conclusion

This paper has focused on highlighting the relative differences between two learning regimes in the context of the classic newsvendor problem. The first is the "one-sided" learning regime, where the firm's current decisions generate information on demand-side characteristics that do not affect the future demand faced by the firm. The second is the "two-sided" learning regime, where the firm's current decisions generate information on product characteristics that directly affect future demand. Essentially, under one-sided learning, future demand for the product is exogenous to the firm's learning process, while under two-sided learning future demand is endogenous to it.

We have argued that the firm's quantity decisions should be viewed in a different light depending on the learning regime in which the firm operates. When learning is one-sided, the firm's policy is a tool for *information extraction* and should be approached in a more-information-is-better fashion. By contrast, when learning is two-sided, the firm's policy is a tool for *information transmission*, and the firm's approach often needs to be reversed. As a consequence, we have established that the well-known prescription to "stock more" should be tempered when learning is two-sided, and may sometimes cease to apply altogether. We have further shown that a firm operating under two-sided learning is able to extract higher expected profits, and that the loss for the firm from failing to account for the appropriate learning regime when planning its operations can be substantial.

In closing, we note that the high-level prescription of stocking less in environments where two-sided learning is prevalent appears to be consistent with the perceived policy of undersupplying the early demand implemented by technology companies such as Apple and Xiaomi. The existing literature and popular press have proposed a number of potential benefits associated with early supply shortages.[9] Many such benefits leverage behavioral/emotional aspects of the consumers' decision-making process; for instance, deliberate supply shortages have been previously associated with "hunger marketing," a strategy intended to trigger a psychological response to scarcity that causes an increase in future demand (e.g., *tech.co* (2015)), while in the presence of peer-to-peer consumer communications, Papanastasiou et al. (2016) highlight that supply shortages induce customer self-selection which can result in favorably biased word-of-mouth. Our paper offers an

[9] We note that rather than forming part of a deliberate policy, stockouts may occur for a variety of reasons including inaccurate demand forecasts, capacity constraints, production lead times, etc.

alternative explanation for the success of understocking, with a basis in rational decision making: Stocking less in the early stages of the selling season may allow the firm to optimally control the information transmission between early and late adopters. In this respect, the mechanism identified in our paper is related to the signaling theory proposed by Stock and Balachander (2005), whereby the firm uses supply shortages as a device to communicate the value of its product to the consumers. Among other differences between the two, the signaling mechanism assumes that the firm has private knowledge of the product's value for the consumers (which implies knowledge of product attributes and performance, as well as consumer tastes) and may therefore be more applicable to cases of more mature technologies and/or less innovative experience goods. By contrast, our information transmission mechanism assumes that the firm faces uncertainty about the product's value for the consumers (either owing to uncertainty about product attributes and performance, or consumer tastes, or both) and may therefore be more applicable to cases of new technologies and/or more innovative goods.

Appendix

A. Analysis of the Firm's Second-Period Problem

A.1. One-Sided Learning

Suppose the firm's quantity decision in the first period is Q_1, and let the number of units leftover at the end of the first period be denoted by $Q_L(Q_1, v) := [Q_1 - D_1(v)]^+$. If a stockout did not occur in the first period (i.e., if $v < v_c(Q_1)$), the firm has a positive amount of leftover inventory and is able to infer the product's value perfectly. The firm's profit in the second period is

$$\pi_{2l}(Q_{2l}) = p \min\left\{(1-\gamma)\left[1-\frac{p}{v}\right]^+, Q_L(Q_1, v) + Q_{2l}\right\} - cQ_{2l}, \tag{6}$$

and the optimal quantity decision is $Q_{2l}^* = \left[(1-\gamma)\left(1-\frac{p}{v}\right) - Q_L(Q_1, v)\right]^+$. That is, if there are enough leftover units to satisfy all second-period demand, there is no additional production in the second period; in the opposite the case, the firm produces as many additional units as are required to meet the demand.

On the other hand, if a stockout occurred in the first period ($v \geq v_c(Q_1)$), there are no leftover units and the firm can only infer the product's value up to $v \geq v_c(Q_1)$. The firm's second-period profit is given by

$$\pi_{2h}(Q_{2h}) = pE\left[\min\left\{(1-\gamma)\left[1-\frac{p}{v}\right]^+, Q_{2h}\right\} \;\middle|\; v \geq v_c(Q_1)\right] - cQ_{2h}. \tag{7}$$

Observe that this is essentially the one-period newsvendor model of §4.1, where we can replace F in Proposition 1 with the appropriate lower-truncated cdf to get

$$Q_{2h}^* = (1-\gamma)\left[1-\frac{p}{v_{2c}(Q_1)}\right]^+, \text{ where } v_{2c}(Q_1) = F^{-1}\left[\frac{p-(1-F[v_c(Q_1)])c}{p}\right] \tag{8}$$

and $v_c(Q_1)$ is given in (3). Lemma 2 catalogues the optimal second-period policy.

LEMMA 2. *When learning is one-sided:*

(i) If a stockout does not occur in the first period ($v < v_c(Q_1)$), the optimal second-period quantity is $Q^*_{2l} = \left[(1-\gamma)\left(1-\frac{p}{v}\right) - Q_L(Q_1, v)\right]^+$.

(ii) If a stockout occurs in the first period ($v \geq v_c(Q_1)$), the optimal second-period quantity is $Q^*_{2h} = (1-\gamma)\left[1 - \frac{p}{v_{2c}(Q_1)}\right]^+$, where $v_{2c}(Q_1) = F^{-1}\left[\frac{p-(1-F[v_c(Q_1)])c}{p}\right]$.

