Dynamic Stochastic Inventory Management with Reference Price Effects

Xin Chen
Department of Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, xinchen@illinois.edu

Peng Hu
School of Management, Huazhong University of Science and Technology, Wuhan, China 430074, hu_peng@hust.edu.cn

Stephen Shum
College of Business, City University of Hong Kong, Kowloon, Hong Kong, swhshum@cityu.edu.hk

Yuhan Zhang
Morgan Stanley, 1585 Broadway 3rd floor New York City NY 10036, yuhan.zhang@morganstanley.com

We analyze the joint inventory and pricing decisions of a firm when demand depends on not only the current selling price but also a memory-based reference price and customers are loss averse. The presence of reference price effect leads to a non-concave one-period expected revenue in price and reference price. We introduce a transformation technique that allows us to prove under some mild assumptions the optimality of a reference-price-dependent base-stock list-price policy, which is characterized by a base-stock level and a target reference price. In addition, we show that the target reference price is increasing in the reference price but except in the loss neutral case the base-stock level is not monotone in the reference price. We also show that in the steady state of the model with the reference price effect the optimal price is lower while the optimal base-stock level is higher than their counterparts in the model without the reference price effect.

1. Introduction

It has been well established in the literature that integrating inventory and pricing decisions can lead to significant profit improvement for a variety of industries (see, e.g., Federgruen and Heching 1999 and Chen and Simchi-Levi 2012). Existing works in that literature typically assume that demand depends on the firm's pricing strategy only through the current selling price, and adopt the standard approach to inductively show that the dynamic programming problem is a concave and/or supermodular maximization problem. This leads to the optimality of the *base-stock list-price* policy: if the initial inventory level is below the base-stock level, then it is increased to the base-stock level and a list price is charged; otherwise nothing is ordered and a price discount is offered, and the higher the initial inventory level the deeper the discount offered (Federgruen and Heching 1999). Nevertheless, recent studies on behavioral studies have recognized that customers are subject to anchoring effects (see, e.g., Kalyanaram and Winer 1995). In particular, customers often develop their own ideas of a "fair price", also referred to as the *reference price*, after observing past prices of the product. If the current selling price is lower than the reference price, customers see it as a gain and hence are more likely to make the purchase. Otherwise, they see it as a loss and would be less inclined to make the purchase. This phenomenon is usually referred to as the *reference price effect*. Customers are called *loss averse* (*loss neutral*) if demand is more responsive (equally responsive) to customers' perceived losses than (as) their perceived gains. Some recent works explore how pricing strategies should account for the reference price effect in the absence of inventory consideration (Greenleaf 1995, Kopalle et al. 1996, Fibich et al. 2003, Popescu and Wu 2007, Nasiry and Popescu 2011, Hu et al. 2016 and Chen et al. 2016).

To our best knowledge, only a few papers have integrated the reference price effect into stochastic inventory models. For example, Urban (2008) analyzes a one-period setting, and Gimpl-Heersink (2008) focuses on a two-period settings and develops theoretical results mainly for the loss neutral case. There are significant technical challenges in extending the analysis to the multi-period setting with loss neutral customers and the more general setting with loss-averse customers. First, the one-period expected revenue is not jointly concave in the state and decision variables (the reference price and the target reference price) in general. Second, the feasible set does not form a lattice in decision variables and system states. As a result, the associated dynamic programming problem is not a concave or supermodular maximization problem, which leaves many questions regarding a firm's optimal decisions in the presence of reference price effect and uncertainty in a dynamic setting unanswered. In particular, does the base-stock list-price policy remain optimal? How does the reference price affect the optimal decisions, and what is the impact of the reference price effect in the long run?

To tackle these challenges, we introduce a transformation technique to generate a modified one-period expected revenue function that is concave. This allows us to prove the optimality of a reference-price-dependent base-stock policy. In addition, our transformation technique coupled with a recently developed result on preserving supermodularity with non-lattice structures (Chen et al. 2013) allows us to prove the optimality of a list-price policy. We also study how the optimal base-stock level and target reference price depend on the current reference price, and compare long-run optimal price and base-stock level to those in the case without the reference price effect.

The rest of this paper is organized as follows. We present the model in Section 2, characterize the optimal policy in Section 3 and conclude in Section 4. All the proofs are relegated to the appendix.

2. The Model

Consider a firm who makes inventory and pricing decisions at the beginning of each period over a planning horizon with T periods. The order is received immediately and incurs a per unit cost c. The price is restricted to some interval $[\underline{p}, \overline{p}]$, where we assume $\underline{p} \geq c$ to ensure the marginal profit to be non-negative. In each period t, the expected demand d_t depends on the price p_t and the reference price r_t, and is given by

$$d_t = (B - Ap_t) + \eta(r_t - p_t),$$

where $B > 0$ represents the market size, $A > 0$ measures the sensitivity of demand with respect to the selling price, and the difference $r_t - p_t$ denotes customers' perceived gain when $r_t > p_t$ and loss when $r_t < p_t$. Moreover, the reference price effect is given by

$$\eta(z) = \eta^+ \max\{0, z\} + \eta^- \min\{0, z\},$$

where the non-negative parameters η^- and η^+ measure the sensitivities of demand associated with the perceived loss and gain, respectively. For simplicity of presentation, we assume that the market size B and customers' response to price and reference price (i.e., A and $\eta^\pm$) are time independent. Our demand model is similar to the ones in Greenleaf (1995) and Nasiry and Popescu (2011). One advantage of this model is that it is much easier to calibrate when compared with more complex demand models. In addition, different parameters in the model can be easily understood by managers and practitioners. Customers are usually called *loss averse* if $\eta^- \geq \eta^+$, *gain seeking* if $\eta^- \leq \eta^+$ and *loss neutral* if $\eta^- = \eta^+$. Prospect theory (Kahneman and Tversky 1979) postulates that loss aversion behavior is more common than gain-seeking behavior at individual customer level and hence we focus on the loss-averse case.

