Motivated by apparel retailers and grocers’ markdown practice over repeated sales horizons, we study a newsvendor who sells a perishable asset over repeated sales seasons to loss averse consumers. Consumers have a given intrinsic valuation for the product. The market size in each sales season is random following a stationary distribution. Consumers are loss averse with stochastic reference points that represent their beliefs about the price and product availability. Given these reference points, they choose purchase plans to maximize their expected utility before visiting the store. In anticipation of consumers’ purchase plans, in each sales season, before demand uncertainty resolves, the firm chooses the initial order quantity; after the uncertainty resolves, the firm chooses a contingent price depending on the demand realization, with the option to clear inventory by charging a sale price. Over repeated seasons, the firm’s operational decisions of ordering and contingent pricing result in a retail price distribution, and in equilibrium, this distribution is consistent with consumers’ stochastic reference points. Focusing on a class of pricing strategies that run contingent sales, we fully characterize the firm’s optimal inventory and contingent pricing policies. Surprisingly, we identify conditions under which the expected price and the firm’s profit can be increasing in the consumer loss aversion level and demonstrate that the firm can prefer low and moderate levels of demand variability over no demand uncertainty. Moreover, we obtain a set of counterintuitive insights on how consumers’ loss aversion affects the firm’s optimal operational policies, which are in stark contrast to those in classic Newsvendor models.

1. Introduction

1.1. Motivation

Firms in many industries, e.g., apparel and grocery, face the problem of selling a fixed initial inventory of perishable assets over a finite sales horizon. For these firms, pricing is the only control to match supply with demand once they set up their initial inventory levels in the beginning of the
sales horizon: if demand is slack during the season, firms run sales to boost demand.\footnote{According to an estimate by the management consulting firm A.T. Kearney, a typical apparel retailer sells between 40\% and 45\% of its inventory at a promotional price (D’Innocenzio 2012). In the microdata collected monthly by the US Bureau of Labor Statistics on goods and services including three major groups (processed food, unprocessed food and apparel), the average sale price is about 25\% to 30\% off the regular price (Klenow and Kryvtsov 2008).} To automate pricing decision making, many retailers have begun using markdown pricing optimization software to determine the depth and frequency of their clearance events.\footnote{Software vendors that offer markdown optimization solutions include DemandTec (now part of IBM), Oracle, Predictix, Revionics and SAS.} Moreover, in many cases, these firms face the same decisions over and over again for repeated sales horizons/seasons. For grocery stores selling fruits and vegetables, each sales horizon can be as short as several days.

Over repeated horizons, markdown pricing practice plays a crucial role in the formation of consumers’ purchase behavior. Media reports on grocers suggest that customers who have been intending to buy at a discounted item very often end up with also purchasing the product when they find the sale has ended and it is listed at the full price (see, e.g., Jargon 2013). However, if the sales are too frequent, consumers can get trained not to buy at the full price. This delicacy alerts that markdown pricing practice needs to strike a fine balance in deciding when and to what extent to run sales, taking into account how consumers react to sales over time.

Experimental studies indicate that consumers often evaluate economic outcomes, e.g., price to pay, relative to a distribution of reference levels, and not merely in terms of absolute levels (see, e.g., Blinder 1998). Such an effect is particularly significant over repeated purchase interactions, as consumers tend to draw on past experiences as benchmarks: they form ideas about what the typical prices are, and judge the value of a product based on the difference between these typical prices and the posted price they observe. Consumers evaluate changes from possible reference levels differently depending on whether the changes are gains or losses. In particular, there is significant empirical evidence indicating that consumers are loss averse, i.e., they weigh losses more heavily than equally-sized gains (see, e.g., Kahneman et al. 1990, Hardie et al. 1993).

Moreover, intuitively, the reference points should be endogenized within consumers, rather than to be given exogenously. This view has gained attraction in the behavioral literature and was formalized by Koszegi and Rabin (2006) with many follow-up works. In particular, the behavioral economics literature (e.g. Heidhues and Koszegi 2005, Koszegi and Rabin 2006) has identified two countervailing effects on firms of running sales when consumers are loss averse with endogenized stochastic reference points: the negative \textit{comparison effect} and the positive \textit{attachment effect}. The comparison effect means that higher sales frequency increases the weight of sale prices in the loss-averse consumers’ reference distribution, making consumers get used to the sale price and less likely to purchase at the full price. The attachment effect, on the contrary, means that higher sales frequency increases consumers’ attachment to forming a habit of purchasing. Hence, to avoid the pain of not obtaining the product when there are no sales, consumers are willing to pay a full price. This paper is motivated by this literature on endogenized reference points and by the markdown practice over repeated sales horizons in the apparel and grocery industries.
1.2. Research Questions and the Model

We are interested in the following research questions in a repeated newsvendor setting with loss averse consumers. First, how should the newsvendor optimally make its ordering and pricing decisions over repeated sales horizons, taking into account purchase behavior of loss averse consumers with stochastic reference points? The answer to this question may provide another justification to the prevalence of the contingent markdown pricing practice. Second, is it possible for the newsvendor to benefit from consumers’ loss averse behavior and if so, when such benefit may prevail? The answer to these questions will provide insights on how firms may improve profitability in the presence of loss averse consumers. Third, like in the classic newsvendor problem, there exists probability of running out of stock. The stockout events can critically factor into loss-averse consumers' purchase decision making over repeated interactions. What is the effect of product unavailability on the firm’s long-term profitability when consumers are loss averse and how should the firm optimally respond? The answer to this question will deepen our understanding of the effect of stocking decisions on firms’ profitability in practical settings where consumers are loss averse.

To answer these questions, we consider a profit-maximizing newsvendor who sells a perishable product on a repeated basis and seeks to maximize its expected long-run average profit. Consumers have a given intrinsic valuation and their market size is random for each sales season following a stationary distribution. Moreover, consumers are loss averse with reference levels that represent their probabilistic beliefs about the product’s price and availability outcomes learned over time. Given these reference points, they develop purchase plans to maximize their total expected utility including gain-loss utility due to their loss-averse behavior. In anticipation of consumers’ purchase plans, in each sales season, before demand uncertainty resolves, the firm chooses an initial order quantity; after the uncertainty resolves, the firm sets full or sale prices contingent on demand realization. While the firm aims to sell the product at the full price, due to the perishability of the product, the firm may choose to run sales to avoid unsold, but already sunk, inventory, or to manipulate customers’ purchase plans. Over repeated seasons, the firm’s operational decisions of ordering and contingent pricing result in a distribution of full and sale prices. In equilibrium, this distribution should be consistent with consumers’ stochastic reference points. In this sense, the consumers’ reference points in our model are endogenously induced and determined by their own purchase plans rather than to be provided as exogenous values.

1.3. Contributions and Main Results

This paper makes three main contributions as follows: First, we are the first to explicitly take into account consumers’ loss aversion with stochastic reference points in firm’s operational decisions, such as the order quantity and contingent pricing policy. An important aspect of this study is the effect of stockouts on the willingness-to-pay of loss-averse consumers, which leads to several counterintuitive insights. Second, we demonstrate that contingent pricing policies allow the firm not only to efficiently match supply with demand, but also to profitably manipulate consumers’
stochastic reference points by varying the frequency of sales. In line with this finding, we numerically show that the firm may prefer small or moderate levels of demand variability, to no demand variability, under the optimal contingent pricing strategy. This is an observation, which is in stark contrast to the results in a classic newsvendor setting. The reason behind this observation is that sales driven by an appropriate level of demand uncertainty, can entice loss-averse consumers to form a purchase plan of buying up to a higher full price. Third, we show how consumers’ loss aversion affects the firm’s optimal operations decisions, with insights significantly different from those in classic newsvendor settings (see, e.g., Petruzzi and Dada 1999). Based on our results, we caution that it is critical to take into account consumers’ loss averse behavior when designing markdown pricing algorithms; otherwise, existing markdown algorithms that ignore consumer loss aversion can lead to substantially sub-optimal solutions.

To be specific, we characterize the optimal ordering and contingent markdown pricing policies, and study their comparative statics. We show that the firm’s optimal two-price contingent pricing policy is to sell the product at a sale price if and only if the realization of the regular consumers’ demand is lower than a threshold, and to sell at a full price otherwise. Under the contingent pricing strategy, demand variability leads to certain frequency of sales. However, this can benefit the firm as it makes the positive attachment effect the dominating force, which increases consumers’ willingness to pay a high price. Hence, the firm may surprisingly be able to charge an expected price (weighted over the optimal full price and sale price by the optimal sales frequency) that is higher than consumers’ intrinsic valuations; moreover, as mentioned, the firm may also prefer low or moderate levels of demand variability.

Furthermore, we show that the dominating attachment effect of sales critically hinges on the firm’s operational decisions. This leads to a set of counterintuitive insights that hold as long as the optimal sales frequency is nonzero. First, the optimal full price in the contingent pricing policy increases in the initial order quantity. This result is in contrast to the common sense from the supply-demand relationship in a non-loss-averse consumer market. The driving force behind this anomaly is the attachment effect. Specifically, an increase in the order quantity boosts consumers’ expectation of buying the product due to a higher fill rate that, in turn, increases their feeling of loss if they do not buy and therefore increases their willingness-to-pay. Second, we demonstrate that the optimal full price decreases in the procurement cost and so does the expected price. This result goes against the common intuition of a positive cost pass-through rate in a non-loss-averse consumer market, where a higher supply cost typically leads to a higher retail price. The driving force behind this result is that an increase in the procurement cost causes the firm to reduce the initial order quantity, which weakens the attachment effect due to a lower fill rate and as a result, the optimal full price has to decrease. Third, the optimal sales frequency increases in the procurement cost. One could expect that when the procurement cost increases, the firm would be more conservative in running sales. The intuition behind this counterintuitive result is again
due to a reduced optimal initial inventory level resulted from an increase in the procurement cost. To induce consumers to form a purchase plan, the firm has to run more frequent promotions to strengthen the attachment effect.

2. Related Literature

There are two bodies of literature that are closely related to our research. The first is the behavioral economics literature on consumers’ loss aversion. The theory of loss aversion was first proposed by Kahneman and Tversky (1979) and extended later by Kahneman et al. (1990) to choices under uncertainty. Examples of recent works include Heidhues and Koszegi (2005, 2008, 2013) and Koszegi and Rabin (2006) that consider stochastic reference points and the impact of consumers’ loss aversion on firms’ pricing strategies. Our research is distinct from these works in several aspects. These papers assume that market demand is deterministic while we consider demand uncertainty. These works ignore inventory decisions and therefore assume that the product is always available. In contrast, we consider the product’s unavailability, which is a key feature in our model leading to several unique insights. Furthermore, in our model, stochastic reference points are driven by contingent pricing policy that is also used to combat demand uncertainty. In summary, the works in behavioral economics typically ignore the operational decisions of inventory and contingent pricing. We consider the joint effect of contingent pricing in operationally matching supply with demand and in profitably manipulating consumers’ stochastic reference points.