We may then write down the firm's optimal expected second-period profit under one-sided learning as

$$\pi^*_{2,1s}(Q_1) = \int_{v_l}^{v_c(Q_1)} \pi_{2l}(Q^*_{2l}; v) f(v) dv + \int_{v_c(Q_1)}^{v_h} \pi_{2h}(Q^*_{2h}; v) f(v) dv$$

where,

$$\pi_{2l}(Q^*_{2l}; v) = p(1-\gamma)\left[1 - \frac{p}{v}\right]^+ - c\left[(1-\gamma)\left(1-\frac{p}{v}\right) - Q_L(Q_1, v)\right]^+$$
$$\pi_{2h}(Q^*_{2h}; v) = p\min\left\{(1-\gamma)\left(1-\frac{p}{v}\right), (1-\gamma)\left(1-\frac{p}{v_{2c}(Q_1)}\right)\right\} - c(1-\gamma)\left(1-\frac{p}{v_{2c}(Q_1)}\right).$$

A.2. Two-Sided Learning

Suppose the firm's first-period quantity decisions is Q_1. Note first that if a stockout does not occur in the first period ($v < v_c(Q_1)$), the firm and the consumers are able to infer the product's value perfectly. In these scenarios the firm's problem is equivalent to that under one-sided learning (under one-sided learning, only the firm is initially uninformed about v; if a stockout does not occur in the first period, then the firm is able to perfectly infer v, so that both the firm and the consumers are informed in the second period, as is the case under two-sided learning). Thus, the firm's profit function when no stockout occurs is again

$$\pi_{2l}(Q_{2l}) = p\min\left\{(1-\gamma)\left[1-\frac{p}{v}\right]^+, Q_L + Q_{2l}\right\} - cQ_{2l},$$

and the optimal quantity decision is $Q^*_{2l} = \left[(1-\gamma)\left(1-\frac{p}{v}\right) - Q_L\right]^+$, as described in §A.1.

The difference between the two learning regimes emerges in scenarios where value is relatively high and a stockout occurs in the first period ($v \geq v_c(Q_1)$). As in the case of one-sided learning, the firm in these scenarios can only infer value up to $v \geq v_c(Q_1)$; however, this is now also the case for the consumers it faces. The firm's profit function becomes

$$\pi_{2h}(Q_{2h}) = p\min\left\{(1-\gamma)\left[1-\frac{p}{\hat{v}(Q_1)}\right]^+, Q_{2h}\right\} - cQ_{2h}, \tag{9}$$

where recall that $\hat{v}(Q_1)$ is the firm and the consumers' updated value expectation following a first-period stockout at quantity Q_1 (see (2)). Notice that the firm's problem when learning is two-sided is simpler, in the sense that the firm no longer faces demand uncertainty in the second period. Since the firm and the consumers hold identical value beliefs, the firm can perfectly predict and act to fulfill all second-period demand. Accordingly, the optimal quantity decision is

$$Q^*_{2h} = (1-\gamma)\left[1-\frac{p}{\hat{v}(Q_1)}\right]^+, \tag{10}$$

and the preceding discussion is summarized in Lemma 3.

Lemma 3. *When learning is two-sided:*

(i) If a stockout does not occur in the first period ($v < v_c(Q_1)$), the optimal second-period quantity is $Q^*_{2l} = \left[(1-\gamma)\left(1-\frac{p}{v}\right) - Q_L\right]^+$.

(ii) If a stockout occurs in the first period ($v \geq v_c(Q_1)$), the optimal second-period quantity is $Q^*_{2h} = (1-\gamma)\left[1-\frac{p}{\hat{v}(Q_1)}\right]^+$, where $\hat{v}(Q_1)$ is given in (2).

Comparing Lemmas 2 and 3, the firm's optimal second-period policy under the two learning regimes is identical when a stockout does not occur in the first period, while when a stockout occurs, $v_{2c}(Q_1)$ in (8) is replaced by $\hat{v}(Q_1)$. As in §A.1, the above expressions can then be used to write down the firm's optimal expected second-period profit under two-sided learning, $\pi^*_{2,2s}(Q_1)$.

B. Endogenous Pricing

In this section we consider the case where the product's price is also a decision variable for the firm. The following corollary is a straightforward extension of Theorem 1.

Corollary 1. *When learning is two-sided, if γ is sufficiently small the optimal price p^*_{2s} and first-period quantity $Q^*_{1,2s}$ satisfy*

$$Q^*_{1,2s} < \gamma\left(1 - \frac{p^*_{2s} - a}{v_r}\right), \text{ where } v_r = F^{-1}\left[\frac{p^*_{2s} - c}{p^*_{2s}}\right]. \tag{11}$$

To arrive at the above result, we consider any price (including the optimal one) and then search for the corresponding optimal quantity decision. Since the result of Theorem 1 holds for any arbitrary price, we next note that a sufficiently small γ ensures that the optimal quantity must be less than the myopic-optimal one given in Proposition 1. While it is possible to draw some conclusions on the value of the optimal price relative to the optimal quantity, perhaps the more interesting question is to what extent the firm's optimal quantity policy (and the associated comparison with the myopic and one-sided policies) is affected by endogenizing the pricing decision. In Figure 4, we conduct experiments where we solve the firm's price-and-quantity problem for a range of value distributions. Overall, the observations from these experiments are consistent with those of the exogenous-price analysis; that is, endogenizing the pricing decision does not appear to affect the firm's optimal quantity decision in any qualitative way. Furthermore, we note that the optimal price under one- and two-sided learning is quite similar and higher than the myopic-optimal price, and the main point of differentiation in the firm's optimal policy is the quantity decision.