The reference price in a period is formed based on the prices observed by customers in all previous periods. We adopt the exponentially smoothed adaptive expectations process (see, e.g., Mazumdar et al. 2005) in which the reference price of the next period is a weighted average of the price and the reference price in the current period. Formally, given price p_t and reference price r_t in period t, the reference price in period $t+1$ evolves according to

$$r_{t+1} = (1-\alpha)p_t + \alpha r_t,$$

where $0 \leq \alpha < 1$ is called the *memory factor*. This reference price evolution model is commonly used both in empirical studies (e.g., Kalyanaram and Little 1994, Greenleaf 1995) and in analytical models (e.g., Popescu and Wu 2007). To avoid the trivial case where past prices have no impact on demand, we assume that $\alpha < 1$. The initial reference price is given by $r_1 \in [\underline{p}, \overline{p}]$, and hence all r_t belong to the interval. Furthermore, we make the following assumption.

Assumption 1. *$\eta^- - \eta^+ \leq 2A(1-\alpha)$.*

Assumption 1 is not restrictive. For example, it is satisfied when customers are loss neutral ($\eta^- = \eta^+$), or when consumers have short memories (α is small) and the direct price effect dominates the reference price effect ($A \geq \eta^-$). Such assumption is consistent with the results from many empirical studies, e.g., Hardie et al. (1993).

To facilitate discussion, we denote r and q as the *current reference price* (r_t) and *target reference price* (r_{t+1}) respectively, and we express price p_t and expected demand d_t as functions of these reference prices as follows:

$$p(r,q) = \frac{q-\alpha r}{1-\alpha} \quad \text{and} \quad d(r,q) = B - A\,\frac{q-\alpha r}{1-\alpha} + \frac{\eta(r-q)}{1-\alpha}.$$

The realized demand in each period t is stochastic and follows the additive model $D_t = d(r,q) + \varepsilon_t$, where ε_t is a random variable with zero mean and independent across time. We assume that D_t is non-negative for any $r, q \in [\underline{p}, \overline{p}]$ in each period t. The realized demand is satisfied by on-hand inventory, where unsatisfied demand is fully backlogged and any excess inventory is carried over to the next period. In addition, let $h^0(I)$ be the associated inventory holding cost when the leftover inventory $I > 0$ and backlogging cost when $I < 0$.

The firm's objective is to find a joint inventory and pricing policy to maximize the total expected discounted profit over the planning horizon. Given the initial inventory level x and reference price r in each period $t = 1, 2, \cdots, T$, the profit-to-go function $v_t^0(x,r)$ at the beginning of period t satisfies the following dynamic programming recursion, denoted as problem (1):

$$v_t^0(x,r) = \underset{y,q}{\text{maximize}} \;\; p(r,q)d(r,q) - c[y + d(r,q) - x] - \mathbb{E}h^0(y-\varepsilon_t) + \gamma \mathbb{E}v_{t+1}^0(y-\varepsilon_t, q), \tag{1a}$$

$$\text{subject to} \;\; d(r,q) + y \geq x, \; p(r,q) \in [\underline{p}, \overline{p}], \tag{1b}$$

where $0 \leq \gamma \leq 1$ is the discount factor, and decision variables y and q denote the expected leftover inventory and the target reference price, respectively. Moreover, the terminal value $v_{T+1}^0(x,r) = cx$, that is, any backlogged demand is satisfied and any leftover inventory incurs reimbursement with the per unit cost/value c at the end of the planning horizon. Similar to Federgruen and Heching (1999), we assume that the function $h(I) = h^0(I) + (1-\gamma)cI$ is convex, $\lim_{|I|\to\infty} h(I) = \infty$, and $\mathbb{E}h(y-\varepsilon_t)$ is well-defined for any y. Under these assumptions, problem (1) admits some finite optimal solution, which we denote as $[y_t(x,r), q_t(x,r)]$. Note that the optimal order-up-to inventory level $s_t(x,r) = y_t(x,r) + d(r, q_t(x,r))$ and the optimal price $p_t(x,r) = p(r, q_t(x,r))$. Throughout this paper, when multiple optimal solutions exist for a problem, we always select the one which gives the lexicographically smallest $[q_t(x,r), -y_t(x,r)]$ for convenience, leading to the lexicographically smallest $[q_t(x,r), -s_t(x,r)]$ among the optimal solutions.

When there is no reference price effect, problem (1) reduces to the model studied by Federgruen and Heching (1999), where the one-period expected revenue, expressed as $pd(p)$, is typically assumed to be a concave function. The authors, following the standard approach, inductively show that the dynamic programming problem is a concave and supermodular maximization problem, and prove that a base-stock list-price policy is optimal. Their method is, however, not applicable in our problem because the term $p(r,q)d(r,q)$ in the objective function (1a) is not jointly concave and the feasible set given in (1b) is not a lattice in all decision variables and parameters. This is why Gimpl-Heersink (2008) cannot extend her theoretical results to multi-period models even for the loss neutral case.

3. Main results

To circumvent the challenge resulted from the lack of joint concavity, we introduce a transformation technique. In particular, define the transformed profit-to-go functions below using a positive λ (see (3) in the appendix for its explicit expression) for all $t = 1, \cdots, T+1$:

$$v_t(x,r) = v_t^0(x,r) - cx - \lambda r^2.$$

From (1) we know that $v_t(x,r)$ satisfies the dynamic programming recursion

$$v_t(x,r) = \underset{y,q}{\text{maximize}} \quad \pi(r,q) - \mathbb{E}h(y-\varepsilon_t) + \gamma\mathbb{E}v_{t+1}(y-\varepsilon_t, q), \tag{2a}$$

$$\text{subject to} \quad d(r,q) + y \geq x, \ p(r,q) \in [\underline{p}, \overline{p}], \tag{2b}$$

where π given below denotes the transformed one-period expected revenue.

$$\pi(r,q) = [p(r,q) - c]d(r,q) - \lambda r^2 + \gamma\lambda q^2.$$

Observe that the optimal solution to problem (1) also solves problem (2), and vice versa. In addition, while the original one-period expected revenue function $\pi^0(r,q) = [p(r,q) - c]d(r,q)$ is not jointly concave in general, it is possible to select an appropriate parameter λ such that the transformed expected revenue $\pi(r,q)$ is jointly concave.