The second closely related stream is on consumer reference-price effects in the pricing and revenue management literature (see Arslan and Kachani 2011 and Özer and Zheng 2012, Section 3.1.2 for recent surveys). Fibich et al. (2003) consider a continuous-time infinite-horizon monopolistic pricing problem with a linear demand model and linear reference effects. The authors show that the optimal pricing policy over time can be either a skimming or penetration strategy and that using the optimal fixed price is close to optimal. Popescu and Wu (2007) study a discrete-time version of Fibich et al. (2003) under a general nonlinear reference-dependent demand model and further confirm the (either increasing or decreasing) monotonicity of the optimal price trajectory. Nasiry and Popescu (2011) study a version of the problem with the reference point as a weighted average of the lowest and most recent prices. Zhao and Stecke (2010) study a newsvendor who can advance-sell to loss averse consumers and the authors solve for the firm’s optimal advance-selling strategy. All of these papers consider a single reference point for decision makers at the time of decision making. While the assumption of a single reference point is a reasonably good approximation to reality, as information becomes more abundant and accessible (due to, e.g., websites such as decide.com provide historic price information), consumers’ reference dependence in their decision making becomes more complicated. In contrast to these papers, we thus consider probabilistic beliefs as consumers’ reference points in determining their purchase decisions. Consistent with Roels and Su (2013) that assume the reference points can be engineered, we echo that the distribution of
reference prices can be manipulated by operational policies. In our context, the reference points for repeated purchasers are driven by demand uncertainty and firm’s inventory and contingent pricing policies.

A related consumer behavior to reference-price effects is regret and disappointment, incurred due to a mismatch between a reference point and a realization. Nasiry and Popescu (2012) characterize the effect of anticipated regret on consumer decisions and on firm profits and policies in an advance selling context where buyers have uncertain valuations. The purchase decisions may have emotional impacts when consumers learn their values and if they have made a wrong choice. Liu and Shum (2013) study a firm’s optimal pricing and rationing decisions over two periods in anticipation of possible consumers’ disappointment due to stockouts, and show that the firm may benefit from such disappointment. Özer and Zheng (2013) study a seller’s optimal pricing and inventory strategies when anticipated regret and misperception of product availability affect consumers’ purchase decisions. While these papers assume a single reference point that leads to regret and disappointment, we focus on stochastic reference points.

Ho et al. (2010) consider manager’s stochastic reference dependence and study the ordering behavior in a multi-location inventory system. The authors consider the reference dependence in the newsvendor’s decision making by assuming that there are psychological costs for the decision maker in facing leftovers and stockouts, whereas leftovers are psychologically more costly than stockouts. In a competitive newsvendor setting, Avcı et al. (2013) study the managers’ loss aversion and status seeking behavior in making newsvendor ordering decisions, with stochastic reference points on the possible competitor’s profit outcomes. These two papers assume that the firms’ managers, as decision makers, are loss averse with stochastic reference points, while we focus on loss averse behavior of consumers and more distinctively, how the firm should react to it.

Several papers consider strategically forward-looking, but loss neutral and nonemotional, consumers’ behavior and its impact on firms’ optimal inventory and pricing decisions. Liu and van Ryzin (2008) and Gallego et al. (2008) consider a capacity rationing problem in a market where the firm can credibly commit to prices. Liu and van Ryzin (2008) find that the firm can optimally set the rationing level in the markdown period to induce high-value consumers to buy in the early period, so to better differentiate the market. Gallego et al. (2008) study the inventory level that should be assigned to sales when strategic consumers adjust their expectations of sales from the firm’s past actions over repeated seasons. These papers focus on capacity rationing decisions, while we consider optimal inventory and contingent pricing decisions. More importantly, while the profit gain for the firm in those contexts comes from better market segmentation, the profit gain for the firm in our model comes from inducing consumers to buy at a full price by manipulating their reference points. For a fixed capacity, Cachon and Feldman (2013) study discount strategies of how often and to what extent to run sales, in a market where consumers face search costs. Different
from the “more frequently than optimal”\textsuperscript{3} discount strategy in their setting that entices consumers to visit the store, the discount strategy in our model serves to induce loss averse consumers to buy at higher prices, and most likely results in “less frequently than optimal” discount strategy (see Corollary 1).

We end this section by a brief comparison between our loss aversion model and those commonly applied in the operations management literature. There are two main differences. First, consistent with the recent works in behavioral economics, e.g., Koszegi and Rabin (2006), we assume that loss averse consumers assess gains and losses in two dimensions, the product and money, separately. While works in behavioral operations management usually consider one dimension of combined consumption utilities (see, e.g., Ho et al. 2010 for managers’ decisions), Koszegi and Rabin (2006) argue that a belief in the two separate dimensions – product and money – may be the most appropriate candidate to be considered for consumers’ reference levels. Second, we consider that by repeatedly visiting the store, consumers take into account the product availability and price outcomes, and their frequencies. Hence, the consumers’ reference level is not necessarily a single point. Instead, it is more likely a distribution of multiple points with respective probabilities. Experimental studies show that in most cases when consumers’ expectations are different from the status quo, probabilistic expectations or beliefs generally result in better predictions of consumers’ behavior (see, e.g., Mellers et al. 1999, Breiter et al. 2001 and Grossman 2011).

3. The Model

We consider a single risk-neutral profit maximizing firm selling a single perishable product over a short horizon on a repeated basis, e.g., a grocery store that sells fruit plates faces a sales horizon measured in days. The firm orders $q$ units of the product at cost $c$ per unit and sells to consumers who request a single unit of the product.

\textbf{Consumer.} There is a random number, $D$, of consumers, where $D \geq 0$ has cumulative distribution function (cdf) $F(\cdot)$ and has an expected value $\mathbb{E}(D) < \infty$. The consumers have a known intrinsic valuation $v(> c)$ of the product. Consumers are loss-averse. Our loss aversion model, to be discussed in further detail, extends the one in Koszegi and Rabin (2006) to consider the impact of operational issues such as stockouts on consumers’ purchase decisions.

\textbf{Sale Price.} We assume that there is a sale price $s(< c)$ at which the firm can clear all on-hand inventory by selling to bargain hunters. We fix the sale price as exogenously given and allow the full price to be optimized. Same qualitative insights can be obtained if one fixes the full price but endogenizes the sale price. The assumption $s < c$ is consistent with the empirical evidence based on data from the US Bureau of Labor Statistics: the sale price typically represents a sizable markdown from the marginal cost (see Shi 2012). This assumption is also commonly used in the

\textsuperscript{3}The “optimal” here refers to the optimal contingent price strategy conditional on the realized demand of a sales season that maximizes the profitability for that particular season.
operations literature, see, e.g., Cachon and Swinney (2009). Consistent with the US Bureau of Labor Statistics which defines a sale as a price cut available to all buyers, we assume that the firm cannot price discriminate among consumers and bargain hunters when running sales. However, as the valuation of bargain hunters is below the procurement cost, the firm primarily targets consumers for profitability. When the sale price is charged, we assume that the firm can prioritize selling to consumers. The one-price-for-all assumption is indeed stylized but it simplifies analysis; moreover, it provides a lower bound on the firm’s profitability when price discrimination is allowed.

**Inventory Availability.** An important novel feature of our model is the consideration of product availability in the framework of stochastic reference points. Since the firm orders a limited amount of order quantity, a consumer may find the product unavailable when demand exceeds supply. This unavailability was ignored in the loss aversion models considered in the behavioral economics literature (e.g., Koszegi and Rabin 2006). But, it is a realistic feature in many applications. In addition, as shown below, the consideration of product unavailability leads to insights that are significantly different from those obtained in previous works. We assume that if there are more consumers than the available products, then the products are rationed among consumers with an equal probability (see, e.g., Su and Zhang 2008 for the same rationing rule). We define the fill rate as the long-run probability that the product is successfully procured when consumers are willing to buy it. The fill rate, together with the price distribution, influences loss averse consumers’ purchase decisions.

**Contingent Pricing Scheme.** Heidhues and Koszegi (2013) show that for a fixed number of loss averse consumers, the optimal strategy is to induce consumers to always purchase the product, and the optimal pricing scheme is to randomize the price in an interval of sale prices with a continuous probability distribution and an isolated atomic full price with a positive mass. We conjecture that the same optimal pricing structure might also propagate in our setting. Nevertheless, even under the much simpler assumptions (i.e., single customer type and no demand uncertainty) as in Heidhues and Koszegi (2013), the closed form of the optimal pricing scheme is not attainable and the implementation of this scheme may be impractical. As the main objective of this paper is to explore the structures of inventory/contingent pricing policies and their profitability under demand uncertainty, we restrict our attention to pricing schemes with easily implementable forms. Specifically, we focus on two-price contingent pricing schemes, in which only two prices – full and sale price – are charged conditional on the realization of consumer demand and that the sale price is exogenously given as $s$. This pricing strategy implies that the firm occasionally runs sales to clear the market. Although it may seem restrictive to consider a two-price pricing scheme, randomizing among a limited number of discrete prices is more practical than randomizing in a continuous price interval as suggested by Heidhues and Koszegi (2013). Moreover, the two-price contingent scheme provides a lower bound on the benefit of contingent pricing over static pricing with a single price.
Let $\bar{p}(>s)$ denote the full price. In addition, let $\Omega$ be the set of demand realizations for which the price is set to be $s$. Then the contingent pricing scheme is

$$p(x) = \begin{cases} 
    s & \text{if } x \in \Omega, \\
    \bar{p} & \text{if } x \in \Omega^c = \mathbb{R}^+ \setminus \Omega.
\end{cases} \quad (1)$$

(We will characterize the optimal set $\Omega$ in Section 4.)

**Figure 1 Sequence of Events**

Sequence of Events. We describe the timing of the model, which is also illustrated in Figure 1. Before the selling seasons: (i) The firm decides and commits to the stock quantity $q^4$, fill rate $\phi$ and the contingent pricing policy $p(x)$ as in (1). The firm announces the fill rate $\phi$, and the price distribution

$$g(p) = \begin{cases} 
    \int_{x \in \Omega} dF(x) & \text{if } p = s, \\
    \int_{x \in \Omega^c} dF(x) & \text{if } p = \bar{p},
\end{cases} \quad (2)$$

induced by the contingent pricing policy. (ii) Consumers commit to a purchase plan. During the selling seasons, the following events occur: (iii) At the beginning of each season, the firm purchases $q$ units of inventory. (iv) The random demand from consumers realizes as $x$. (v) The firm sets the price $p = p(x)$ according to its contingent pricing policy. (vi) Consumers observe the price $p$ and the availability of the product, and make their purchase decisions according to their mentally committed plans. If the product is available and is charged with a price at which the consumers have planned to buy, they will purchase it. If there are more consumers willing to purchase the product than the available inventory, units are randomly rationed.