C. Proofs

Proof of Proposition 1

Profit in the first period is

$$\begin{aligned}\pi_1(Q_1) &= pE[\min\{\gamma D(v), Q_1\}] - cQ_1 \\ &= p\int_{v_l}^{v_h}[\min\{\gamma D(v), Q_1\}]f(v)dv - cQ_1,\end{aligned}$$

where $D(v) = 1 - G\left(\frac{p}{v}\right) = \left[1-\frac{p}{v}\right]^+$. Recall $v_c(Q_1)$ in (3) is the critical product value that induces a stockout in the first-period when the quantity decision is Q_1 for $Q_1 \in [0, Q_{\max}]$. Then

$$\pi_1(Q_1) = pE[\min(\gamma D(v), Q_1)] - cQ_1$$

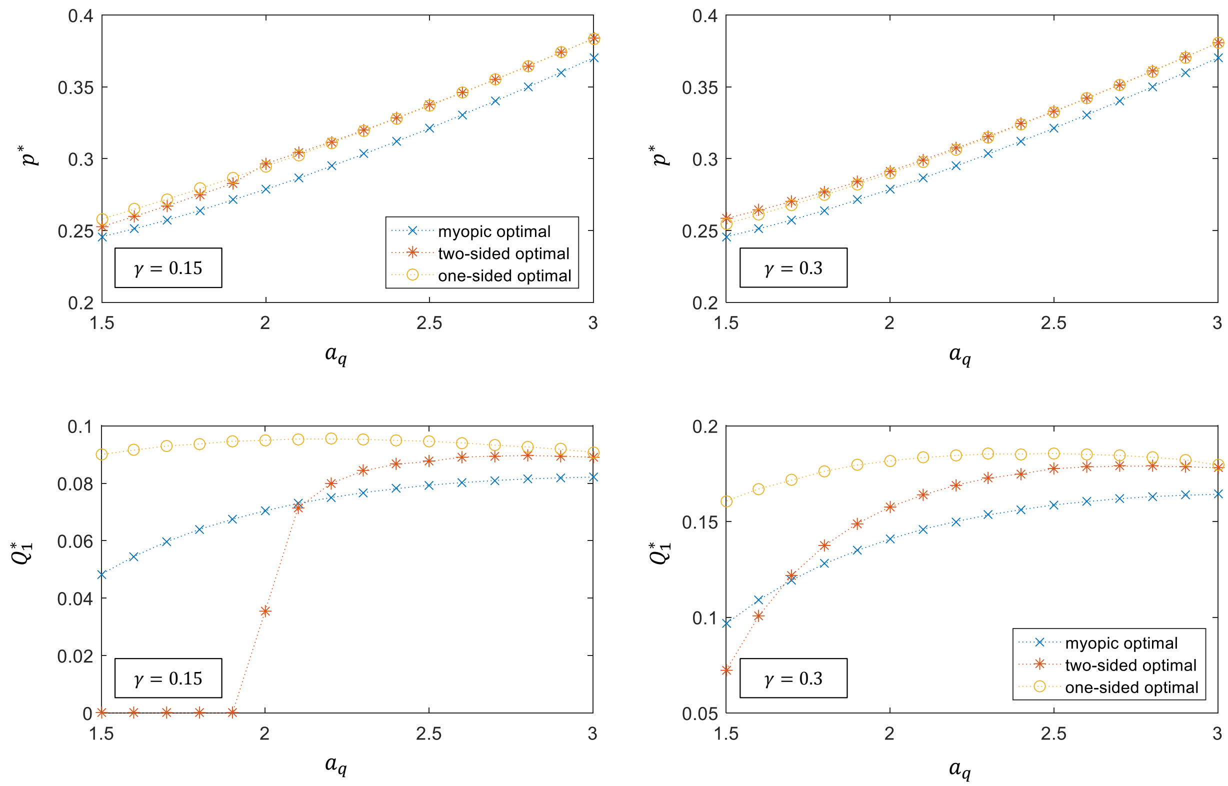

Figure 4 **Optimal prices and first-period quantities under one- and two-sided learning. Parameter values:** $c = 0.1$, F **Beta-distributed with** $b_q = 4.5 - a_q$.

$$= p\left(\int_{v_l}^{v_c(Q_1)} \gamma D(v)f(v)dv + \int_{v_c(Q_1)}^{v_h} Q_1 f(v)dv\right) - cQ_1$$

$$= p\left(\int_{v_l}^{v_c(Q_1)} \gamma D(v)f(v)dv + Q_1\bar{F}(v_c(Q_1))\right) - cQ_1$$

The first-order derivative with respect to Q_1 is

$$\frac{d\pi_1}{dQ_1} = p\left(\gamma D(v_c(Q_1))f(v_c(Q_1))\frac{dv_c}{dQ_1} + \bar{F}(v_c(Q_1)) - Q_1 f(v_c(Q_1))\frac{dv_c}{dQ_1}\right) - c$$

By definition of $v_c(Q_1)$, we have $\gamma D(v_c(Q_1)) = Q_1$ so that

$$\frac{d\pi_1}{dQ_1} = p\bar{F}(v_c(Q_1)) - c \tag{12}$$

Note that $v_c(Q_1)$ is strictly increasing in $Q_1 \in [0, Q_{\max}]$ so that $\bar{F}(v_c(Q_1))$ is strictly decreasing and therefore so is $\frac{d\pi_1}{dQ_1}$. It follows that $\pi_1(Q_1)$ is unimodal in $Q_1 \in [0, Q_{\max}]$. We next note that $\frac{d\pi_1}{dQ_1}|_{Q_1 = Q_{\max}} < 0$, and that (12) has a unique root at Q_1 such that

$$v_c(Q_1) = F^{-1}\left[\frac{p-c}{p}\right] =: v_m,$$

assuming such a Q_1 exists in the interval $[0, Q_{\max}]$. If it does exist, then the optimal quantity is given by $Q^*_{1,m} = \gamma\left(1 - \frac{p}{v_m}\right)$, while if it does not exist, then $\pi_1(Q_1)$ is strictly decreasing in $Q_1 \in [0, Q_{\max}]$ and the optimal quantity is $Q^*_{1,m} = 0$. Hence, we have

$$Q^*_{1,m} = \gamma D\left(v_m\right) = \gamma\left[1 - \frac{p}{v_m}\right]^+,$$

Proof of Lemma 2

See the analysis preceding the result in the main text.