Proposition 1 **(Concavity of Transformed Problem)**. *There exists a positive λ such that $\pi(r,q)$ is jointly concave when $p(r,q) \in [\underline{p}, \overline{p}]$.*

We now provide some intuition on the transformation technique. For the loss neutral case, though the original one-period expected revenue function $\pi^0(r,q)$ is quadratic with non-positive coefficients for r^2 and q^2, it is not jointly concave in r and q because the determinant of its Hessian is negative. By subtracting λr^2 from the profit-to-go function in each period, we make the coefficient of r^2 more negative yet that of q^2 less negative in the transformed one-period expected revenue function $\pi(r,q)$. Proposition 1 implies that it is always possible to find an appropriate λ which bends the curvature of the function π^0 to the extent that yields joint concavity in π. It also implies this observation carries over to the loss-averse case as long as it is not too far away from the loss neutral case, i.e., under Assumption 1.

With the help of Proposition 1, we are ready to characterize $v_t(x,r)$ and the optimal ordering policy. Consider the problem below by relaxing the constraint $d(r,q)+y \geq x$ from (2b):

$$\underset{y,q}{\text{maximize}} \;\left\{\pi(r,q) - \mathbb{E}h(y-\varepsilon_t) + \gamma\mathbb{E}v_{t+1}(y-\varepsilon_t, q) : p(r,q) \in [\underline{p}, \overline{p}]\right\}.$$

Let $[y_t(r), q_t(r)]$ be an optimal solution of the above problem, and $s_t(r) = y_t(r) + d(r, q_t(r))$.

Theorem 1 **(Optimality of Base-stock Policy)**. *$v_t^0(x,r)$ is increasing in r, and $v_t(x,r)$ is decreasing in x and jointly concave in (x,r). Moreover, a base-stock policy with the reference-price-dependent base-stock level $s_t(r)$ is optimal, and $q_t(x,r) = q_t(r)$ and $p_t(x,r) = \frac{q_t(r)-\alpha r}{1-\alpha}$ if $x \leq s_t(r)$.*

Theorem 1 suggests that there is a reference-price-dependent base-stock level $s_t(r)$ such that an order is placed to raise the inventory up to $s_t(r)$ if $x < s_t(r)$, and no order is placed otherwise. Furthermore, if $x \leq s_t(r)$, then we should set the selling price to reach a target reference price $q_t(r)$ independent of x.

Next we study how the target reference price $q_t(x,r)$ depends on the initial inventory level x. In particular, we are interested to know whether a list-price policy is still optimal, that is, whether the optimal price $p_t(x,r) = p(r, q_t(x,r))$ or equivalently $q_t(x,r)$ is decreasing in x. A widely used approach for answering such question is to show that the objective function is supermodular and the feasible set forms a lattice in decision variables and parameters, hence then the optimal solution is increasing in the parameters. Unfortunately, the requirement on the feasible set is not satisfied in our problem. Interestingly, the concavity of the profit-to-go functions enabled by our transformation technique allows us to apply results developed by Chen et al. (2013) on preserving concavity and supermodularity in a class of parametric optimization problems with non-lattice structures and show that a list-price policy is optimal.

Theorem 2 **(Optimality of List-price policy)**. *$q_t(x,r)$ and $p_t(x,r)$ are decreasing in x.*

We now move to study how the base-stock level $s_t(r)$, the target reference price $q_t(r)$ and the list price $p_t(r) = p(r, q_t(r))$ depend on the current reference price r. It should be noticed that the transformation technique is also crucial in the proof of the following theorem.

Theorem 3 **(Impact of Reference Price)**. *$q_t(r)$ is increasing in r, and there exist $Q_t^- \leq Q_t^+$ such that $q_t(r) > r$ and $p_t(r) > r$ if $r < Q_t^-$, $q_t(r) = p_t(r) = r$ if $Q_t^- \leq r \leq Q_t^+$, and $q_t(r) < r$ and $p_t(r) < r$ if $r > Q_t^+$. Moreover, $s_t(r)$ is increasing in r when customers are loss neutral, but it does not exhibit any specific monotonic pattern in general.*

Theorem 3 shows two results. First, the target reference price is increasing in the initial reference price, and there exist thresholds $Q_t^- \le Q_t^+$ such that the optimal price $p_t(r)$ is higher (or lower) than the current reference price if the current reference price is lower (or higher) than Q_t^- (or Q_t^+). However, it remains an open problem whether $p_t(r)$ is increasing in r when $r \notin [Q_t^-, Q_t^+]$. In the case without demand uncertainty and inventory decision, Popescu and Wu (2007) identify three sufficient conditions on the demand model for the monotonicity of $p_t(r)$. However, in our setting, except for the special case with $\alpha = 0$ (i.e., $p_t(r) = q_t(r)$), only the linear loss neutral demand model satisfies one of their other two conditions, and even in this case, their analysis cannot be generalized.

Second, while a higher current reference price leads to a higher base-stock level when customers are loss neutral, the base-stock level is not monotone in the reference price when customers are loss averse. The reason is that, the optimal price is equal to the current reference price when the current price lies in the region $[Q_t^-, Q_t^+]$. At this optimal price, consumers perceive no gain or loss and the impact of reference price on demand is reduced to zero. Thus, when the current reference price (which is also equal to the optimal price) increases, the demand decreases and correspondingly the base-stock level decreases. This does not happen when customers are loss neutral since the region $[Q_t^-, Q_t^+]$ reduces to a singleton in this case. In particular, in the loss neutral case, the optimal price is equal to the current reference price only at this singleton, and the reference price appears in the demand function linearly. Thus, when the current reference price increases, demand still increases if the increase of the selling price (if at all) is not too large, leading to a higher optimal base-stock level.

To further study the impact of reference price effect, we compare two models, one without and the other one with the reference price effect, in an infinite horizon setting where demand uncertainty ε_t are assumed to be identically distributed and the discount factor $\gamma < 1$. For technical reasons, we assume the probability that realized demand $d(r,q) + \varepsilon_t$ is zero is less than 1 for any $r, q \in [\underline{p}, \overline{p}]$. Notice that in the model without the reference price effect, the firm, starting with a low inventory level, should order up to a base-stock level s^* and charge a list price p^* (see Federgruen and Heching 1999). In the model with the reference price effect, we have the following theorem.