4 The stocking decision can be delayed to the beginning of each selling season before demand uncertainty is resolved, which does not change the results of the model.
Two comments with respect to the realism of this sequence of events are in order: First, in reality, even without the firm’s announcement, consumers can infer the fill rate $\phi$ and the induced price distribution $g(p)$ over repeated interactions. Heidhues and Koszegi (2013) propose that even if a firm does not announce its induced price distribution, it can develop a stable “reputation” for having committing to a price distribution. Second, we assume that the firm can observe the demand realization at the beginning of each sales horizon before setting the price. The same stylized assumption is made by many recent papers, e.g., Cachon and Feldman (2013). In addition, we note that it is quite often that after observing a relatively small fraction of the total demand, an accurate forecast of the total can be obtained (see, e.g., Fisher and Raman 1996). Thus, the assumption that the firm observes the actual demand at the beginning of the period may be less restrictive than it seems to be.

3.1. Consumer’s Problem

A consumer’s expected utility is the sum of her expected consumption utility and her expected gain-and-loss utility. The expected consumption utility results from the expected consumption outcome, which depends on the consumer’s purchase decision and the availability of the product. The expected gain-and-loss utility captures the consumer’s loss aversion when she compares a realized consumption outcome to other possible outcomes in her reference distribution.

Before we discuss consumer’s decision making in further detail, we briefly discuss how consumer’s utility and reference points are derived. Let the binary variable $b \in \{0, 1\}$ denote consumer’s purchase outcome, where $b = 1$ indicates that the consumer successfully procures the product and $b = 0$ otherwise. Note that the outcome $b = 0$ occurs either because a consumer chooses not to purchase or because the product is unavailable even if the consumer chooses to purchase. A consumer’s utility function has two components: product and money. Denote the consumption outcome by $k = (k^v, k^p)$, where $k^v = vb$ is the valuation drawn from a purchase outcome, and $k^p = -pb$ is the minus of the payment made for a purchase outcome. Hence, the combined consumption utility is

$$C(k) = k^v + k^p = (v - p)b.$$

In addition, a loss averse consumer compares her actual consumption outcome $k = (k^v, k^p)$ to a possible consumption outcome $r = (r^v, r^p)$ (i.e., a reference point) in her reference point distribution, where $r^v$ is the reference product valuation and $r^p$ is the reference out-of-pocket cost. The advantage of this framework proposed by Koszegi and Rabin (2006) is that the stochastic reference points are endogenized by the decision maker, rather than being arbitrary or exogenously dictated. That is, a change in price or availability changes consumers’ expectation of possible outcome, leading to a change in their purchase plan. Given the firm’s contingent pricing policy (1), there exist three reference points, i.e., $r \in \{(v, -s), (v, -\bar{p}), (0, 0)\}$, if a consumer plans to buy at both the full and
sale price. Comparing her actual consumption outcome to a reference point, the consumer obtains a gain-and-loss utility along both dimensions of product and money:

\[ W(k|r) = \eta (k^v - r^v)^+ + \eta \lambda (k^v - r^v)^- + \eta (k^p - r^p)^+ + \eta \lambda (k^p - r^p)^- , \]

where \( \eta > 0, \lambda > 1, a^+ = \max\{a, 0\} \) and \( a^- = \min\{a, 0\} \) for any real number \( a \). Note that \( \lambda \geq 1 \) implies that the consumer feels losses stronger than equally sized gains. Therefore, customer’s total utility of a consumption outcome \( k \) conditional on a reference point \( r \) is

\[ u(k|r) = C(k) + W(k|r) . \]

As a consumer’s reference is her probabilistic beliefs about the possible outcomes, we use \( \Gamma(\cdot) \) to denote the probability distribution over \( r \). We call \( \Gamma(\cdot) \) the customer’s reference distribution in order to distinguish it from a deterministic reference point. Therefore, the expected utility of a consumption outcome \( k \) conditional on the customer’s reference distribution is

\[ U(k|\Gamma) = \sum_r u(k|r) \cdot \Gamma(r) . \] (3)

The utility function (3) indicates that in evaluating an actual consumption outcome, the consumer compare this outcome to every other possible outcome in her reference distribution.

Heidhues and Koszegi (2013) show that the consumer’s purchase plan follows a cut-off structure: the consumer chooses to buy at any price lower than or equal to the cut-off price, and not to buy at any price higher. Then, to induce consumers to always make a purchase, the full price \( p \) must be the cut-off price (we will use them interchangeably) and satisfy:

\[ U((v,-p)|\Gamma) = U((0,0)|\Gamma), \] (4)

where the random reference distribution \( \Gamma \) is endogenously induced by the purchase plan with the cut-off price \( p \) (see Heidhues and Koszegi 2013, Definition 1).

Equation (4) holds even if product unavailability is taken into account. At any cut-off price \( \bar{p} \), the utilities of purchasing and not purchasing, conditional on the same reference distribution \( \Gamma \), should be equal. Explicitly, if the availability of the product is \( \phi \) when the consumer intends to make a purchase, equating the two utilities gives

\[ \phi \cdot U((v,-p)|\Gamma') + (1 - \phi) \cdot U((0,0)|\Gamma') = U((0,0)|\Gamma') , \]

where \( \Gamma' \) is the reference distribution when unavailability is taken into account. The reduced equation is in the same form as (4).

Since each purchase plan is uniquely determined by the cut-off price \( \bar{p} \), with a little abuse of notation, we use \( \bar{p} \) to denote the consumer’s purchase plan. The purchase plan serves as a personal equilibrium (PE) (see Heidhues and Koszegi 2013). For consumers, alternative purchase plans such as not to purchase regardless of prices and to purchase only at the sale price can also serve as personal equilibria. However, they lead to trivial results. Hence, we restrict our focus to the personal equilibrium in which consumers can be induced to buy at both full and sale prices.
Now, we explain how a consumer forms her reference distribution $\Gamma$, based on her plan as well as the information about $\phi$ and $g(p)$, which are provided by the firm. The consumer anticipates to find the product available at the sale price $s$, and the full price $p$ with probabilities $\phi g(s)$ and $\phi g(p)$, respectively. In both cases, the consumer expects to buy. So, the outcome vector is $(v, -s)$ when $p = s$ and $(v, -p)$ when $p = \bar{p}$. In addition, the consumer expects the product to be unavailable with probability $1 - \phi$, in which case her expected outcome vector is $(0, 0)$. Thus, the consumer’s reference distribution, $\Gamma(\cdot)$, is

$$
\Gamma(r; g(\cdot), \phi, p) = \begin{cases} 
\phi g(s) & \text{if } r = (v, -s), \\
\phi g(p) & \text{if } r = (v, -p), \\
1 - \phi & \text{if } r = (0, 0).
\end{cases}
$$

We use $\Gamma(g, \phi, p)$ to denote the reference distribution induced by consumer’s purchase plan $p$. We can characterize the value of the credible cut-off price $p$ by using (3), (4) and (5) as follows.

**Lemma 1** Given pricing scheme (1) and inventory, $q$, the threshold (cut-off) price $p$ satisfies

$$(v - p) + (\eta v - \eta \lambda p) + \eta (\lambda - 1) \sum_{p=\bar{p}, s} (v + p) \cdot \phi g(p) = 0,$$

so the cut-off price $p$ can be written as

$$
p = v + \frac{\phi s \int_{\Omega} dF(x) - [1 - \phi (1 + \int_{\Omega} dF(x))]}{1 + \eta [\lambda - \phi (\lambda - 1) \int_{\Omega} dF(x)]} (\lambda - 1) \eta.
$$

3.2. Firm’s Problem

The firm may want to deliberately ration the order quantity among consumers (see, e.g., Liu and van Ryzin 2008) and may be more likely to do so when consumers are loss averse (see Liu and Shum 2013). Let $\xi(p, q, x)$ denote the availability of product to consumers when the full price is $p$, the order quantity is $q$ and the demand is realized as $x$. Given consumers’ purchase plan $p$, the firm will optimally set the full price at $p$ and need to further decide the order quantity $q$, the set of consumer demand realizations $\Omega$ that triggers the sales, and product availability $\xi(p, q, x)$ at each demand realization $x$, to maximize its expected profit. Since the fill rate is dependent on both consumers’ purchase plan and the inventory level, we use $\phi(p, q)$ to emphasize such dependency.

By our definition, the fill rate can be written as

$$
\phi(p, q) = \int_{0}^{\infty} \xi(p, q, x) dF(x).
$$

Accordingly, the firm’s expected profit can be written as

$$
\Pi(p, q, \Omega, \xi(p, q, x)) = sq \int_{x \in \Omega} dF(x) + p \int_{x \in \Omega} \xi(p, q, x) xdF(x) - cq.
$$

We note that: First, the firm can always clear the market at price $s$ and collect the expected profit $sq \int_{x \in \Omega} dF(x)$, no matter how consumers choose their purchase plans. However, the firm can
collect the additional profit \( \bar{p} \int_{x \in \Omega} \xi(\bar{p}, q, x) x dF(x) \), only when consumers are willing to purchase the product at the full price \( \bar{p} \). Second, to simply the analysis and the exposition, we assume that consumer’s decision making is influenced by the fill rate only, namely, the \textit{aggregated} product availability, but not by the product availability at each price. This assumption is reasonable because in practice most consumers are only sophisticated to a certain degree. However, we can see from (7) that the same fill rate \( \phi(\bar{p}, q) \) can result from different selections of the product availability function \( \xi(\bar{p}, q, x) \), and by (8), that the selection of the product availability function will affect the firm’s profit. Therefore, when determining the product availability for each demand realization \( x \), the firm should not only consider its impact on consumers’ perception, but also consider its impact on profitability. In summary, the firm’s problem is \( \max_{(\bar{p}, q, \Omega, \xi(\bar{p}, q, x))} \Pi(\bar{p}, q, \Omega, \xi(\bar{p}, q, x)) \) subject to that the fill rate \( \phi(\bar{p}, q) \) is consistent with the firm’s choice of \( \xi(\bar{p}, q, x) \) (see (7)).

4. Market Equilibrium

In the market equilibrium, neither the firm nor consumers have incentives to deviate from their decisions. Specifically, given the firm’s decisions on contingent pricing, inventory, and fill rate, consumers’ purchase plan is a personal equilibrium; in turn, given consumers’ purchase plan, the firm’s decisions maximize its expected profit.