Proof of Proposition 2

Let $D_1(v) = \gamma\left[1 - \frac{p}{v}\right]^+$ and $D_2(v) = (1-\gamma)\left[1 - \frac{p}{v}\right]^+$. We have

$$\pi^*_{2C}(Q_1) := c\int_{v_l}^{v_c(Q_1)} \min\left\{D_2(v), Q_L(Q_1, v)\right\} f(v)dv$$

Note that for any Q_1, $D_2(v)$ is nondecreasing in v, while Q_L is nonincreasing v. Therefore, we have

$$\pi^*_{2C}(Q_1) := c\int_{v_l}^{v_w} D_2(v)f(v)dv + c\int_{v_w}^{v_c}(Q_1 - D_1(p,v))f(v)dv$$

for some v_w such that $D_2(v_w) = Q_1 - D_1(v_w) > 0$. Taking derivative with respect to Q_1,

$$\frac{\partial \pi^*_{2C}}{\partial Q_1} = c\frac{\partial v_w}{\partial Q_1}D_2(v_w)f(v_w) - c\frac{\partial v_w}{\partial Q_1}(Q_1 - D_1(v_w))f(v_w) + c\frac{\partial v_c}{\partial Q_1}(Q_1 - D_1(v_c))f(v_c) + c\int_{v_w}^{v_c} f(v)dv.$$

By definition of v_w we have $D_2(v_w) = Q_1 - D_1(v_w)$, and by definition of v_c we have $D_1(v_c) = Q_1$, so that

$$\frac{\partial \pi^*_{2C}}{\partial Q_1} = c\int_{v_w}^{v_c} f(v)dv$$

Since $v_w < v_c$, it follows that $\frac{\partial \pi^*_{2C}}{\partial Q_1} > 0$.

Proof of Proposition 3

Under one-sided learning, we have for $D_2(v) = (1-\gamma)[1 - \frac{p}{v}]^+$

$$\begin{aligned}\pi^*_{2I} &= (p-c)\int_{v_l}^{v_c(Q_1)} D_2(v)f(v)dv + p\int_{v_c(Q_1)}^{v_h} \min\{D_2(v), Q^*_{2h}(Q_1)\}f(v)dv - cQ^*_{2h}(Q_1)\\ &= (p-c)\int_{v_l}^{v_c(Q_1)} D_2(v)f(v)dv + p\int_{v_c(Q_1)}^{v_r(Q^*_{2h}(Q_1))} D_2(v)f(v)dv + Q^*_{2h}(Q_1)[p\bar{F}(v_r(Q^*_{2h}(Q_1))) - c],\end{aligned}$$

for some $v_c(Q_1) < v_r(Q^*_{2h}(Q_1)) < v_h$. From the proof of Proposition 1, we have that at the optimal second-period quantity $Q^*_{2h}(Q_1)$ is such that $[p\bar{F}(v_r(Q^*_{2h}(Q_1))) - c] = 0$, so that the above reduces to

$$\pi^*_{2I} = (p-c)\int_{v_l}^{v_c(Q_1)} D_2(p,v)f(v)dv + p\int_{v_c(Q_1)}^{v_r(Q^*_{2h}(Q_1))} D_2(p,v)f(v)dv.$$

Note that there is a one-to-one correspondence between $v_c(Q_1)$ and Q_1 (these are positively related) and a one-to-one correspondence between $v_c(Q_1)$ and $v_r(Q^*_{2h}(Q_1))$ (these are also positively related). It is convenient to treat $v_c(Q_1)$ directly as the decision variable. We then have

$$\begin{aligned}\frac{\partial \pi^*_{2I}}{\partial v_c} &= (p-c)D_2(p,v_c)f(v_c) - pD_2(p,v_c)f(v_c) + p\frac{\partial v_r}{\partial v_c}D_2(p,v_r)f(v_r)\\ &= -cD_2(p,v_c)f(v_c) + p\frac{\partial v_r}{\partial v_c}D_2(p,v_r)f(v_r) \qquad (13)\end{aligned}$$

Note next that v_r satisfies $F_t(v_r) = \frac{p-c}{p}$, where F_t is the cdf of v truncated from below at v_c (see proof of Proposition 1). Therefore, we have

$$F_t(v_r) = \frac{F(v_r) - F(v_c)}{1 - F(v_c)} = \frac{p-c}{p}$$
$$F(v_r) = F(v_c) + \frac{p-c}{p}(1 - F(v_c))$$

Differentiating with respect to v_c,

$$f(v_r)\frac{\partial v_r}{\partial v_c} = f(v_c)\frac{c}{p},$$

and inserting into (13) we get

$$\frac{\partial \pi_{2I}^*}{\partial v_c} = cf(v_c)[D_2(p, v_r) - D_2(p, v_c)] > 0,$$

where the last inequality follows from $v_r > v_c$. Furthermore, since the consumers' type distribution is the standard uniform distirbution, the last expression simplifies to

$$\begin{aligned}\frac{\partial \pi_{2I}^*}{\partial v_c} &= cf(v_c)(1-\gamma)\left[\frac{p}{v_c} - \frac{p}{v_r}\right] \\ &= cf(v_c)(1-\gamma)p\left[\frac{v_r - v_c}{v_r v_c}\right]\end{aligned}$$

Proof of Proposition 4

The first-order derivative of the profit function (5) with respect to Q_1 is

$$\frac{\partial \pi_{t,1s}}{\partial Q_1} = \frac{\partial \pi_1}{\partial Q_1} + \frac{\partial \pi_{2C}^*}{\partial Q_1} + \frac{\partial \pi_{2I}^*}{\partial Q_1}.$$

The myopic-optimal quantity decision $Q_{1,m}^*$ given in Proposition 1 is such that $\frac{\partial \pi_1}{\partial Q_1}\big|_{Q_1 = Q_{1,m}^*} = 0$, and $\pi_1(Q_1)$ is unimodal so that $\frac{\partial \pi_1}{\partial Q_1}\big|_{Q_1 < Q_{1,m}^*} > 0$. Furthermore, from the Propositions 2 and 3 we have that $\frac{\partial \pi_{2C}^*}{\partial Q_1}, \frac{\partial \pi_{2I}^*}{\partial Q_1} > 0$ for all $Q_1 \in [0, Q_{\max}]$; it then follows that $Q_{1,1s}^* > Q_{1,m}^*$.