Theorem 4 **(Impact of Reference Price Effect)**. *Reference prices r_t under the optimal policy converge to some steady state reference price with probability 1 as t goes to infinity, where all steady state reference prices form an interval $[R^-, R^+]$. Moreover, for any $p^\infty \in [R^-, R^+]$, $p^\infty \le p^*$ and the associated base-stock level $s(p^\infty) \ge s^*$.*

Theorem 4 shows that, the existence of reference price effect leads to a lower price in the long run, which is consistent with the result and follows the same intuition in Popescu and Wu (2007) dealing with the case without demand uncertainty, and consequently, the associated base-stock level $s(p^\infty) \geq s^*$. It is appropriate to point out that the steady state price p^∞ is actually the same as that in the case without demand uncertainty. This is because the uncertainty involved in the additive demand is independent of the price and the reference price, and can be fully subsumed by inventory decisions. Yet, unlike the case without demand uncertainty, the price trajectory in our model is not necessarily monotone due to the existence of demand uncertainty. In fact, the system in our model evolves exactly the same as the one in the model without uncertainty only after the inventory level drops to below the base-stock level.

4. Concluding Remarks

We study a joint inventory and pricing model taking into account the reference price effect. Despite the technical challenges resulted from the increase in the dimension of the dynamic program (due to the reference price effect), the non-smooth demand function (due to the kink function η in our model) and the lack of properties commonly used in the literature (e.g., the joint concavity and lattice structure) to characterize optimal policies, we develop a transformation technique and manage to prove that a base-stock list-price policy is optimal under a mild condition. We study how the optimal policy, characterized by the base-stock level and the target reference price, changes with the reference price. We also show that the existence of reference price effect leads to a lower price and higher base-stock level in the long run.

It is noteworthy to mention that Assumption 1 can be relaxed for some of our results. For instance, for Proposition 1, the condition specified by inequality (4) is sufficient. Thus, Theorem 1 and Theorem 4, whose proofs follow from the joint concavity of $\pi(r,q)$, hold under the less restrictive condition (4). Similarly, as can be seen from the proof, Theorem 3 holds under a weaker condition specified by inequality (8). Most of our results can be extended, by following a similar approach, to cases with non-stationary parameters under suitable conditions.

Furthermore, under some conditions, it is possible to extend our transformation technique to more general demand models. For instance, consider the piece-wise linear demand $d = \min_n\{B_n - A_n p + \eta_n(r-p) : 1 \leq n \leq N\}$, which may provide good approximations to a general demand model that is concave in p and r for a sufficiently large N. If Assumption 1 holds for $\eta^- = \max_n \eta_n$, $\eta^+ = \min_n \eta_n$ and $A = \min_n A_n$, then it can be verified that Theorems 1 and 2 remain valid. Of course, in more general settings, functions other than λr^2 may be needed in the transformation. We should also point out the transformation technique can be extended to the case with multiplicative demand uncertainty (see Hu 2011 for details).

This paper should only be taken as an initial attempt at studying inventory and pricing models with reference effects. Two extensions are particularly noteworthy. First, our discussion focuses on a single product, hence an interesting extension is to study the optimal pricing and inventory policy in a multi-product setting. It is very likely that concavity is also absent in this problem, and we expect that an extension of our transformation technique will still be useful. Of course, this problem is significantly more challenging, due to the large dimension of the state space and the dependence of a product on the reference price and inventory levels of other products.

Second, besides the impact of the reference price, other types of reference effects also affect a firm's decisions (see, e.g., Yang et al. 2014a, Yang et al. 2014b and the references therein). For instance, a reference fill-rate may have a significant impact on inventory and pricing decisions. This line of inquiry is pursued by Liu and van Ryzin (2011) for the case of no demand uncertainty. When there is demand uncertainty, the investigation becomes challenging because the reference fill-rate is a kinked function of demand and the inventory level and properties such as concavity and supermodularity are likely to be absent. While this problem deserves a separate study, the techniques used in this paper may be helpful. In particular, our transformation technique can be potentially applied to general dynamic programming problems of the form below:

$$v_t^0(x) = \underset{y}{\text{maximize}} \; \{\mathbb{E}\pi_t^0(x,y,\varepsilon_t) + \gamma \mathbb{E} v_{t+1}^0(\varphi_t(x,y,\varepsilon_t)) : y \in \mathcal{S}_t(x)\},$$

where γ denotes the discounted factor, $v_t^0(x)$ is the profit-to-go associated with the state x observed at the beginning of period t, $\pi_t^0(x,y,\varepsilon_t)$ represents the one-period profit associated with state x, decision y and uncertainty term ε_t, the system dynamic is given by $x_{t+1} = \varphi_t(x,y,\varepsilon_t)$, and $\mathcal{S}_t(x)$ denotes the feasible set of decision y associated with state x. Observe that given a function sequence $\{\lambda_t(x)\}$, we can define the transformed profit-to-go function $v_t(x) = v_t^0(x) + \lambda_t(x)$, which satisfies the dynamic programming recursion

$$v_t(x) = \underset{y}{\text{maximize}} \; \{\pi_t(x,y) + \gamma \mathbb{E} v_{t+1}(\varphi_t(x,y,\varepsilon_t)) : y \in \mathcal{S}_t(x)\},$$

where $\pi_t(x,y) = \mathbb{E}[\pi^0(x,y,\varepsilon_t) + \lambda_t(x) - \gamma\lambda_{t+1}(\varphi_t(x,y,\varepsilon_t))]$. The two problems above are equivalent in the sense that an optimal solution to one problem also solves the other. Moreover, the sequence $\{\lambda_t(x)\}$ provides flexibility to "reshape" the original one-period profit function π_t^0 to the transformed one π_t. Thus, depending on the expressions of π_t^0 and φ_t, it is possible to construct an equivalent dynamic programming problem with additional (hidden) structural properties in its objective function by appropriately selecting $\{\lambda_t(x)\}$.