4.1. Optimal Pricing Policy

This section characterizes the firm’s optimal contingent pricing scheme for any given order quantity level \( q \). (We study the optimal choice of \( q \) in §4.2.) Additionally, we study how the pricing decision changes with respect to the loss aversion level \( \lambda \) and order quantity \( q \). To streamline the analysis, we assume:

**Assumption 1** \( \frac{v}{s} \geq \frac{1 + \eta \lambda}{1 + \eta} \).

Assumption 1 requires that consumers’ valuation of the product is sufficiently higher than the sale price \( s \) in proportion that is dictated by their degree of loss aversion. If \( \eta = 1 \) and \( \lambda = 2 \) as suggested by empirically observed values of the loss aversion levels (see Ho and Zhang 2008), then we require \( v \geq \frac{3}{2} s \). Under this assumption, we can avoid the trivial solution that the optimal full price collapses to the sale price. In the rest of the paper, we also ignore the trivial cases that the optimal contingent price scheme degenerates to a single price and consumers’ equilibrium purchase plan is to only purchase at the sale price. Hence, the following results (Propositions 1 to 7) are only applicable within the range where the optimal sales frequency is \textit{nonzero}, as is usually observed in practice. Similarly, our numerical results in Section 4.3 are focused on this range as well.

**Proposition 1** (Optimal Contingent Pricing Policy) \textit{Given any initial order quantity \( q \), the optimal contingent pricing policy must have the following structure, which induces consumers to purchase at both prices.}
(i) The pricing scheme is of a threshold form:

\[ p = \begin{cases} 
    s & \text{if } x \leq \tau^*, \\
    \overline{p} & \text{if } x > \tau^*,
\end{cases} \]

where

\[ \overline{p} = v + \frac{\phi^* s F(\tau^*) - [1 - \phi^*(2 - F(\tau^*))]}{1 + \eta[\lambda - \phi^*(\lambda - 1)(1 - F(\tau^*))]} (\lambda - 1) \eta > s, \]

(9)

\( F(\tau^*) \) is the likelihood of running a sales at the sale price \( s \), namely, optimal sales frequency, and

\[ \phi^* = \int_0^\infty \xi(\overline{p}, q, x) dF(x) = \int_0^\infty \frac{\min(x, q)}{x} dF(x) \]

(10)

is the optimal fill rate given the order quantity \( q \).

(ii) The nonzero optimal sales threshold \( \tau^* \leq q \) is the solution to

\[ sq - \overline{p} \tau + \frac{\partial \overline{p}}{\partial F} \int_\tau^\infty \min\{x, q\} dF(x) = 0, \]

(11)

where \( \partial \overline{p} / \partial F \) is the derivative of the optimal full price with respect to sales likelihood \( F(\tau^*) \).

Proposition 1 provides several insights into the optimal contingent pricing policy. First, the optimal two-price scheme is in the form of a threshold structure: the firm will not run sales unless demand is sufficiently low. Second, since the firm’s optimal strategy is to sell as much as possible to consumers given inventory permits, the product is always available to consumers when sales take place by our assumption of priority rationing rules. The psychological gain from always having the product at the sale price reinforces consumers’ positive attachment effect; that is, the loss of not having the product at the full price increases consumers’ willingness to pay for the product at a high full price to prevent this loss. This effect still exists but is weaker when the firm cannot successfully prioritize consumers over bargain hunters. So our results can be thought of as a bound on the firm’s optimal profit in the absence of such prioritization. Third, when consumers are loss averse and demand is uncertain, the firm may price the product at a level higher than consumers’ intrinsic valuation. By contrast, the firm can only profitably charge the product at lower than or equal to the consumer valuation, when consumers are not loss averse. By (9), the optimal full price \( \overline{p}^* \) is greater than \( v \) if and only if the numerator of the second term in the right hand side of (9) is positive. The next proposition summarizes this result.

Proposition 2 (Attachment Effect and Fill Rate) Under the optimal contingent pricing policy, consumers pay a full price higher than their valuation, if and only if, the sales frequency \( F(\tau^*) \) and the fill rate \( \phi^* \) simultaneously satisfy the following two conditions:

\[ \phi^* > \frac{1}{2}, \]  
\[ F(\tau^*) \leq \frac{2\phi^* - 1}{\phi^*} \frac{v}{v - s} \equiv \widehat{F}. \]  

(12)  
(13)
The implications of Proposition 2 are twofold. First, Condition (12) indicates that the optimal fill rate needs to be relatively high in order to make buying at the full price an attractive purchase plan to form at the first place. Second, Condition (13) indicates that the optimal sales frequency needs to be sufficiently low to prevent consumers from not buying at the full price. The intuition is as follows, when considering the trade-off between the positive attachment effect and the negative comparison effect, we observe that: When the sales frequency increases, the consumers’ expectation of obtaining the product increases because of an increased sales chance. This attachment effect increases the consumers’ feeling of loss if they do not buy when the full price happens to be charged. However, when the sales frequency becomes overly high, then compared to the higher possibility of buying the product at the sale price $s$, the consumer views buying the product at the full price $p^*$ to be a loss. And, the more the sales likelihood is, the more loss the consumers feel if they buy the product at the full price. To dampen this negative comparison effect, the sales frequency should not be too high. As an immediate corollary of Proposition 2, we have the following result.

**Corollary 1 (When to Discount Less Frequently Than “Optimal”)** *Under Conditions (12) and (13), the optimal sales frequency is lower than the “optimal” sales frequency in the absence of consumers’ loss aversion (i.e., the sales frequency under the optimal contingent price strategy conditional on the realized demand of a sales season that maximizes the profitability for that particular period).*

Corollary 1 is in contrast to Cachon and Feldman (2013), where the firm benefits from discounting more frequently than the frequency under the “optimal” contingent discount policy given the realized demand. As mentioned in the literature review, the driving forces behind the two results are totally different. In our setting, frequent deep discounting is detrimental because it enhances the negative comparison effect and the firm would want to avoid such frequent discounting if consumers are loss averse.

Next, we investigate how the optimal full price $p^*$, the optimal sales threshold $\tau^*$, and the expected price change with respect to the initial order quantity level $q$ and the consumer loss aversion parameter $\lambda$. Temporarily ignoring the ordering cost, given a fixed consumer loss aversion parameter $\lambda$, we state the comparative statics with respect to the initial order quantity.

**Proposition 3 (Higher Availability, Larger Attachment Effect)** *Given any loss aversion level $\lambda$, the optimal full price $p^*$ is increasing\(^5\) and the optimal sales threshold $\tau^*$ is decreasing, in the order quantity $q$. Therefore, the expected price $\hat{p}^* = sF(\tau^*) + p^*(1 - F(\tau^*))$ is increasing in the order quantity $q$.*

\(^5\)Unless otherwise specified, the monotonicity is in its weaker sense.
The increasing monotonicity of the optimal full price in the order quantity \( q \) is in contrast to the common intuition of the price-quantity relationship in a consumer market without loss aversion (i.e., \( \lambda = 0 \)): As the firm increases its initial order, the fill rate increases, but the risk of unsold inventory (i.e., overstocking risk) increases as well. Hence, one would intuitively expect the firm to reduce the price to avoid unsold inventory. The seemingly anomaly is due to that in a loss-averse consumer market, the firm can take advantage of the positive attachment effect by creating moderate sales and the attachment effect is reinforced by higher product availability.

In a classic newsvendor setting with non-loss-averse consumers, the fill rate does not play a role in influencing consumers’ willingness-to-pay, and its impact on the firm’s profitability is always along the single dimension of inventory risks. In a loss-averse consumer market, however, an increase in the fill rate can increase consumers’ expectation of obtaining the product, and hence increases the attachment effect, i.e., the loss that consumers feel if they decide not to buy at the full price increases, because they are more attached to the idea of buying the product. Consequently, consumers are willing to pay a higher full price to avoid such a loss. Nevertheless, an increased fill rate also increases the overstocking risk to the firm. Overall, unlike in the classic newsvendor setting, Proposition 3 shows that the marginal understocking cost \((\hat{p}^* - c)\) is increasing in the order quantity. Therefore, the presence of consumer loss aversion allows the firm to mitigate the overstocking risk by raising the understocking cost. As a result, the firm tends to order more quantity and earn more from the loss-averse consumer market. Proposition 3 highlights the subtlety in determining the optimal order quantity when selling to a loss-averse consumer market, which we will discuss in §4.2.

Next, we consider the impact of consumers’ loss aversion on the firm’s optimal sales frequency and profitability. Note that given a fixed order quantity \( q \), the optimal fill rate \( \phi^* \) is not affected by consumers’ loss aversion \( \lambda \) explicitly; see Proposition 1(i). However, from Proposition 1(ii), we can see that the optimal sales threshold \( \tau^* \) depends on \( \lambda \) explicitly, and so does the optimal sales frequency \( F(\tau^*) \); we use the notation \( \tau^*(\lambda) \) and \( F(\tau^*(\lambda)) \) to emphasize this dependency.

**Proposition 4 (Comparative Statics on Loss Aversion Level)** Given any fixed order quantity \( q \), if \( F(\tau^*(\lambda)) < \hat{F} \), where \( \hat{F} \) is defined in (13), the optimal sales frequency \( F(\tau^*(\lambda)) \) is decreasing in \( \lambda \) and the optimal full price \( \bar{p}^* \) is increasing in \( \lambda \).

Proposition 4 says that the optimal sales frequency \( F(\tau^*(\lambda)) \) is decreasing in \( \lambda \) when it is below the threshold \( \hat{F} \). As discussed above, the attachment effect can increase consumers’ willingness-to-pay, so it has a positive impact on the firm’s profitability. The comparison effect, in contrast, restrains the firm from charging a higher full price, so it has a negative impact on the firm’s profitability. As consumers’ loss aversion increases, both effects become more significant. Proposition 4 shows that when the optimal sales frequency is not too high, as consumers’ loss aversion increases, the firm should lower the optimal sales frequency \( F(\tau^*) \) to dampen consumers’ feeling of loss caused
Recall that $F(\tau^*) < \hat{F}$ is a necessary condition of that the optimal full price is higher than consumers’ valuation of product $v$ (see Proposition 2). Hence, Proposition 4 also suggests that if the optimal full price is greater than $v$ for some level of loss aversion, it will be always greater than $v$ if consumers become more loss averse. In view of (13), the condition that $F(\tau^*) < \hat{F}$ is more likely to hold when the initial order level, $q$, is higher and when the valuation ratio, $v/s$, is larger.