Proof of Lemma 3

See the analysis preceding the result in the main text.

Proof of Proposition 5

We prove the result for the more general case of a demand function $D(v)$ which is strictly concave in v for $v \in [\underline{v}_l, v_h]$, where $\underline{v}_l := \{\min v : v \in [v_l, v_h], D(v) \geq 0\}$ (note that this holds under our specific model assumptions, according to which $D(v) = \left[1 - \frac{p}{v}\right]^+$). We show that under this condition, $\pi_{2I}^*(Q)$ is strictly decreasing in Q for all $Q \in [0, Q_{\max}]$.

We have

$$\pi_{2I}^*(Q) = (1-\gamma)(p-c)\left[\int_{\underline{v}_l}^{v_c(Q)} D(v)f(v)dv + \int_{v_c(Q)}^{v_h} D(\hat{v}(Q))f(v)dv\right]$$

Fix some $Q_1 \in (0, Q_{\max}]$, and note that this then uniquely determines $\bar{v}_1 := v_c(Q_1)$ through (3). Now consider some Q_2 such that $Q_2 < Q_1$, which in turn determines $\bar{v}_2 := v_c(Q_2)$ and note that $\bar{v}_2 < \bar{v}_1$. Define also $\hat{v}_1 := \hat{v}(Q_1)$ and $\hat{v}_2 := \hat{v}(Q_2)$ (see (2)). We will show that if $D(v)$ is strictly concave in $v \in [\underline{v}_l, v_h]$, then $\pi_{2I}^*(Q_1) < \pi_{2I}^*(Q_2)$. We have

$$\pi_{2I}^*(Q_1) = (p-c)\left[\int_{\underline{v}_l}^{\bar{v}_1} D_2(v)f(v)dv + D_2(\hat{v}_1)\int_{\bar{v}_1}^{v_h} f(v)dv\right]$$

$$\pi_{2I}^*(Q_2) = (p-c)\left[\int_{\underline{v}_l}^{\bar{v}_2} D_2(v)f(v)dv + D_2(\hat{v}_2)\int_{\bar{v}_2}^{v_h} f(v)dv\right]$$

Since $\bar{v}_2 < \bar{v}_1$ and the two functions are therefore equivalent in the range $v \le \bar{v}_2$, it suffices to show that

$$\int_{\bar{v}_2}^{\bar{v}_1} D_2(v)f(v)dv + D_2(\hat{v}_1)\int_{\bar{v}_1}^{v_h} f(v)dv < D_2(\hat{v}_2)\int_{\bar{v}_2}^{v_h} f(v)dv. \tag{14}$$

Taking the left-hand side of the inequality, we have

$$\begin{aligned}
&\int_{\bar{v}_2}^{\bar{v}_1} D_2(v)f(v)dv + D_2(\hat{v}_1)\int_{\bar{v}_1}^{v_h} f(v)dv \\
&\qquad = [F(\bar{v}_1) - F(\bar{v}_2)]\int_{\bar{v}_2}^{\bar{v}_1} \frac{D_2(v)f(v)}{[F(\bar{v}_1) - F(\bar{v}_2)]}dv + [1 - F(\bar{v}_1)]D_2(\hat{v}_1) \\
&\qquad < [F(\bar{v}_1) - F(\bar{v}_2)]D_2(\hat{v}_{21}) + [1 - F(\bar{v}_1)]D_2(\hat{v}_1),
\end{aligned}$$

where $\hat{v}_{21} := E[v \mid \hat{v}_2 \le v \le \hat{v}_1]$ and the inequality follows from strict concavity of $D_2(v)$ by applying Jensen's inequality. Next, define $P_{21} := \frac{F(\bar{v}_1)-F(\bar{v}_2)}{1-F(\bar{v}_2)}$ and $P_1 = \frac{1-F(\bar{v}_1)}{1-F(\bar{v}_2)}$, so that

$$\begin{aligned}
[F(\bar{v}_1) - F(\bar{v}_2)]D_2(\hat{v}_{21}) + [1 - F(\bar{v}_1)]D_2(\hat{v}_1) &= (1 - F(\bar{v}_2))\left(P_{21}D_2(\hat{v}_{21}) + P_1 D_2(\hat{v}_1)\right) \\
&< (1 - F(\bar{v}_2))\, D_2(P_{21}\hat{v}_{21} + P_1\hat{v}_1) \\
&= (1 - F(\bar{v}_2))\, D_2(\hat{v}_2),
\end{aligned}$$

where the inequality follows by applying Jensen's inequality, and the last equality holds by the law of total expectation. The last expression is the right-hand side of inequality (14), thus proving the result.

Proof of Theorem 1

For the first statement, note that

$$\frac{\partial \pi_{t,1s}}{\partial Q_1} = \frac{\partial \pi_1}{\partial Q_1} + \frac{\partial \pi_{2C,1s}^*}{\partial Q_1} + \frac{\partial \pi_{2I,1s}^*}{\partial Q_1}, \quad \text{and} \quad \frac{\partial \pi_{t,2s}}{\partial Q_1} = \frac{\partial \pi_1}{\partial Q_1} + \frac{\partial \pi_{2C,2s}^*}{\partial Q_1} + \frac{\partial \pi_{2I,2s}^*}{\partial Q_1}.$$

The first-period is identical under the two learning regimes and the effect of the inventory channel is unchanged, so that $\frac{\partial \pi_{1C,1s}^*}{\partial Q_1} = \frac{\partial \pi_{2C,2s}^*}{\partial Q_1}$. However, from Propositions 3 and 5, we have for all $Q_1 \in [0, Q_{\max}]$ that $\frac{\partial \pi_{2I,1s}}{\partial Q_1} > \frac{\partial \pi_{2I,2s}}{\partial Q_1}$. It then follows by the unimodality of $\pi_{t,1s}(Q_1)$ that $Q_{1,2s}^* < Q_{1,1s}^*$.