Acknowledgments
We thank Zhenyu Hu, Zhan Pang, the Area Editor Costis Maglaras, one anonymous associate editor and two anonymous referees for their valuable comments. The authors also acknowledge the second author, Peng Hu, as the corresponding author of this paper. This research is partly supported by NSF Grants CMMI-0926845 ARRA, CMMI-1030923, CMMI-1363261 and CMMI-1538451, China NSFC Grants 71228203, 71520107001 and 71201066, and Hong Kong GRF Grant 643310.

Appendix: Proofs

Proof of Proposition 1

Define constants $m^{\pm} = \eta^{\pm}/[A(1+\alpha)]$ and

$$\lambda = \left\{[A(1+\alpha)+\eta^-] - \sqrt{[A(1+\alpha)]^2 + 2A(1+\alpha)\eta^-}\right\}/[2(1-\alpha)]. \tag{3}$$

We first prove that $\pi(r,q)$ is jointly concave in (r,q) for $p(r,q) \geq \underline{p}$ if the following inequality holds:

$$m^- - m^+ \leq \sqrt{1+2m^-} + \sqrt{1+2m^+}. \tag{4}$$

To see it, by $p(r,q) - c \geq \underline{p} - c \geq 0$ and $\eta(z) = \min\{\eta^+ z, \eta^- z\}$, we can express

$$\pi(p+z, p+\alpha z) = \min\left\{\pi^+(p,z), \pi^-(p,z)\right\} - (1-\gamma)\lambda(p+\alpha z)^2,$$

where $\pi^{\pm}(p,z) = (p-c)(B - Ap + \eta^{\pm} z) - \lambda(p+z)^2 + \lambda(p+\alpha z)^2$. Because $\pi(r,q)$ is jointly concave in (r,q) if and only if $\pi(p+z, p+\alpha z)$ is jointly concave in (p,z), the minimum of two concave functions is also concave, and $\pi^{\pm}(p,z)$ have the Hessian matrices as follows:

$$\begin{bmatrix} -2A, & \eta^{\pm} - 2\lambda(1-\alpha) \\ \eta^{\pm} - 2\lambda(1-\alpha), & -2\lambda(1-\alpha^2) \end{bmatrix},$$

we only need to prove determinants of these matrices are non-negative, i.e., $4A\lambda(1-\alpha^2) \geq [\eta^{\pm} - 2\lambda(1-\alpha)]^2$. By definition of $m^{\pm}$, it is equivalent to $\Delta^{\pm}(\Lambda) \geq 0$ with $\Delta^{\pm}(x) = 2x - (m^{\pm} - x)^2$ and

$$\Lambda = [2(1-\alpha)\lambda]/[A(1+\alpha)] = 1 + m^- - \sqrt{1+2m^-},$$

where the second equation can be verified from definitions of λ and m^-. By some basic algebra, it is straightforward to see that Λ given above is indeed a root of function $\Delta^-(x)$, and it also lies between the two roots of the concave and quadratic function $\Delta^+(x)$. Thus, we conclude the joint concavity of $\pi(r,p)$ by $\Delta^{\pm}(\Lambda) \geq 0$ provided inequality (4).

It remains to prove inequality (4). In fact, by definitions of $m^{\pm}$, Assumption 1 is equivalent to $m^- - m^+ \leq 2(1-\alpha)/(1+\alpha)$. It leads to $m^- - m^+ \leq 2$ and hence inequality (4) immediately. □

Proof of Theorem 1

Notice that $v^0_{t+1}(x,r)$ is increasing in r at $t = T$. Suppose it is true for some $t \leq T$. Let $d^0(p,r) = B - Ap + \eta(r-p)$ and reformulate $v^0_t(x,r)$ given by problem (1) as below:

$$\begin{aligned} &\underset{y,p}{\text{maximize}} \quad (p-c)d^0(p,r) - c(y-x) + \mathbb{E}h^0(y-\varepsilon_t) + \gamma\mathbb{E}v^0_{t+1}(y-\varepsilon_t, \alpha r + (1-\alpha)p), \\ &\text{subject to} \quad d^0(p,r) + y \geq x, \; p \in [\underline{p}, \overline{p}]. \end{aligned}$$

Because $\underline{p} \geq c$ and $d_t^0(p,r)$ is increasing in r, its objective increases and its feasible set expands as r increases, implying $v_t^0(x,r)$ is increasing in r. The monotonicity of $v_t(x,r)$ in x is straightforward because for problem (2), its objective is independent of x and its feasible set shrinks as x increases.

For joint concavity of v_t for any $1 \leq t \leq T$, suppose it is true in period $t+1$, which is trivial for $t = T$ by $v_{T+1}(x,r) = -\lambda r^2$. In period t, by Proposition 1, convexity of h and the inductive assumption, the objective function in (2a) is jointly concave in (r,q,y). In addition, because $d(r,q)$ is concave and $p(r,q)$ is linear, the feasible set in (2b) is convex in (x,r,q,y). Thus, problem (2) is a concave maximization problem, implying $v_t(x,r)$ is also jointly concave.

Because problem (1) has the same optimal solutions as the concave maximization problem (2), it immediately follows the optimality of the base-stock policy with the base-stock level $s_t(r)$, as well as the optimality of the specified target reference price $q_t(r)$ and price $p_t(x,r)$ if $x \leq s_t(r)$. □

Proof of Theorem 2

We first show that $\pi(r,q)$ is supermodular in r and q when $p(r,q) \geq \underline{p}$. Because $\eta(z)$ is increasing and concave, it is straightforward to see $(p-c)[B - Ap + \eta(z)]$ is component-wise concave and supermodular in (p,z) when $p \geq c$. By Corollary 2 in Chen et al. (2013), $\pi(r,q)$ is supermodular when $p(r,q) \geq \underline{p} \geq c$.