Figure 2(a) shows a numerical example that illustrates Proposition 4. Parameters $s = 2.5, c = 3, v = 4$ are used for Figures 2(a) and 2(b), and parameters $s = 1.5, c = 2, v = 4$ are used for Figure 2(c). Demand follows a normal distribution $N(50, \sigma)$, where $\sigma \in \{15, 17\}$ (so the chance of negative demand is negligible). In Figure 2, the primary (left) y-axis is used for the sales threshold and the full price, while the secondary (right) y-axis is used for the firm’s profit. One setting that fits these parameters can be a grocery store selling fruit plates to patrons. Products like fruit plates have very short life cycle and consumers have high frequent purchases, so it is natural for consumers to form a reference price distribution. We use $\eta = 1$ and vary consumer’s loss aversion parameter, $\lambda$, from 1 to 3 to be consistent with the empirical evidence (see Ho and Zhang 2008). In Figure 2(a), $q$ is high leading to $\phi^* = 0.8335$ and $\hat{F} = 2.1340$, so given any sales threshold $\tau^*(\lambda)$, $F(\tau^*(\lambda)) < \hat{F}$ holds. The claim of Proposition 4 is illustrated by Figure 2(a), where the sales threshold is strictly decreasing while the full price is strictly increasing in the consumers’ loss aversion parameter. However, if the condition $F(\tau^*(\lambda)) < \hat{F}$ of Proposition 4 is violated, both Figure 2(b), where $q = 20$ results in $\phi^* = 0.4518$ (so even (12) is not satisfied) and $\hat{F} = -0.5687$, and Figure 2(c), where $q = 22.81$ but a lower variance of demand results in $\phi^* = 0.5002$ and $\hat{F} = 0.0012$, show that the sales threshold could be decreasing in consumer loss aversion, the full price could be strictly decreasing (see Figure 2(b)) or be constant (see Figure 2(c)) in consumer loss aversion, while the profit could be strictly decreasing (see Figure 2(b)) or have a $U$-shape (see Figure 2(c)) in consumer loss aversion.
The next proposition discusses how the firm’s optimal profit changes with respect to the consumers’ loss aversion level. Given a fixed initial inventory level $q$, the firm’s revenue under the optimal contingent pricing policy is

$$\Pi^* = \Pi(p^*, q, \tau^*, \xi(p^*, q, x)) = sqF(\tau^*) + p^* \int_{\tau^*}^{\infty} \min\{x, q\} dF(x).$$ (14)

We can see that the profitability depends on the loss aversion level, through the optimal sales threshold. Let $\tau^*(1)$ be the value of the sales threshold $\tau^*(\lambda)$ at $\lambda = 1$ and $\tau^*(\infty) = \lim_{\lambda \to \infty} \tau^*(\lambda)$, then we have the following results.

**Proposition 5 (When Loss Aversion Benefits)** Given any fixed order quantity $q$,

(i) If $F(\tau^*(1)) \leq \hat{F}$, the firm’s profit $\Pi^*$ is strictly increasing in $\lambda$;

(ii) If $F(\tau^*(\infty)) \geq \hat{F}$, the firm’s profit $\Pi^*$ is strictly decreasing in $\lambda$;

(iii) Otherwise, the firm’s profit $\Pi^*$ has a U-shape in $\lambda$, with a unique minimum $\lambda_{\min}$ such that $F(\tau^*(\lambda_{\min})) = \hat{F}$.

Proposition 5 states that the firm’s profit is either strictly monotone (increasing or decreasing) or it has a unique minimum in the loss-aversion parameter. The firm’s profitability is influenced by the two competing effects – the positive attachment effect and the negative comparison effect. Both effects become more significant when consumer loss aversion increases. Part (i) states that if the sales frequency in the loss neutral market (i.e., $\lambda = 1$) is less than the threshold $\hat{F}$, then the positive attachment effect always dominates, and the firm’s profit will increase as the consumers become more loss averse. In this case, the firm strictly benefits from loss aversion. Again, this is more likely the case when the initial order level $q$ is higher and when the valuation ratio $v/s$ is larger. Part (ii) specifies an extreme where the negative comparison effect always dominates. In this case, the firm’s profit is hurt by the consumers’ loss averse behavior. By the definition of $\hat{F}$, this extreme is more likely the case when the initial order level $q$ is lower and when the valuation ratio $v/s$ is smaller. Part (iii) shows an intermediate case that may be more common in practice, where the negative comparison effect first dominates for lower levels of loss aversion, and beyond some point, the positive attachment effect takes over. In this case, the firm may benefit from consumers’ loss aversion, if the level of loss aversion is sufficiently high. The three cases of Proposition 5 are illustrated by Figures 2(a), 2(b) and 2(c) respectively.

### 4.2. Optimal Order Quantity

In this section, we discuss the firm’s choice of the order quantity, given that the firm implements the optimal contingent pricing scheme specified in Proposition 1 for a given order quantity. According to the pricing scheme in Proposition 1, the firm’s decision of order quantity $q$ determines the optimal fill rate $\phi^*$; the optimal fill rate $\phi^*$ and the firm’s decision of sales threshold $\tau$ will jointly determine the optimal full price $p^*$ and the sales frequency $F(\tau)$. Therefore, the firm’s expected
profit $\Pi$ can be seen as a function of the firm’s two decision variables $q$ and $\tau$. Given a fixed order quantity $q$, the optimal sales threshold $\tau^*$ should satisfy the first order condition (11). Given a fixed sales threshold $\tau$, the partial derivative of the expected profit function $\Pi(\cdot)$ with respect to $q$ is

$$
\frac{\partial \Pi(q, \tau)}{\partial q} = sF(\tau) + p^*(1 - F(q)) + \frac{\partial p^*}{\partial x} \frac{\partial \phi^*}{\partial q} \int_{\tau}^{\infty} \min\{x, q\} dF(x) - c. \quad (15)
$$

Note that $\frac{\partial \Pi(0, \tau)}{\partial q} = sF(\tau) + p^* - c > 0$. In addition, since $\lim_{q \to \infty} (1 - F(q)) = 0$ and $\lim_{q \to \infty} \frac{\partial \phi^*}{\partial q} = 0$, we have $\lim_{q \to \infty} \frac{\partial \Pi(q, \tau)}{\partial q} = sF(\tau) - c < 0$. Therefore, based on Bolzano’s Theorem, there exists an optimal order quantity $q^* > 0$ at which point $\frac{\partial \Pi(q^*, \tau)}{\partial q} = 0$. Given the consumer loss aversion level $\lambda$ and the procurement cost $c$, the optimal order quantity $q^*$ and the optimal sales threshold $\tau^*$ must simultaneously satisfy the two first order conditions (11) and (15).

By examining the solutions to these two equations, we have the following results on comparative statics of optimal order quantity with respect to the loss aversion level and procurement cost.

**Proposition 6 (More Loss Aversion, Larger Initial Order)** One of the following two scenarios must prevail:

(i) The firm’s optimal order quantity $q^*$ is increasing in $\lambda$;

(ii) The firm’s optimal order quantity $q^*$ has a U-shape in $\lambda$.

Proposition 6 indicates that there is a threshold on the loss aversion level $\lambda$, beyond which the more loss averse consumers are, the higher the firm should set its initial order quantity. This result is consistent with Proposition 3, since a higher initial order quantity would lead to a higher fill rate and hence enhance the positive attachment effect. This insight is different from the message in Liu and Shum (2013) that limiting supply in a loss-averse (in their words, disappointment-averse) market can benefit the firm. In Liu and Shum (2013), loss averse behavior of high-valued consumers with a deterministic reference point can reinforce intertemporal market segmentation. However, we demonstrate that even with a single segment of consumers of a homogenous valuation but in the presence of stochastic reference points, the firm is better off by expanding supply to the market as much as possible. The ample supply coupled with contingent sales can induce the loss-averse consumers with stochastic reference points to buy at the (higher) full price.

Next, we consider how the firm’s optimal order quantity decisions may change with respect to the procurement cost. One may typically expect that the firm would sell the product at a higher price if it procures the units more expensively. This is true for a market with non-loss-averse consumers. However, if the firm sells to consumers who are loss-averse, it should sell at a cheaper price when the procurement cost increases. The next proposition states this result.

**Proposition 7 (Higher Cost, More Sales and Lower Full Price)** The following statements hold.

(i) The optimal order quantity $q^*$ is decreasing in the procurement cost $c$;
(ii) The optimal sales threshold $\tau^*$ is increasing in the procurement cost $c$;
(iii) Both the optimal full price $p^*$ and the expected price $\hat{p}^*$ are decreasing in the procurement cost $c$.

Part (i) of Proposition 7 says that when the procurement cost increases, the firm orders less. This result is consistent with a market without loss-averse consumers. Surprisingly, parts (ii) and (iii) of Proposition 7 show that a higher procurement cost would lead to a higher optimal sales frequency (due to an increased sales threshold) and a lower optimal full price, both of which are in consumers’ favor. In particular, the finding that the optimal full price decreases in the procurement cost is in stark contrast to predictions by the marketing literature on cost pass-throughs in a non-loss-averse consumer market: the optimal full price is (i) equal to consumers’ valuation regardless of the procurement cost if there is a single segment of consumers with a homogenous valuation, or (ii) increasing in the procurement cost if there are multiple or a continuum of heterogenous consumer valuations. The reason behind this counterintuitive finding is that an increase in the procurement cost causes the firm to reduce the optimal order quantity (see part (i)). By Proposition 3, less availability reduces consumers’ willingness-to-pay at the full price, and the firm compensates for this reduction by reducing the optimal full price. Another natural response to reduced inventory availability is to slightly increase the sales frequency to boost the attachment effect while sustaining the full price.

4.3. Numerical Results
In this subsection, we numerically illustrate the interesting observation that small or medium degrees of demand variability could be in the firm’s favor when consumers are loss averse. When the newsvendor sells to loss neutral consumers, the firm can use the contingent pricing strategy to mitigate the negative impact of demand variability on profitability. But regardless of how successful the mitigation is, demand variability always cuts into the firm’s profitability, so the firm always prefers deterministic demand to uncertain demand in the canonical newsvendor setting. However, this is no longer true when the newsvendor sells to loss averse consumers with stochastic reference points. In this case, the contingent pricing strategy not only mitigates the negative impact of demand variability, but also induces consumers to accept a price higher than their intrinsic valuation. Therefore, the benefit of contingent pricing can be higher in the loss averse market than in the loss neutral market. This, again, stresses the important role of consumer loss-averse behavior in studying firm’s inventory and pricing decisions.

Figure 3 illustrates three representative numerical examples showing how the firm’s profit changes in response to various levels of demand variability for a given initial order quantity, when the firm uses the optimal contingent pricing specified in Proposition 1. The parameters chosen for these numerical examples are $s = 1.5, c = 2, v = 4, \eta = 1, \lambda = 2$, a normally distributed random market size $D$ with mean $\mathbb{E}(D) = 50$, and we allow the standard deviation $\sigma$ of $D$ to vary. In Figure
Figure 3  Impact of Demand Variability on Profit: $D \sim N(50, \sigma^2)$

3(a), the firm’s profit with demand variability $\sigma \in (0, 17)$, represented by the solid line, is higher than the firm’s profit with no demand variability $\sigma = 0$, represented by the dashed line. Similar findings emerge in Figures 3(b) and 3(c), where the firm earns a higher profit when $\sigma \in (0, 7.7)$ and $\sigma \in (0, 3.1)$ respectively. Therefore, in these cases, the firm prefers some level of demand variability to no demand variability. The same observation can remain at the optimal initial order level.