For the second statement, note that the firm's total profit is given by $\pi_{t,2s}(Q_1) = \pi_1(Q_1) + \pi_{2C}^*(Q_1) + \pi_{2I}^*(Q_1)$, where

$$\begin{aligned}
\pi_1(Q_1) &= pE\left[\min\left\{\gamma\left[1 - \frac{p}{v}\right]^+, Q_1\right\}\right] - cQ_1 \\
\pi_{2C}^*(Q_1) &= \int_{v_l}^{v_c(Q_1)} c\min\left\{(1-\gamma)\left[1 - \frac{p}{v}\right]^+, Q_L(Q_1, v)\right\} f(v)dv \\
\pi_{2I}^*(Q_1) &= \int_{v_l}^{v_c(Q_1)} (p-c)(1-\gamma)\left[1 - \frac{p}{v}\right]^+ f(v)dv + \int_{v_c(Q_1)}^{v_h} (p-c)(1-\gamma)\left[1 - \frac{p}{\hat{v}(Q_1)}\right]^+ f(v)dv.
\end{aligned}$$

Note that there is a one-to-one correspondence between v_c and Q_1. Treating v_c as the decision variable, we have $\pi_{t,2s}(v_c) = \pi_1(v_c) + \pi_{2C}^*(v_c) + \pi_{2I}^*(v_c)$,

$$\pi_1(v_c) = pE\left[\min\left\{\gamma\left[1 - \frac{p}{v}\right]^+, Q_1(v_c)\right\}\right] - cQ_1(v_c)$$

$$\pi^*_{2C}(v_c) = \int_{v_l}^{v_c} c \min\left\{(1-\gamma)\left[1-\frac{p}{v}\right]^+, Q_L(Q_1(v_c), v)\right\} f(v)dv$$
$$\pi^*_{2I}(v_c) = \int_{v_l}^{v_c} (p-c)(1-\gamma)\left[1-\frac{p}{v}\right]^+ f(v)dv + \int_{v_c}^{v_h} (p-c)(1-\gamma)\left[1-\frac{p}{\hat{v}(Q_1(v_c))}\right]^+ f(v)dv.$$

We next consider the derivative of $\pi_{t,2s}(v_c)$ in the limit $\gamma \to 0$. The derivatives of π_1 and π^*_{2C} are given by (see also the proofs of Propositions 1 and 2)

$$\frac{\partial \pi_1}{\partial v_c} = \gamma\left(\frac{p}{v_c^2}\right)\left(p\bar{F}(v_c) - c\right)$$
$$\frac{\partial \pi^*_{2C}}{\partial v_c} = \gamma\left(\frac{p}{v_c^2}\right) c \int_{v_m}^{v_c} f(v)dv,$$

so that $\lim_{\gamma\to 0}\frac{\partial \pi_1}{\partial v_c} = \lim_{\gamma\to 0}\frac{\partial \pi^*_{2C}}{\partial v_c} = 0$ for any $v_c > 0$. Next, from Proposition 5 we have $\lim_{\gamma\to 0}\frac{\partial \pi^*_{2I}}{\partial Q_1} < 0$, which implies $\lim_{\gamma\to 0}\frac{\partial \pi^*_{2I}}{\partial v_c} < 0$ so that $\lim_{\gamma\to 0} v^*_c = \underline{v}_l$. It follows by continuity of v^*_c in γ that for sufficiently small γ, we have $v^*_c < v_m$ (where v_m is given in Proposition 1), which is equivalent to $Q^*_{1,2s} < Q^*_{1,m}$.

Proof of Proposition 6

Under one-sided learning, let the optimal first-period quantity be Q^*_1. Now suppose that under two-sided information the firm employs the same policy. Then first-period profit under the two learning regimes is identical; we will show that expected second-period profit is higher in the case of two-sided learning. Note first that the second-period profit is the same under the two regimes for any $v < v_c(Q^*_1)$. Now consider cases $v \geq v_c(Q^*_1)$. From Proposition 3 we know that if the consumers are informed (as is the case under one-sided learning), the firm also prefers to be informed. That is, if π^*_c denotes expected profit across scenarios $v \geq v_c(Q^*_1)$ when the consumers are informed but the firm is uninformed, and π^*_b denotes the (hypothetical) profit when both are informed, then $\pi^*_b > \pi^*_c$. Next, from Proposition 5 we have that if being informed means that consumers are also informed, the firm prefers both parties to be uninformed. Letting π^*_u denote expected profit across scenarios $v \geq v_c(Q^*_1)$ when both the firm and the consumers are uninformed, this implies that $\pi^*_u > \pi^*_b$. Combining the two inequalities, we have $\pi^*_u > \pi^*_b > \pi^*_c$ completing the proof.