Let $w_t(y,q) = \gamma v_{t+1}(y,q) - h(y)$ and reformulate problem (2) as below:

$$v_t(x,r) = \underset{s,q}{\text{maximize}} \ \left\{ \pi(r,q) + \mathbb{E} w_t(s - d(r,q) - \varepsilon_t, q) : s \geq x, \ p(r,q) \in [\underline{p}, \overline{p}] \right\}.$$

To see $q_t(x,r)$ is decreasing in x, by Theorem 2.2.8 in Simchi-Levi et al. (2014), we only need to prove the submodularity of $w_t(s - d(r,q), q)$ in (s,q), i.e., for any $S \geq s$ and $Q \geq q$,

$$w_t(S - d(r,Q), Q) - w_t(s - d(r,Q), Q) \leq w_t(S - d(r,q), q) - w_t(s - d(r,q), q). \tag{5}$$

Note that in cases with multiple optimal solutions, Theorem 2.2.8 in Simchi-Levi et al. (2014) is applicable given our choice of the lexicographically smallest $[q_t(x,r), -s_t(x,r)]$.

Observe that by letting $\theta^{\pm} = (A + \eta^{\pm})/(1-\alpha)$ and $\delta^{\pm} = (\alpha A + \eta^{\pm})/(1-\alpha)$, we can express

$$d(r,q) = \min\{(\delta^+ r - \theta^+ q), (\delta^- r - \theta^- q)\} + B. \tag{6}$$

We next show inequality (5) holds if both $w_t(s + \theta^{\pm} q, q)$ are submodular in (s,q). In fact, if either $Q \leq r$ or $q \geq r$, then $d(r,q) = \delta r - \theta q + B$ and $d(r,Q) = \delta r - \theta Q + B$ in inequality (5) for either $(\theta,\delta) = (\theta^+, \delta^+)$ or $(\theta,\delta) = (\theta^-, \delta^-)$. In this case inequality (5) immediately follows from the submodularity of $v_{t+1}(s + \theta^{\pm} q, q)$. Moreover, if $q \leq r \leq Q$, then by what we just proved,

$$\begin{aligned} w_t(S - d(r,Q), Q) - w_t(s - d(r,Q), Q) &\leq w_t(S - d(r,r), r) - w_t(s - d(r,r), r) \\ &\leq w_t(S - d(r,q), q) - w_t(s - d(r,q), q). \end{aligned}$$

That is, inequality (5) also holds in this case. Thus, we only need to verify the submodularity of $w_t(s+\theta^{\pm}q,q) = \gamma v_{t+1}(s+\theta^{\pm}q,q) - h(s+\theta^{\pm}q)$ in (s,q). Furthermore, because $-h(s+\theta^{\pm}q)$ is submodular in (s,q) by convexity of h, it suffices to verify that $v_{t+1}(s+\theta^{\pm}q,q)$ are submodular in (s,q) or equivalently, $v_{t+1}(\theta^{\pm}r-x,r)$ are supermodular in (x,r).

Now we inductively prove the supermodularity of $v_{t+1}(\theta^{+}r-x,r)$ for all $1\le t\le T$. Note that the statement holds for $t=T$ since $v_{T+1}(\theta^{+}r-x,r)=-\lambda r^2$. Suppose $v_{t+1}(\theta^{+}r-x,r)$ is supermodular in (x,r) for some $t\le T$. Then by problem (2), $v_t(\theta^{+}r-x,r)$ is determined by

$$\underset{q,\tilde{y}}{\text{maximize}} \;\; \pi(r,q) - \mathbb{E}h(\theta^{+}q-\tilde{y}-\varepsilon_t) + \gamma\mathbb{E}v_{t+1}(\theta^{+}q-\tilde{y}-\varepsilon_t,q), \tag{7a}$$

$$\text{subject to} \;\; d(r,q)+(\theta^{+}q-\tilde{y}) \ge (\theta^{+}r-x), \quad p(r,q)\in[\underline{p},\overline{p}], \tag{7b}$$

where the objective function in (7a) is jointly concave when $p(r,q)\in[\underline{p},\overline{p}]$ as proved in Theorem 1, and it is also supermodular by supermodularity of π, convexity of h and the inductive assumption. Furthermore, from expression (6), we can express the first constraint in (7b) as

$$\begin{cases}(\theta^{+}-\delta^{+})r+(\theta^{+}-\theta^{+})q+\tilde{y}+s_1=x, & s_1\ge -B,\\ (\theta^{+}-\delta^{-})r+(\theta^{-}-\theta^{+})q+\tilde{y}+s_2=x, & s_2\ge -B,\end{cases}$$

where s_i, $i=1,2$, are slack variables. Hence, by introducing $\tilde{r}=r$, problem (7) is of the form

$$\underset{(\tilde{r},q,\tilde{y},s_1,s_2)\in\mathcal{S}}{\text{maximize}} \;\; \{g(\tilde{r},q,\tilde{y},s_1,s_2): M[\tilde{r},q,\tilde{y},s_1,s_2]' = N[x,r]'\},$$

where g is concave and supermodular on $\mathcal{S}=\{(\tilde{r},q,\tilde{y},s_1,s_2): p(\tilde{r},q)\in[\underline{p},\overline{p}],\ s_i\ge -B\}$, X' denotes the transpose of a matrix X, and matrices M and N are given by

$$M=\begin{bmatrix}1, & 0, & 0,\,0,\,0\\ \theta^{+}-\delta^{+}, & \theta^{+}-\theta^{+}, & 1,\,1,\,0\\ \theta^{+}-\delta^{-}, & \theta^{-}-\theta^{+}, & 1,\,0,\,1\end{bmatrix} = \begin{bmatrix}1, & 0, & 0,\,0,\,0\\ A, & 0, & 1,\,1,\,0\\ A-\frac{\eta^{-}-\eta^{+}}{1-\alpha}, & \frac{\eta^{-}-\eta^{+}}{1-\alpha}, & 1,\,0,\,1\end{bmatrix}, \; N=\begin{bmatrix}0,\,1\\1,\,0\\1,\,0\end{bmatrix}.$$

By Assumption 1, it can be verified that $N'M$ is a non-negative matrix, and $N'N$ is a positive diagonal matrix. Thus, $v_t(\theta^{+}r-x,r)$ is supermodular by Remark 1 in Chen et al. (2013).