This observation is in stark contrast to the result in the classic newsvendor setting. When consumers are not loss averse, the optimal contingent pricing strategy is to set $p = s$ when demand realization is lower than $sq/v$ and to set $p = v$ otherwise. Although the use of this contingent pricing strategy partly mitigate the negative impact of demand variability on the firm’s profitability, the firm would never be able to achieve the same level of profit as in the case of no demand variability. On the contrary, when consumers are loss averse, certain levels of demand variability would lead to moderate sales, which induce consumers to accept a higher full price. When the loss of revenue due to sales and demand variability is more than compensated by this higher full price, the firm’s profit will increase. On the other hand, Figure 3 also suggests that when the demand variability is sufficiently high, the negative impact of demand variability would eventually dominate the positive benefit of charging a higher full price. In conclusion, these numerical examples suggest that small or medium degrees of demand variability may benefit the firm, but sufficiently high demand variability is always detrimental. We note that there exist examples in which any demand variability hurts the firm.

In addition, we explore the interactions of the sale price $s$, the loss aversion parameter $\lambda$ and the demand variability $\sigma$, on the firm’s profitability. Numerical analysis suggests that the higher the sale price is, the wider the range of demand variability that benefits the firm. Indeed, a higher sale price not only generates a higher sales revenue when demand realization is low, but also yields a higher full price according to Proposition 1. The same observation holds for the loss aversion
Figure 4  Profit under Optimal Order Quantity

parameter, namely, the higher the loss aversion level, the wider is the range of demand variability where the firm is better off.

Lastly, Figure 4 plots numerical examples in which the firm strictly benefits from consumer’s loss aversion when the initial order quantity is optimized. For each value of the loss aversion parameter $\lambda$, both the initial order quantity and the sales threshold are optimized to maximize the firm’s profit, assuming that buying at both full and sale prices is consumers’ personal equilibrium. We can see that the profit is strictly increasing in the consumer loss aversion level. This monotonicity of profitability in consumer loss aversion is indirectly contributed to the strictly increasing optimal order quantity in increasing the product availability, and is directly due to the strictly increasing optimal full price. However, we also observe that the optimal sales threshold may not be monotone in the consumer loss aversion level. Numerical examples, in which the parameters $v, c$ and $s$ vary for other demand distributions, demonstrate similar patterns.

5. Conclusion

Consumers’ loss-averse behavior with endogenized stochastic reference levels has largely been overlooked in the operations management literature. In addition, in the behavioral economics literature, the loss-averse consumers are assumed to have a deterministic market size while the firms are assumed to have no supply constraints. We attempt to fill the gap between these two bodies of literature by considering the trade-offs in the inventory and contingent pricing policies of a newsvendor-type firm over repeated sales horizons. The firm orders a limited amount of inventories and sells them to loss-averse consumers with a random market size whose stochastic price reference levels are influenced by the firm’s pricing strategies. We fully characterize the optimal inventory and pricing policies, which include a contingent pricing strategy of a threshold form. We demonstrate that the firm can benefit from consumers’ loss averse behavior. The model reveals
somewhat counterintuitive insights on how consumers’ loss aversion affects the firm’s optimal decisions. Specifically, we show that when consumers are loss averse, (i) the optimal full price increases in the order quantity; (ii) the optimal full price decreases in the procurement cost; and (iii) the optimal sales frequency increases in the procurement cost. Again, we caution that these results hold up to the point where the optimal sales frequency becomes degenerate to zero.

There are several limitations of our model. First, to simplify analysis, we assume that under contingent pricing, the firm can quickly learn the demand for the entire period. However, in reality, demand uncertainty may unfold over time and the firm may not learn it all at once. Second, we assume that the sales horizons unfold repeatedly with a stationary demand distribution over time. In reality, market conditions can shift as time goes by. Third, we assume that consumers do not learn the firm’s contingent pricing scheme by associating observed price and fill rate to possible demand realizations. The strategic learning behavior of consumers may lead to a different reference distribution from what we studied. Fourth, we consider the equilibrium behavior over repeated interactions between loss averse consumers and a forward-looking monopolistic newsvendor. Hence, we caution that our comparative statics results may not be applicable in predicting transient or competitive market interactions.

Despite these limitations, our stylized model captures the core tensions of how consumers’ loss aversion with endogenized stochastic price-reference points influences the firm’s optimal inventory and pricing decisions. The obtained insights may have several implications. For example, as limited initial inventory hinders the firm from profitably selling at the full price to loss averse consumers, our results suggest that the firm may want to aggressively build up initial stocks, more than when selling to consumers who are not loss averse. Moreover, as running sales may profitably manipulate loss-averse consumers’ reference prices, the firm can be satisfied with limited replenishment ability, because beneficial occasional sales can be a natural outcome of the contingent pricing strategy driven by market uncertainty. As a result, the additional benefit of quick replenishment initiatives beyond contingent pricing policies may be limited in a market with loss averse consumers.

Appendix. A. Proofs.

Proof of Lemma 1. When the consumer faces price $p$ and decides to buy (i.e., the left hand side of (4)), then she buys it if it is available (i.e., the outcome vector becomes $k = (v, -p)$). Thus, she obtains a consumption utility $(v - p)$. In addition, the consumer gains a gain-loss utility by comparing buying at $p$ to the other two possibilities in her reference distribution as follows: (i) compared to the possibility that the consumer could have obtained the product at the sale price $s < p$ (i.e., $r = (v, -s)$), buying the product at $p$ does not generate any feeling of gain or loss in the product since in both cases consumer obtains the product. However, the consumer feels a loss of $(p - s)$ dollars in money (i.e., $k - r = (0, s - p)$). Hence, relative to
this possibility, she obtains a gain-loss utility $-\eta \lambda (p - s) \phi g(s) < 0$. (ii) Compared to the possibility that the consumer might not have found the product available (i.e., $r = (0, 0)$), buying at $\bar{p}$ is the gain of the value $v$ for obtaining the product and the loss of $\bar{p}$ dollars (i.e., $k - r = (v, -\bar{p})$). Hence, relative to this possibility the consumer obtains a gain-loss utility $(\eta v - \eta \lambda \bar{p})(1 - \phi)$. Thus, according to (3), the consumer’s total utility from buying at $\bar{p}$ is

$$U((v, -\bar{p})|G(g, \phi, \bar{p})) = (v - \bar{p}) - \eta \lambda (\bar{p} - s) \cdot \phi g(s) + (\eta v - \eta \lambda \bar{p}) \cdot (1 - \phi),$$

Now, if the consumer faces price $\bar{p}$ and decides not to buy (i.e., the right hand side of (4)) then the outcome vector becomes $k = (0, 0)$, and she does not obtain any consumption utility. However, she obtains a gain-loss utility $(\eta s - \eta \lambda v) \phi g(s)$, from comparing not buying with the possibility that she could have obtained the product at the sale price $s$ (i.e., $r = (v, -s)$), which is assessed as a gain of $s$ dollars and a loss of $v$ for not obtaining the product. Similarly, the consumer feels a gain-loss utility $(\eta \bar{p} - \eta \lambda v) \phi g(\bar{p})$ from not buying compared to the possibility that she was expecting to buy at $\bar{p}$. Hence, the consumer’s total utility from not buying at $\bar{p}$ (the right hand side of (4)) is

$$U((0, 0)|G(g, \phi, \bar{p})) = \sum_{p=\bar{p},s} (\eta p - \eta \lambda v) \cdot \phi g(p).$$

By equating $U((v, -\bar{p})|G(g, \phi, \bar{p}))$ and $U((0, 0)|G(g, \phi, \bar{p}))$, and after some algebra, we have the desired result. □

**Proof of Proposition 1.** First, we characterize the full price and optimal fill rate when $q$ is given. If the firm can successfully induce consumers to buy at both prices, then according to (6) in Lemma 1, the full price $\bar{p}^*$ must satisfy

$$\bar{p}^* = v + \frac{\phi \psi s - (1 - \phi (2 - \psi)) v}{1 + \eta [\lambda - \phi (\lambda - 1)(1 - \psi)]}(\lambda - 1)\eta, \quad (16)$$

where $\psi$ is the probability of $s$ being charged. We see from (16) that $\bar{p}^*$ is increasing in $s$, which justifies the modeling assumption that the sale price should not be lower than $s$. By Assumption 1, it follows that $\bar{p}^* > s$. In this case, the pricing scheme is

$$p = \begin{cases} s & \text{if } x \in \Omega, \\ \bar{p}^* & \text{if } x \in \Omega^c, \end{cases}$$

where $\Omega$ satisfies $\int_{\Omega} dF(x) = \psi$. Note that the fill rate $\phi$ and sales frequency $\psi$ are two independent decision variables for now.
Therefore, regardless how \( \psi \) to consumers. As a result, the optimal fill rate should be \( \phi \) products to consumers under any demand realization \( x \), i.e., the best strategy to set the product availability for each demand realization is \( \xi(p^*, q, x) = \min_{x, q}(x, y) \). This strategy increases both the full price and the sales to consumers. As a result, the optimal fill rate should be \( \phi^* = \int_0^\infty \xi(p^*, q, x) dF(x) = \int_0^\infty \min_{x, q}(x, y) dF(x) \).

Next, the remaining task is to determine the sales frequency and the set of demand realizations chosen for sales. For a fixed \( \psi \), we show that it is optimal for the firm to set \( \Omega = [0, \tau] \) such that \( \psi = \int_0^\tau dF(x) \). First, note that given any \( \Omega \) that satisfies \( \int_{\Omega} dF(x) = \psi \), the firm’s expected profit is

\[
\Pi = sq \int_{\Omega} dF(x) + \bar{p}^* \int_{\bar{t}}^\infty \xi(\bar{p}^*, q, x) x dF(x) - cq = sq\psi + \bar{p}^* \int_{\bar{t}}^\infty \min_{x, q}(x, y) dF(x) - cq,
\]

Hence, to maximize profit is equal to maximize \( \int_{\Omega} \min_{x, q}(x, y) dF(x) \). Subject to \( \int_{\Omega} dF(x) = \psi \), the best choice of \( \Omega \) to maximize \( \int_{\Omega} \min_{x, q}(x, y) dF(x) \) is \( \Omega = [0, \tau] \) then \( \psi = \int_{\Omega} dF(x) = F(\tau) \), and \( \int_{\Omega} \min_{x, q}(x, y) dF(x) = \int_{\tau}^\infty \min_{x, q}(x, y) dF(x) \). Consequently, the sales frequency \( \psi \) and the threshold for running sales \( \tau \) are inter-changeable decision variables to the firm. We will focus on determining \( \tau \).