Proof of Proposition 7

We treat the first-period stockout threshold v_c as the firm's decision variable (note that for any given γ there is a one-to-one correspondence between v_c and Q_1; see (3)). Let $v^*_{c,m}$, $v^*_{c,1s}$, and $v^*_{c,2s}$ denote the optimal myopic, one-sided and two-sided thresholds, respectively. Now, note that $v^*_{c,m}$ is independent of γ, while from Propositions 3 and 5 we have $\lim_{\gamma\to 0} v^*_{c,1s} = v_h$ and $\lim_{\gamma\to 0} v^*_{c,2s} = \underline{v}_l$ so that $v^*_{c,2s} < v^*_{c,m} < v^*_{c,1s}$ (this is because $\lim_{\gamma\to 0}\pi_{t,2s}(Q_1) = \pi^*_{2I}(Q_1)$ and $\pi^*_{2I}(Q_1)$ is monotonic increasing in Q_1 under one-sided learning and monotonic decreasing under two-sided learning). This implies that $Q^*_{1,2s} < Q^*_{1,m} < Q^*_{1,1s}$. It then follows by monotonicity of $\pi^*_{2I}(Q_1)$ in Q_1 (see Proposition 5) that in the limit $\gamma \to 0$ we have $\pi_{t,2s}(Q^*_{1,1s}) < \pi_{t,2s}(Q^*_{1,m}) < \pi_{t,2s}(Q^*_{1,2s})$.

Proof of Proposition 8

Since the demand for the product in the first period is deterministic and given by $D_1(b_1)$, under any quantity decision $Q_1 \leq Q_{max} = D_1(b_1)$ the amount sales achieved by the firm, and therefore the volume of product reviews generated in the first period, is $n = Q_1$ (it is straightforward to show that any quantity decision

$Q_1 > Q_{max}$ is strictly suboptimal and therefore need not be considered). Moreover, under any quantity decision $Q_1 \le Q_{max}$, the consumers' updated value belief b_2 is viewed in the first period as a random variable (as it depends on the ex ante uncertain content of product reviews); let $\psi(\cdot;Q_1)$ denote the corresponding pdf.

Next, note that the firm's second-period profit can be expressed as $\pi_2^*(Q_1) = \pi_{2C}^*(Q_1) + \pi_{2I}^*(Q_1)$ in accordance with the analysis of §6. (Notice that in this model we have $\pi_{2C}^*(Q_1) = 0$ for any $Q_1 \le Q_{max}$, since under any such quantity decision there are no leftover units in the first period and all units sold in the second period must therefore be newly produced; we thus have $\pi_2^*(Q_1) = \pi_{2I}^*(Q_1)$.) To calculate $\pi_{2I}^*(Q_1)$, we point out that conditional on the content of product reviews and the resulting updated belief b_2, demand in the second period is deterministic and given by $D_2(b_2) = (1-\gamma)\left[1 - \frac{p}{b_2 v_h + (1-b_2)v_l}\right]$, so that the optimal second-period quantity is simply $Q_2^*(b_2) = D_2(b_2)$, and the corresponding optimal profit is $\pi_2^*(b_2) = pD_2(b_2)$. To find the expected second-period profit as a function of Q_1, we take expectation with respect to the updated belief b_2, noting that its pdf depends on Q_1; in particular, we have

$$\pi_{2I}^*(Q_1) = \int_0^1 \pi_2^*(b_2)\psi(b_2;Q_1)db_2 = \int_0^1 pD_2(b_2)\psi(b_2;Q_1)db_2$$

To establish that the above expression is strictly decreasing in Q_1 for all $Q_1 \in [0, Q_{max}]$, we note that (a) $D_2(b_2)$ is positive, increasing and strictly concave in b_2, and (b) by Proposition 1(ii) of Yu et al. (2015), the distribution of the updated belief b_2 is strictly second-order stochastically decreasing in the first-period quantity Q_1. Observations (a) and (b) together imply that $\pi_{2I}^*(Q_1)$ is strictly decreasing in Q_1.

Proof of Proposition 9

By the strict monotone likelihood ratio property of the agents' signals, we have that the updated value expectation $\mu(s_i)$ of agent i is strictly increasing in her private signal s_i. The agent's willingness to pay for the product is then given by $z_i = x_i\mu(s_i)$. Let $W(z;v)$ denote the cdf of z_i; note that $W(z;v)$ is strictly decreasing in v. Demand in the first period is then given by $D_1(v) = \gamma(1 - W(p;v))$.

Since $D_1(v)$ is strictly increasing in v, if the firm stocks Q_1 units in the first period and a stockout does not occur, product value can be perfectly deduced through $D_1(v) = S_1 = Q_1$. On the other hand, if a stockout does not occur, then product value can be deduced only up to $D_1(v) \ge S_1 = Q_1$, or equivalently, only up to $v \ge v_{c2}(Q_1)$ for an appropriate function $v_{2c}(Q_1)$. Moreover, note that strict monotonicity of $D_1(v)$ in v implies that $v_{c2}(Q_1)$ is strictly increasing in Q_1. To prove the result, we may then replace $v_c(Q_1)$ with $v_{2c}(Q_1)$ in the proof of Proposition 5.

Proof of Proposition 10

The firm's total expected profit as a function of Q_1 is $\pi_{t,2s}(Q_1) = \pi_1(Q_1) + \pi_2^*(Q_1)$, where $\pi_2^*(Q_1)$ denotes expected profit under the optimal second-period quantity. As in §6, we write $\pi_2^*(Q_1) = \pi_{2C}^*(Q_1) + \pi_{2I}^*(Q_1)$, and focus on $\pi_{2I}^*(Q_1)$ in order to establish that $\frac{d\pi_{2I}^*}{dQ_1} < 0$. Recall that $\pi_{2I}^*(Q_1)$ is the firm's optimal expected profit assuming all units sold in the second period are newly produced. When the followers receive first-period information (and therefore learning occurs) with probability ζ, we have $\pi_{2I}^*(Q_1) = \zeta\pi_{2I,L}^*(Q_1) + (1-\zeta)\pi_{2I,N}^*(Q_1)$, where $\pi_{2I,L}^*(Q_1)$ ($\pi_{2I,N}^*(Q_1)$) is the firm's profit when learning occurs (when learning does not occur). Now, observe that when learning does not occur, demand in the second period is independent of Q_1, so that $\pi_{2I,N}^*(Q_1)$ is also independent of Q_1 and $\frac{d\pi_{2I,N}^*}{dQ_1} = 0$. Next, note that when learning does occur, we have $\frac{d\pi_{2I,L}^*}{dQ_1} < 0$, as this is the case covered by Proposition 5. Therefore, we have $\frac{d\pi_{2I}^*}{dQ_1} = \zeta\frac{d\pi_{2I,L}^*}{dQ_1} < 0$.