It remains to show the supermodularity of $v_{t+1}(\theta^{-}r-x,r)$. Note that $\theta^{-}\ge\theta^{+}$ and

$$v_{t+1}(\theta^{-}r-x,r)=v_{t+1}(\theta^{+}r-[x-(\theta^{-}-\theta^{+})r],r).$$

Because $v_{t+1}(\theta^{+}r-x,r)$ is supermodular and jointly concave in (x,r), by Corollary 2 in Chen et al. (2013), $v_{t+1}(\theta^{-}r-x,r)$ given above is supermodular in (x,r). □

Proof of Theorem 3

Let $u_t(q) = \max_y \mathbb{E}[\gamma v_{t+1}(y-\varepsilon_t, q) - h(y-\varepsilon_t)]$. By Theorem 1, $q_t(r)$ solves the problem

$$\underset{q}{\text{maximize}} \ \left\{\pi(r,q) + u_t(q) : p(r,q) \in [\underline{p}, \overline{p}]\right\},$$

where the objective function is supermodular and strictly concave in q (see the proof of Proposition 1), and thus $q_t(r)$ is increasing in r by Theorem 2.2.8 in Simchi-Levi et al. (2014) and continuous by Berge's Maximum Theorem.

For the existence of $Q_t^{\pm}$, it suffices to show that $z_t(r) = r - q_t(r)$ is increasing in r since in this case, we can select $Q_t^- = \max\left\{r \in [\underline{p}, \overline{p}] : z_t(r) < 0\right\}$ and $Q_t^+ = \min\left\{r \in [\underline{p}, \overline{p}] : z_t(r) > 0\right\}$, where $Q_t^- = \underline{p}$ when $z_t(\underline{p}) \geq 0$, and $Q_t^+ = \overline{p}$ when $z_t(\overline{p}) \leq 0$. To see it, note that $z_t(r)$ solves the problem:

$$\underset{z}{\text{maximize}} \ \left\{\pi(r, r-z) + u_t(r-z) : r - z/(1-\alpha) \in [\underline{p}, \overline{p}]\right\}.$$

We can verify that $\frac{\partial}{\partial r}\pi(r, r-z)$ depends on z via the term $[2Az + \eta(z)]/(1-\alpha) - 2\gamma\lambda z$ with λ given in (3). Thus, $\pi(r, r-z)$ is supermodular in (r,z) if $2A + \eta^+ \geq 2(1-\alpha)\lambda$, that is,

$$\eta^- - \eta^+ \leq A(1-\alpha) + \sqrt{[A(1+\alpha)]^2 + 2A(1+\alpha)\eta^-}. \tag{8}$$

Notice that the above inequality holds under Assumption 1. Therefore, $\pi(r, r-z)$ is indeed supermodular in (r,z). Furthermore, by joint concavity of v_{t+1}, u_t is concave and hence $u_t(r-z)$ is also supermodular in (r,z). Therefore $z_t(r)$ is increasing in r.

To characterize $s_t(r)$ in the loss neutral case $\eta(z) = \eta z$, we adopt the transformation technique again. In particular, consider the transformed profit-to-go functions below for all t:

$$\hat{v}_t(x,r) = v_t^0(x,r) - \hat{\lambda} r^2 - cx,$$

where we select $\hat{\lambda} = (\alpha A + \eta)\eta/[2(1-\alpha)(A+\eta)]$. Since the expect demand function $d = d(r,q)$ is linear in this case, we can express price $p = \hat{p}(d,q)$ and reference price $r = \hat{r}(d,q)$ as linear functions of expected demand d and target reference price r, respectively. Similar to the proof of Theorem 1, $s_t(r)$ maximizes the following function in term of s for any fixed r:

$$\hat{u}_t(s,r) = \underset{y,d}{\text{maximize}} \ \ \hat{\pi}(d,q) - \mathbb{E}h(y-\varepsilon_t) + \gamma \mathbb{E}\hat{v}_{t+1}(y-\varepsilon_t, q), \tag{9a}$$

$$\text{subject to} \ \ \hat{r}(d,q) = r, \ d + y = x, \ \hat{p}(d,q) \in [\underline{p}, \overline{p}]. \tag{9b}$$

where $\hat{\pi}_t$ defined below represents the transformed one-period expected revenue:

$$\hat{\pi}(d,q) = [\hat{p}(d,q) - c]\, d - \hat{\lambda}[\hat{r}(d,q)]^2 + \gamma\hat{\lambda} q^2.$$

Observe that the selected $\hat{\lambda}$ ensures that $\hat{\pi}(d,q) = C_1 d^2 + C_2 q^2$ for some non-positive coefficients C_1 and C_2. Thus, $\hat{\pi}$ is jointly concave and supermodular. Furthermore, note that

$$\hat{v}_t(x,r) = \underset{s}{\text{maximize}} \ \{\hat{u}_t(s,r) : s \geq x\}. \tag{10}$$

By applying Theorem 1 in Chen et al. (2013) to problem (9) and Theorem 2.2.8 in Simchi-Levi et al. (2014) to problem (10), it can be inductively proved that $\hat{u}_t$ and $\hat{v}_t$ are jointly concave and supermodular for all t. Thus, $s_t(r)$ is increasing in r.

When consumers are loss averse, as illustrated in Remark 5.3 in Gimpl-Heersink (2008), $s_t(r)$ does not exhibit a specific monotonic pattern even in the single-period case. □

Proof of Theorem 4

By a routine technique in dynamic programming, we can show that the profit-to-go function satisfies the following Bellman equation:

$$v(x,r) = \underset{y,q}{\text{maximize}} \ \ \pi(r,q) - \mathbb{E}h(y-\varepsilon) + \gamma \mathbb{E}v(y-\varepsilon, q), \tag{11a}$$

$$\text{subject to} \ \ d(r,q) + y \geq x, \ \ p(r,q) \in [\underline{p}, \overline{p}]. \tag{11b}$$