Because \( \Omega = [0, \tau] \), the firm’s profit function can be rewritten as

\[
\Pi = sq \int_{0}^{\tau} dF(x) + \bar{p}^* \int_{\bar{t}}^\infty \min_{x, q}(x, y) dF(x) - cq.
\]

Taking derivative of \( \Pi \) with respect to \( \tau \) gives

\[
\frac{\partial \Pi}{\partial \tau} = f(\tau) (sq - \bar{p}^* \cdot \min_{\{\tau, q\}}) + \frac{\partial \bar{p}^*}{\partial \tau} \int_{\bar{t}}^\infty \min_{x, q}(x, y) dF(x).
\]

(17)

Next, invoking the chain rule to compute the derivative of \( \bar{p}^* \) with respect to \( \tau \) yields

\[
\frac{d\bar{p}^*}{d\tau} = \frac{\partial \bar{p}^*}{\partial F} \frac{dF(\tau)}{d\tau} = f(\tau) \frac{\partial \bar{p}^*}{\partial F}.
\]

(18)

Finally, applying (18) to (17) yields

\[
\frac{\partial \Pi}{\partial \tau} = f(\tau) \left[ sq - \bar{p}^* \cdot \min_{\{\tau, q\}} + \frac{\partial \bar{p}^*}{\partial F} \int_{\bar{t}}^\infty \min_{x, q}(x, y) dF(x) \right],
\]

(19)

Note that

\[
\frac{\partial \bar{p}^*}{\partial F} = \frac{\eta \phi(\lambda - 1)}{[1 + \eta \lambda - \eta \phi(\lambda - 1) + \eta \phi F(\tau)(\lambda - 1)]^2 \cdot \{s(1 + \eta \lambda) - v(1 + \eta) - \eta \phi(\lambda - 1)(v + s)\}} < 0,
\]

(20)
\[
\frac{\partial^2 \bar{p}^*}{\partial F^2} = -\frac{2\eta\phi(\lambda - 1)}{1 + \eta\lambda - \eta\phi(\lambda - 1) + \eta\phi F(\tau)(\lambda - 1)} \frac{\partial \bar{p}^*}{\partial F} > 0. 
\]

Equation (20) follows because according to Assumption 1, we must have \(s(1 + \eta\lambda) - v(1 + \eta) \leq 0\).

For \(\tau > q\), we have \(sq - \bar{p}^* \cdot \min\{\tau, q\} < 0\), so according to (19), we have \(\partial \Pi / \partial \tau < 0\), i.e., it is not optimal to set the sales threshold \(\tau\) above \(q\).

In summary, either \(\partial \Pi / \partial \tau < 0\) for all \(\tau \in [0, q]\), or there exists \(\tau^* \in [0, q]\) such that the first order condition is satisfied, i.e., \(\partial \Pi / \partial \tau|_{\tau^*} = 0\). Because we assume \(f(\tau) > 0, \forall \tau > 0\), (19) is equivalent to (11). □

Proof of Corollary 1. In the proof of Proposition 1, we have shown \(\partial \bar{p}^* / \partial F < 0\); see (20). Then by (11), \(sq - \bar{p}^* \tau^* \geq 0\). Hence, \(\tau^* \leq sq / \bar{p}^*\). Under Conditions (12) and (13), \(\bar{p}^* \geq v\) by Proposition 2 and then \(\tau^* \leq sq / v\). The result follows by realizing that \(sq / v\) is the “optimal” contingent discount threshold in the absence of consumers’ loss aversion, i.e., the optimal contingent pricing strategy is to set \(p = s\) when demand is lower than \(sq / v\) and to set \(p = v\) otherwise. □

Proof of Proposition 3. By (10), there is a one-to-one increasing correspondence between \(q\) and \(\phi^*\). Hence comparative statics with respect to \(q\) are equivalent to those with respect to \(\phi^*\), which we focus on in the remaining proof. Note that \(\phi^*\) affects \(\bar{p}^*\) through two channels: the first is the direct effect of \(\phi^*\) on \(\bar{p}^*\) and the second is the indirect effect through the sales frequency decision \(F(\tau^*)\) (or equivalently, the optimal threshold for sales \(\tau^*\)). Under the assumption that \(\tau^*\) is determined by (11), we can write \(\tau^*\) as a function of \(\phi^*\), i.e., \(\tau^*(\phi^*)\), and the full price is a function of \(\phi^*\) as \(\bar{p}^*(\phi^*, \tau^*)\).

To show this result, by the chain rule, we can write

\[
\frac{\partial \bar{p}^*}{\partial \phi^*} = \frac{\partial \bar{p}^*}{\partial \tau^*} \cdot \frac{\partial \tau^*}{\partial \phi^*}, \quad (21)
\]

where \(\partial \bar{p}^* / \partial \phi^*\) in (21) can be obtained from (9) as

\[
\frac{\partial \bar{p}^*}{\partial \phi^*} = \frac{[sF(\tau^*) + u(2 - F(\tau^*))] \cdot [s\eta\lambda F(\tau^*) + \eta\phi(\lambda + 1 - F(\tau^*))]}{(1 + \eta[\lambda - \phi^*(\lambda - 1)](1 - F(\tau^*)))^2} (\lambda - 1)\eta > 0.
\]

By (20), we know that \(\partial \bar{p}^* / \partial \tau^* = \partial \bar{p}^* / \partial F \cdot f(\tau^*) < 0\). We next show that \(\partial \tau^* / \partial \phi^* < 0\) by applying the Implicit Function Theorem\(^6\) to the first order condition in (11):

\[
\frac{\partial \tau^*}{\partial \phi^*} = -\frac{\partial^2 \Pi(\tau^*, \phi^*)}{\partial \tau^* \partial \phi^*} \Bigg/ \frac{\partial^2 \Pi(\tau^*, \phi^*)}{\partial \tau^* \partial \phi^*}.
\]

\(^6\) Here and below we note that the conditions required for the implicit function theorem to hold are all satisfied whenever the optimal sales threshold is determined by (11).
in which $\partial^2 \Pi(\tau^*, \phi^*)/\partial \tau^2 < 0$ as long as $\tau^*$ is a local interior maximizer. For $\partial^2 \Pi(\tau^*, \phi^*)/\partial \tau^* \partial \phi^*$, we have

$$
\frac{\partial^2 \Pi(\tau^*, \phi^*)}{\partial \tau^* \partial \phi^*} = f(\tau^*) \left\{ -\tau^*, \frac{\partial \bar{p}^*}{\partial \phi^*} + \frac{\partial^2 \bar{p}^*}{\partial \tau^* \partial \phi^*} \int_{\tau^*}^{\infty} \min\{x, q\} f(x) dx \right\}. \tag{23}
$$

We know that $\partial \bar{p}^*/\partial \tau^* = \partial \bar{p}^*/\partial F \cdot f(\tau^*)$, and the expression of $\partial \bar{p}^*/\partial F$ is (20) with $\phi^*$ replacing $\phi$ and $\tau^*$ replacing $\tau$. Note that the denominator of (20) is positive and decreasing in $\phi$, and the numerator of (20) is negative and decreasing in $\phi$, so $\partial \bar{p}^*/\partial F$ is decreasing in $\phi^*$, i.e., $\partial^2 \bar{p}^*/\partial \tau^* \partial \phi^* < 0$. Together with (23), we must have $\partial^2 \Pi(\tau^*, \phi^*)/\partial \tau^* \partial \phi^* < 0$, and by (22), we have $\partial \tau^*/\partial \phi^* < 0$.

Based on the above results, and by (21), we have $d\bar{p}^*/d\phi^* > 0$. The average price is $\hat{p}^* = F(\tau^*)s + (1 - F(\tau^*))\bar{p}^*$, then we have

$$
\frac{\partial \hat{p}^*}{\partial \phi^*} = (s - \bar{p}^*) f(\tau^*) \frac{\partial \tau^*}{\partial \phi^*} + (1 - F(\tau^*)) \frac{\partial \bar{p}^*}{\partial \phi^*} > 0. \quad \square
$$

*Proof of Proposition 4.* By the Implicit Function Theorem, we have

$$
\frac{\partial \tau^*}{\partial \lambda} = -\frac{\partial^2 \Pi^*(\tau^*, \lambda)}{\partial \tau^* \partial \lambda} \bigg/ \frac{\partial^2 \Pi^*(\tau^*, \lambda)}{\partial \tau^2}. \tag{24}
$$

Note that $\partial^2 \Pi^*(\tau^*, \lambda)/\partial \tau^2 \leq 0$ because $\tau^*$ is the local maximization. Hence, the sign of $\partial \tau^*/\partial \lambda$ is the same as $\partial^2 \Pi^*(\tau^*, \lambda)/\partial \tau^* \partial \lambda$. Taking derivative of (19) with respect to $\lambda$ gives

$$
\frac{\partial^2 \Pi^*(\tau^*, \lambda)}{\partial \tau^* \partial \lambda} = f(\tau^*) \left( \frac{\partial^2 \bar{p}^*}{\partial F \partial \lambda} \int_{\tau^*}^{\infty} \min\{x, q\} dF(x) - \frac{\partial \bar{p}^*}{\partial \lambda} \tau^* \right). \tag{25}
$$

Define

$$
B \equiv 1 + \eta[\lambda - \phi^*(\lambda - 1)(1 - F(\tau^*))], \quad C \equiv s\phi^* F(\tau^*) - (1 - \phi^*(2 - F(\tau^*)))v, \tag{26}
$$

then we have

$$
\frac{\partial^2 \bar{p}^*}{\partial F \partial \lambda} = \frac{\eta(1 + \eta)}{B^2} \phi^*(s - v) + \frac{1}{B^2} \left\{ 2\eta(\lambda - 1) \frac{\partial B}{\partial F} \frac{\partial B}{\partial \lambda} - \eta(\lambda - 1) B \frac{\partial^2 B}{\partial F \partial \lambda} - B \frac{\partial B}{\partial F} \right\} C. \tag{27}
$$

In the derivation of (27), we used that $\partial^2 C/\partial F \partial \lambda = 0, \partial C/\partial \lambda = 0$.

Define $\hat{\tau} = F^{-1} \left( \min \{1, \tilde{F}\} \right)$, then when $\tau^*(\lambda) \leq \hat{\tau}$, one can verify that

$$(s - v) < 0, \ \eta(\lambda - 1) \left[ 2 \frac{\partial B}{\partial F} \frac{\partial B}{\partial \lambda} - B \frac{\partial^2 B}{\partial F \partial \lambda} \right] - B \frac{\partial B}{\partial F} < 0, \ C > 0,$$

so we must have

$$
\frac{\partial^2 \bar{p}^*}{\partial F \partial \lambda} < 0.
$$
Taking derivative from (9) with respect to \( \lambda \) gives
\[
\frac{\partial \varphi^*}{\partial \lambda} = \frac{s\phi^* F(\tau^*) - [1 - \phi^*(2 - F(\tau^*))]v}{[1 + \eta][\lambda - \phi^*(\lambda - 1)(1 - F(\tau^*))]}(1 + \eta).
\]
(28)

Therefore, the sign of \( \frac{\partial \varphi^*}{\partial \lambda} \) is determined by the sign of \( s\phi^* F(\tau^*) - [1 - \phi^*(2 - F(\tau^*))]v \). When \( \tau^*(\lambda) \leq \hat{\tau} \), \( s\phi^* F(\tau^*) - [1 - \phi^*(2 - F(\tau^*))]v > 0 \), so \( \frac{\partial \varphi^*}{\partial \lambda} > 0 \). In this case, by (25), we know \( \partial^2 \Pi^*(\tau^*, \lambda) / \partial \tau^* \partial \lambda < 0 \), and by (24), we know that \( \partial \tau^* / \partial \lambda < 0 \).