Proof of Corollary 1

Fix any price p. By Proposition 1, the quantity decision $Q_m = \gamma\left(1 - \frac{p}{v_r}\right)$, where $v_r = F^{-1}\left[\frac{p-c}{p}\right]$ maximizes the firm's expected first-period profit. By Theorem 1(ii), for γ sufficiently small the optimal quantity is strictly smaller than Q_m. Since Theorem 1 holds for any arbitrary price, it follows that the optimal policy $\{p^*_{2s}, Q^*_{1,2s}\}$ satisfies $Q^*_{1,2s} < \gamma\left(1 - \frac{p^*_{2s}}{v_r}\right)$.

References

Alizamir, S., F. de Vericourt, S. Wang. 2018. Warning against recurring risks: An information design approach *Working paper, Yale University.*

Apple Inc. 2017. Apple reports record first quarter results. (Jan. 31st).

Arrow, K., T. Harris, J. Marschak. 1951. Optimal inventory policy. *Econometrica* **19**(3) 250–272.

Azoury, K. 1985. Bayes solution to dynamic inventory models under unknown demand distribution. *Management science* **31**(9) 1150–1160.

Bensoussan, A., S. Sethi. 2009. A note on "the censored newsvendor and the optimal acquisition of information". *Operations Research* **57**(3) 791–794.

Bergemann, D., J. Välimäki. 1997. Market diffusion with two-sided learning. *RAND Journal of Economics* **28**(4) 773–795.

Besbes, O., A. Muharremoglu. 2013. On implications of demand censoring in the newsvendor problem. *Management Science* **59**(6) 1407–1424.

Bimpikis, K., M. Markakis. 2015. Inventory pooling under heavy-tailed demand. *Management Science* **62**(6) 1800–1813.

Bloomberg. 2018. Here's what analysts say about the new samsung smartphones. (Feb. 27th).

Cachon, G., R. Swinney. 2009. Purchasing, pricing, and quick response in the presence of strategic consumers. *Management Science* **55**(3) 497–511.

Chen, L., A. Mersereau. 2015. Analytics for operational visibility in the retail store: The cases of censored demand and inventory record inaccuracy. *Retail Supply Chain Management*. Springer, 79–112.

Cnet.com. 2018. Apple iPhone XS, XS Max preorders kick off, some models sell out. (Sep. 14th).

Courty, P., J. Nasiry. 2016. Product launches and buying frenzies: A dynamic perspective. *Production and Operations Management* **25**(1) 143–152.

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

Debo, L., G. Van Ryzin. 2009. Creating sales with stockouts *Working paper, University of Chicago.*

DeGraba, P. 1995. Buying frenzies and seller-induced excess demand. *The RAND Journal of Economics* **26**(2) 331–342.

Ding, X., M. Puterman, A. Bisi. 2002. The censored newsvendor and the optimal acquisition of information. *Operations Research* **50**(3) 517–527.

Feldman, P., Y. Papanastasiou, E. Segev. 2018. Social learning and the design of new experience goods. *Management Science* .

Hu, M., J. Milner, J. Wu. 2015. Liking and following and the newsvendor: Operations and marketing policies under social influence. *Management Science* **62**(3) 867–879.

Lariviere, M., E. Porteus. 1999. Stalking information: Bayesian inventory management with unobserved lost sales. *Management Science* **45**(3) 346–363.

Liu, T., P. Schiraldi. 2012. New product launch: herd seeking or herd preventing? *Economic theory* **51**(3) 627–648.

Lovejoy, W. 1990. Myopic policies for some inventory models with uncertain demand distributions. *Management Science* **36**(6) 724–738.

Lu, X., J. Song, K. Zhu. 2005. On "the censored newsvendor and the optimal acquisition of information". *Operations Research* **53**(6) 1024–1026.

Mersereau, A. 2015. Demand estimation from censored observations with inventory record inaccuracy. *Manufacturing & Service Operations Management* **17**(3) 335–349.

Milgrom, Paul. 1981. Good news and bad news: Representation theorems and applications. *Bell Journal of Economics* **12**(2) 380–391.

Papanastasiou, Y., N. Bakshi, N. Savva. 2016. Scarcity strategies under quasi-bayesian social learning. *Working Paper, London Business School.*

Papanastasiou, Y., K. Bimpikis, N. Savva. 2017. Crowdsourcing exploration *Management Science, forthcoming.*

Papanastasiou, Y., N. Savva. 2016. Dynamic pricing in the presence of social learning and strategic consumers. *Management Science* **63**(4) 919–939.

Peter, J Paul, Jerry Corrie Olson, Klaus G Grunert. 1999. Consumer behavior and marketing strategy .

Petruzzi, N., M. Dada. 2002. Dynamic pricing and inventory control with learning. *Naval Research Logistics* **49**(3) 303–325.

Rogers, Everett M. 2010. *Diffusion of innovations*. Simon and Schuster.

Scarf, H. 1959. Bayes solutions of the statistical inventory problem. *The annals of mathematical statistics* **30**(2) 490–508.

Stock, A., S. Balachander. 2005. The making of a "hot product": A signaling explanation of marketers' scarcity strategy. *Management Science* **51**(8) 1181–1192.

Su, X., F. Zhang. 2009. On the value of commitment and availability guarantees when selling to strategic consumers. *Management Science* **55**(5) 713–726.

Swinney, R. 2011. Selling to strategic consumers when product value is uncertain: The value of matching supply and demand. *Management Science* **57**(10) 1737–1751.

tech.co. 2015. Xiaomi vs apple: The influence of hunger marketing. (Apr. 23rd).

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