We now study the asymptotic property of any state path $\{(x_t, r_t) : t \geq 1\}$ under the optimal policy of problem (11). Because a base-stock policy is optimal, and the realized demand is positive with non-zero probability as assumed, there exists a finite τ such that an order is placed at period $t = \tau$ with probability 1. Construct $\{(\tilde{x}_t, \tilde{r}_t) : t \geq 1\}$ such that $(\tilde{x}_t, \tilde{r}_t) = (x_t, r_t)$ when $t \leq \tau$, and $\{(\tilde{x}_t, \tilde{r}_t) : t \geq \tau\}$ be a state path of the problem below:

$$\tilde{v}(r) = \underset{y,q}{\text{maximize}} \ \{\pi(r,q) - \mathbb{E}h(y-\varepsilon_t) + \gamma \tilde{v}(q) : p(r,q) \in [\underline{p}, \overline{p}]\}. \tag{12}$$

Notice that the above problem indicates a system allowing the firm to return the product to the manufacturer and obtain a full refund. In any period $t > \tau$, clearly $\tilde{v}(\tilde{r}_t) \geq v(\tilde{x}_t, \tilde{r}_t)$, and the optimal target inventory level of problem (12) is $y^m + d(\tilde{r}, \tilde{r}_{t+1})$ with $y^m = \arg\min_y \mathbb{E}h(y - \varepsilon_t)$. Because ε_t are identically distributed and $d(r,q) + \varepsilon_t$ is non-negative for any $r, q \in \mathcal{P}$, we have $\tilde{x}_{t+1} = y^m - \varepsilon_t < y^m + d(\tilde{r}_{t+1}, \tilde{r}_{t+2})$. Thus, orders are placed in all periods $t > \tau$ for problem (12), implying that $\{(\tilde{x}_t, \tilde{r}_t) : t \geq 1\}$ is also a state path of problem (11) under some possibly suboptimal policy and hence $\tilde{v}(\tilde{r}_t) \leq v(\tilde{x}_t, \tilde{r}_t)$. In summary, we have $\tilde{v}(\tilde{r}_t) = v(\tilde{x}_t, \tilde{r}_t)$, that is, a state path under the optimal policy of problem (11) mimics that of problem (12) once an order is placed. Hence, both problems have the same steady state reference prices. On the other hand, for problem (12), because the optimal y is given by y^m, it reduces to the optimality equation corresponding to the dynamic pricing model with reference price effect analyzed in Popescu and Wu (2007). Specifically, they establish that all steady state reference prices of problem (12) form an interval $[R^-, R^+]$ with $R^\pm = R(\eta^\pm)$, where $R(\eta)$ is the steady sate reference price of the dynamic pricing problem with a loss neutral reference price effect function $\eta(z) = \eta z$ and is decreasing in η. Because the list price p^* for the model without the reference price effect is equal to $R(0)$, we conclude that $p^\infty \leq p^*$ and $s(p^\infty) = y^m + d(p^\infty, p^\infty) \geq y^m + d(p^*, p^*) = s^*$ for any $p^\infty \in [R^-, R^+]$. □

References

Chen, X., P. Hu, S. He. 2013. Preservation of supermodularity in parametric optimization problems with non-lattice structures. *Operations Research* **61**(5) 1166–1173.

Chen, X., P. Hu, Z. Hu. 2016. Efficient algorithms for dynamic pricing problem with reference price effect. *Submitted to Management Science* .

Chen, X., D. Simchi-Levi. 2012. Pricing and inventory management. *Oxford Hanbook of Pricing Management, eds. Philips P and Ozer, O., Oxford University Press, United Kingdom* 784–822.

Federgruen, A., A. Heching. 1999. Combined pricing and inventory control under uncertainty. *Operations Research* **47**(3) 454–475.

Fibich, G., A. Gavious, O. Lowengart. 2003. Explicit solutions of optimization models and differential games with nonsmooth (asymmetric) reference-price effects. *Operations Research* **51**(5) 721–734.

Gimpl-Heersink, D. I. L. 2008. Joint pricing and inventory control under reference price effects. Ph.D. thesis, Vienna University of Economics and Business Administration.

Greenleaf, E.A. 1995. The impact of reference price effects on the profitability of price promotions. *Marketing Science* **14**(1) 82–82.

Hardie, B.G.S., E.J. Johnson, P.S. Fader. 1993. Modeling loss aversion and reference dependence effects on brand choice. *Marketing Science* **12**(4) 378–394.

Hu, P. 2011. Coordinated dynamic pricing and inventory management. Ph.D. thesis, University of Illinois at Urbana-Champaign.

Hu, Z., X. Chen, P. Hu. 2016. Dynamic pricing with gain-seeking reference price. *Operations Research* **64**(1) 150–157.

Kahneman, D., A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* **47**(2).

Kalyanaram, G., J.D.C. Little. 1994. An empirical analysis of latitude of price acceptance in consumer package goods. *Journal of Consumer Research* **21**(3) 408–418.

Kalyanaram, G., R. S. Winer. 1995. Empirical generalizations from reference price research. *Marketing Science* **14**(3) G161–G169.

Kopalle, P.K., A.G. Rao, J.L. Assunção. 1996. Asymmetric reference price effects and dynamic pricing policies. *Marketing Science* **15**(1) 60–85.

Liu, Q., G. van Ryzin. 2011. Strategic capacity rationing when customers learn. *Manufacturing & Service Operations Managment* **13**(1) 89–107.

Mazumdar, T., Raj S. P., Sinha I. 2005. Reference price research: Review and propositions. *Journal of Marketing* **69**(4) 84–102.

Nasiry, J., I. Popescu. 2011. Dynamic pricing with loss-averse consumers and peak-end anchoring. *Operations Research* **59**(6) 1361–1368.

Popescu, I., Y. Wu. 2007. Dynamic pricing strategies with reference effects. *Operations Research* **55**(3) 413–429.

Simchi-Levi, D., X. Chen, J. Bramel. 2014. *The logic of logistics: Theory, algorithms, and applications for logistics management (3rd edition)*. Springer-Verlag, New York, NY.

Urban, T.L. 2008. Coordinating pricing and inventory decisions under reference price effects. *International Journal of Manufacturing Technology and Management* **13**(1) 78–94.

Yang, L., F. de Vericourt, P. Sun. 2014a. Time-based competition with benchmark effects. *Manufacturing & Service Operations Management* **16**(1) 119–132.

Yang, L., P. Guo, Y. Wang. 2014b. Service pricing with loss averse customers. Working Paper.