The net effect of consumers’ loss aversion on the firm’s regular price is
\[
\frac{d\varphi^*}{d\lambda} = \frac{\partial \varphi^*}{\partial \lambda} + \frac{\partial \varphi^*}{\partial F(\tau^*)} \frac{\partial \tau^*}{\partial \lambda}.
\]
(29)

The sign of \( \frac{\partial \varphi^*}{\partial F} \) is taken from the proof of Proposition 1 and \( \partial F(\tau^*) / \partial \tau^* = f(\tau^*) > 0 \). According to the results above, when \( \tau^*(\lambda) < \hat{\tau} \), \( \tau^*(\lambda) \) is decreasing in \( \lambda \) (i.e., \( \partial \tau^* / \partial \lambda < 0 \)), and so is \( F(\tau^*) \). By Proposition 2, we must have \( s\phi^* F(\tau^*) - [1 - \phi^*(2 - F(\tau^*))]v > 0 \), so \( \frac{\partial \varphi^*}{\partial \lambda} > 0 \). Plugging these results into (29), we must have \( \frac{d\varphi^*}{d\lambda} > 0 \). □

**Proof of Proposition 5.** First, we show that for any local extreme point \( \lambda^* \) of the profit function \( \Pi^*(\tau^*, \lambda) \), if such \( \lambda^* \) exists, the optimal sales threshold \( \tau^* \), as a function of \( \lambda \), has the property \( \partial \tau^* / \partial \lambda |_{\lambda^*} < 0 \). That is, \( \tau^* \) is decreasing at \( \lambda = \lambda^* \).

Invoking the Envelope Theorem at \( \tau^* \), we have
\[
\frac{d\Pi^*(\tau^*, \lambda)}{d\lambda} = \frac{\partial \Pi^*(\tau^*, \lambda)}{\partial \lambda}.
\]

Taking derivative from (14) with respect to \( \lambda \), we have
\[
\frac{d\Pi^*(\tau^*, \lambda)}{d\lambda} = \frac{\partial \Pi^*(\tau^*, \lambda)}{\partial \lambda} = \frac{\partial \varphi^*}{\partial \lambda} \int_{\tau^*}^{\infty} \min\{x, q\} dF(x),
\]
(30)

where \( \partial \varphi^*(\tau^*, \lambda) / \partial \lambda \) is given by (28).

Since \( \lambda^* \) is a local extreme point, we must have \( \partial \Pi^*(\tau^*, \lambda) / \partial \lambda |_{\lambda^*} = 0 \), and because \( \tau^* < q \), according to (30), that is equivalent to \( \partial \varphi^* / \partial \lambda |_{\lambda^*} = 0 \). By (28), this holds, if and only if, \( \lambda^* \) satisfies
\[
F(\tau^*(\lambda^*)) = \frac{2\phi^* - 1}{\phi^*} \frac{v}{v - s} = \hat{F},
\]
or equivalently,
\[
\tau^*(\lambda^*) = \hat{\tau}.
\]

By (30) and (28), for any other value \( \lambda \neq \lambda^* \), \( \Pi^*(\tau^*, \lambda) \) is increasing in \( \lambda \) if \( \tau^*(\lambda) < \hat{\tau} \) and decreasing if \( \tau^*(\lambda) > \hat{\tau} \).
According to the proof of Proposition 4, \( \partial \tau^*/\partial \lambda \) is expressed as (24), and its sign is determined by the sign of \( \partial^2 \Pi^*(\tau^*, \lambda) / \partial \tau^* \partial \lambda \), as given in (25). Then at \( \lambda^* \), we have \( C(F(\tau^*), \lambda^*) = 0 \), and according to (27), we further have

\[
\frac{\partial^2 \Pi^*}{\partial F \partial \lambda} \bigg|_{\lambda^*} = \frac{\eta(1 + \eta)}{B^2} \phi^*(s - v) < 0.
\]

Since \( \partial \Pi^*/\partial \lambda |_{\lambda^*} = 0 \), the second term in (25) is zero at \( \lambda = \lambda^* \), and \( \partial^2 \Pi^*(\tau^*, \lambda) / \partial \tau^* \partial \lambda |_{\lambda^*} < 0 \), which implies that

\[
\frac{d \tau^*}{d \lambda} \bigg|_{\lambda^*} < 0. \tag{31}
\]

Equation (31) indicates that the sales threshold \( \tau^* \), as a function of \( \lambda \), can only cross the horizontal line \( \hat{\tau} \) at most once from above. Therefore, three possibilities can emerge:

- **Case 1:** \( \tau^*(\lambda) > \hat{\tau} \) for all \( \lambda > 1 \), in which case \( \Pi^*(\tau^*, \lambda) \) is monotonically decreasing in \( \lambda \);
- **Case 2:** \( \tau^*(\lambda) < \hat{\tau} \) for all \( \lambda > 1 \), in which case \( \Pi^*(\tau^*, \lambda) \) is monotonically increasing in \( \lambda \);
- **Case 3:** \( \tau^*(1) > \hat{\tau} \) and \( \tau^*(\infty) < \hat{\tau} \), and \( \tau^*(\lambda) \) crosses \( \hat{\tau} \) at a single point \( \lambda^* \), so that \( \Pi^*(\tau^*, \lambda) \) is monotonically decreasing in \( \lambda \in [1, \lambda^*] \) and monotonically increasing in \( \lambda \in [\lambda^*, \infty] \). □

**Proof of Proposition 6.** By the Implicit Function Theorem, we have

\[
\frac{dq^*}{d\lambda} = -\frac{\partial^2 \Pi^*}{\partial q^* \partial \lambda} \frac{1}{\partial^2 \Pi^* / \partial q^* \partial \lambda}. \tag{32}
\]

Since \( q^* \) is a maximizer of profit function, we must have \( \partial^2 \Pi^*/\partial q^* \partial \lambda < 0 \), and therefore the sign of \( \partial q^*/\partial \lambda \) is the same as \( \partial^2 \Pi^*/\partial q^* \partial \lambda \). By (15), we can explicitly write out \( \partial^2 \Pi^*/\partial q^* \partial \lambda \) as

\[
\frac{\partial^2 \Pi^*}{\partial q^* \partial \lambda} = \frac{\partial \Pi^*}{\partial \lambda} (1 - F(q^*)) + \frac{\partial \Pi^*}{\partial \phi^* \partial \lambda} \frac{\partial \phi^*}{\partial q^*} \int_{\tau}^\infty \min\{x, q^*\} dF(x). \tag{33}
\]

where the expression of \( \partial \Pi^*/\partial \lambda \) is given in (28).

Taking derivative from (28) with respect to \( \phi^* \) gives

\[
\frac{\partial^2 \Pi^*}{\partial \phi^* \partial \lambda} = sF(\tau^*) + (2 - F(\tau^*))v + \eta \lambda sF(\tau^*) + \eta \lambda v (2 - F(\tau^*)) - \eta v (\lambda - 1)(1 - F(\tau^*)) - \eta (1 + \eta) > 0.
\]

By (28), when \( s\phi^* F(\tau^*) - [1 - \phi^* (2 - F(\tau^*))] v > 0 \), i.e., the condition under which \( \Pi^* > v \), we have \( \partial \Pi^*/\partial \lambda > 0 \), so \( dq^*/d\lambda > 0 \), i.e., the optimal order quantity is increasing in consumers’ loss-averseness. As the order quantity increases, according to Proposition 3, the full price \( \overline{p} \) increases in the order quantity and is greater than \( v \), so the inequality \( s\phi^* F(\tau^*) - [1 - \phi^* (2 - F(\tau^*))] v > 0 \) still holds, by which we must have \( \partial \overline{p}^*/\partial \lambda > 0 \). Hence, both terms in the right hand side of (33) are positive. Together with (32), we can further conclude \( dq^*/d\lambda > 0 \).
In summary, once $\lambda$ is such that $dq^*/d\lambda > 0$, the sign of $dq^*/d\lambda$ will not become negative as $\lambda$ increases. A possible cut-off point $\hat{\lambda}$ immediately follows. □

Proof of Proposition 7. To find the net effect $dp^*/dc$, we use the chain rule and get

$$\frac{dp^*}{dc} = \frac{\partial p^*}{\partial \phi^*} \frac{\partial \phi^*}{dq^*} \frac{dq^*}{dc} + \frac{\partial p^*}{\partial F} \frac{\partial F}{\partial \tau^*} \frac{\partial \tau^*}{dq^*} \frac{dq^*}{dc}.$$  

These signs hold because from Proposition 3, we have $\partial p^*/\partial \phi^* > 0$ and $\partial \tau^*/\partial q^* < 0$, and from (23), we have $\partial p^*/\partial F < 0$ and $\partial F(\tau^*)/\partial \tau^* = f(\tau^*) > 0$, and because the fill rate is increasing in the order quantity, we have $\partial \phi^*/\partial q^* > 0$ holds because the more order quantity, the higher the optimal fill rate. Therefore, the sign of $dp^*/dc$ is the same as the sign of $\partial q^*/dc$.

To determine the sign of $dq^*/dc$, we note that $\partial \Pi^*/\partial q^* = 0$, and by the Implicit Function Theorem:

$$\frac{dq^*}{dc} = -\frac{\partial^2 \Pi^*}{\partial q^* \partial c} \frac{\partial \Pi^*}{\partial q^*}.$$  

It can be verified that $\partial^2 \Pi^*/\partial q^* \partial c = -1$, and we have $\partial^2 \Pi^*/\partial q^*^2 \leq 0$ because $q^*$ is a maximizer of the profit function. Therefore, we have

$$\frac{dq^*}{dc} \leq 0.$$  

Furthermore, we have

$$\frac{dF(\tau^*)}{dc} = \frac{\partial F(\tau^*)}{\partial \tau^*} \frac{\partial \tau^*}{dq^*} \frac{dq^*}{dc} > 0.$$  

□

References


Cachon, G., P. Feldman. 2013. Price commitments with strategic consumers: Why it can be optimal to discount more frequently... than optimal. Working paper, UC Berkeley, CA.


Özer, Ö., Y. Zheng. 2013. Markdown or everyday-low-price? The role of consumer regret and availability misperception Working paper, UT Dallas and MIT.


