

Multi-Product Dynamic Pricing with Reference Effects Under Logit Demand

Mengzi Amy Guo¹, Hansheng Jiang², Zuo-Jun Max Shen³

Department of Industrial Engineering and Operations Research, University of California, Berkeley

¹ mengzi_guo@berkeley.edu, ² hansheng.jiang@rotman.utoronto.ca, ³ maxshen@berkeley.edu

Problem definition: We consider the dynamic pricing problem of multiple products under (asymmetric) reference effects over an infinite horizon. Unlike existing literature mostly focused on the single-product setting, our multi-product setting takes into account the cross-product effects among substitutes and incorporates the memory-based reference prices into the multinomial logit (MNL) demand model. Even with the single-product logit demand, the structure of the optimal pricing policy is intractable. Therefore, we focus on the long-run patterns of the optimal pricing policy and also discuss the performance of the myopic pricing policy. **Methodology/results:** We first provide a comprehensive characterization of the myopic pricing policy, including its solution, long-run convergence behavior, and optimality gap. For the optimal pricing policy, we show an intricate connection between its long-run dynamics and types of reference effects. We demonstrate that the presence of any gain-seeking product renders a long-run constant pricing policy suboptimal. Conversely, the constant policy (or optimal steady state) can exist in both loss-neutral and loss-averse scenarios, where we provide a sufficient condition for such existence. We further show that when pricing perfect substitutes, the true optimal policy under the multi-product framework is more likely to yield a long-run cyclic pattern than the policy derived from the single-product framework, a phenomenon that aligns well with the periodic discounts in real-world markets. This highlights the importance of employing the multi-product framework, as sticking to the single-product framework when pricing multiple products can misrepresent long-run dynamics and result in sub-optimality. **Managerial implications:** The discrepancy in the long-run behaviors between multi-product-based and single-product-based policies underscores the critical need for incorporating cross-product effects while dealing with more than one substitute. In the multi-product domain, our model suggests that retailers are more likely to benefit from appropriate price variations than maintaining a constant price.

Key words: multi-product dynamic pricing, multinomial logit model, reference effect, retailing

1. Introduction

In modern retailing, a monopolist seller commonly adopts frequent price changes and various promotional activities to learn consumer preferences and maximize their revenue. It is commonly assumed that demand at each period is negatively correlated with prices offered and that revenue is maximized at some fixed price. Consumers' decision processes, however, are known to exhibit bounded rationality. For repeatedly visiting consumers, they may compare the current prices with some reference prices when making purchase decisions. Reference prices represent the anchor price

points formed by consumers, and are usually shaped by historical prices. With the rapid rise of e-commerce and online selling, prices are easier to track, and thus historical prices are becoming more accessible to consumers. For example, camelcamelcamel.com offers a complimentary price tracking service for the leading US online retailer Amazon (as shown in Figure 1a). Similarly, Alibaba, the largest online shopping platform in China, provides a functionality that allows users to monitor the price history (as shown in Figure 1b).

Figure 1 Price Tracking Tools for Amazon (left) and Alibaba (right)



Reference price effects have roots in the renowned prospect theory in economics, dating back to Kahneman and Tversky (1979). The prospect theory postulates that consumers respond more strongly to losses (when the price is higher than the reference price) than gains (when the price is lower than the reference price) of equal quantity, and therefore consumers behave in a *loss-averse* manner. The opposite, known as the *gain-seeking* behavior, when consumers respond more strongly to gains than losses, has also been established both empirically and theoretically (Greenleaf 1995, Hu et al. 2016, Hu and Nasiry 2018). Both loss-averse and gain-seeking behaviors are accompanied by asymmetric reference effects and directly influence consumer demand under dynamic pricing policies. Hence, it is crucial that sellers take reference effects into consideration when designing their pricing strategies and promotional activities.

While the concept of reference effects originates in economics and marketing fields, there is a growing trend in OM literature that theoretically studies the intertemporal pricing problem under reference effects (see, e.g., Wu et al. (2015), Wang (2016), Hu et al. (2016), Chen et al. (2017), Chen and Nasiry (2020), den Boer and Keskin (2021), Jiang et al. (2022)). Existing research observes that the optimal pricing policy under reference effects tends to be periodically cyclic with gain-seeking consumers (Popescu and Wu 2007, Hu et al. 2016) and more likely to converge to a static price with loss-averse consumers (Greenleaf 1995, Kopalle et al. 1996, Popescu and Wu

2007). More recently, several works argue that the loss-aversion does not necessarily prevent cyclic pricing patterns when new components are incorporated, e.g., varied consumer shopping schedules (Wang 2016), consumer heterogeneity (Chen and Nasiry 2020), strong seasonality effect on the demand function (Chen et al. 2017), or non-uniformity in consumer evaluation (which can create a mismatch between individual-level behaviors and aggregated demand responsiveness, according to Hu and Nasiry (2018)). However, all the aforementioned works focus on single-product scenarios, and little is known about the pricing policy under a multi-product setting, where the demand for each product is affected not only by its own price and reference price but also by those of other products in the same choice set. Our research aims to fill in this blank by studying the intertemporal pricing problem with reference effects in the context of multiple products.

We consider a monopolist selling a variety of products over an infinite horizon, with the demand function characterized by the discrete choice model—the multinomial logit (MNL) model (McFadden 1974). To account for the reference effects, we incorporate the memory-based reference price into the utility function of the MNL model. By doing so, we simultaneously account for the vertical effect of price comparison in time and the horizontal effect of product comparison within the offer set. Furthermore, we allow product-differentiated responsiveness parameters for both prices and reference effects in utility functions, so that consumers may exhibit different behaviors with respect to different products. Aligning with the current trend in single-product research, we are interested in exploring the long-run dynamics of the optimal pricing policy in this multi-product framework. In particular, we seek to investigate the following questions: *(i) moving from a single to multiple products, whether the gain-seekingness still leads to the cyclic pricing in the long run; (ii) to what extent does loss-aversion antagonize cyclic pricing patterns; and (iii) when pricing multiple alternatives, how does the price pattern under our multi-product framework differ from that of the baseline approach, where each product is priced separately using the single-product model.* Intuitively, adhering to the single-product model when pricing multiple substitutes may not only be sub-optimal but also result in misleading long-run behaviors due to the ignorance of cross-product effects. Our work theoretically shows that when managing similar substitutes, taking into account their cross-product effects in the multi-product framework facilitates long-run cyclic price patterns. More broadly, understanding the long-run pricing behavior is crucial for retailers from a managerial standpoint, as it directly associates with demand stability, which further influences the upstream operations in supply chain management.

1.1. Main Contributions

To our knowledge, this is the first study in the OM literature that explores reference effects in a multi-product, multi-period, discrete-time setting. We propose a dynamic pricing framework with

the MNL demand, which accommodates both cross-period effects induced by reference effects and cross-product effects in a multi-product context. The main contributions are summarized below.

1. Our myopic pricing policy, which involves maximizing the single-period revenue with asymmetric reference effects under the MNL demand, extends the existing literature by incorporating product-differentiated sensitivities to both price and reference price, as well as allowing different reference prices across products in the choice set. From a technical point of view, we introduce a novel linear transformation to the original first-order condition. This transformed first-order condition is the key to deriving structural characteristics of the optimal single-period price (i.e., myopic price), with which we manage to simplify this problem of single-period revenue optimization from a multi-dimensional nonsmooth and nonconcave problem to a one-dimensional binary search on the admissible price range. Consequently, we can efficiently compute the myopic pricing policy with an ϵ -approximation in $\mathcal{O}(\log(1/\epsilon))$ number of floating-point operations, a significant improvement over the $\mathcal{O}(1/\epsilon^n)$ complexity of greedy search methods, where n is the number of products. More broadly, the technique developed to solve the myopic price could also shed light on other single-period multi-product pricing problems. In addition, when repeatedly applying the myopic pricing policy, we show that the long-run convergence behavior completely depends on whether some products are associated with gain-seeking reference effects.

2. Using the myopic value function, we provide a performance lower bound for the optimal pricing policy (see Proposition 2). The optimality gap shrinks (and becomes zero in extreme cases) under any of the three situations: (i) the weight on future revenues decreases; (ii) the memory effect of reference prices increases; or (iii) the reference effect diminishes.

3. Given the computational intractability of the optimal pricing policy due to the curse of dimensionality, we focus on the policy's long-run behavior. In particular, with the presence of any gain-seeking products in the choice set, the optimal policy never admits a steady state, resulting in a cyclic policy over the long term. By contrast, the optimal steady state can exist in loss-neutral and loss-averse scenarios (i.e., no gain-seeking products in the set), for which we characterize the structure of the optimal steady state. We then show for a special loss-averse case (i.e., $c_i^- \geq c_i^+ = 0$, $\forall i \in N$) that the optimal policy always admits a steady state. From a technical standpoint, we propose two distinct types of perturbation analyses on the optimal value function. We believe these methodologies have potential applications to a broader class of dynamic programming problems.

4. From a managerial perspective, our research (see Theorem 7) underscores the importance of employing the multi-product framework when pricing similar substitutes. Attempting to use the single-product model to price multiple products separately may result in misleading long-run behavior, as this approach is more likely to produce a constant policy while the true optimal strategy is cyclic. The numerical experiments further highlight an intriguing observation that the cyclic optimal policy periodically hits the price ceiling.

1.2. Related Literature

Our work is related to several lines of literature, including dynamic pricing under reference effects (mostly in the single-product setting), the modeling of reference effects in the multi-product setting, and price optimization under the MNL type demand models.

1.2.1. Dynamic Pricing with Reference Effects. Under a discrete-time setting, the works most relevant to us are on dynamic pricing under reference effects, including Popescu and Wu (2007), Hu et al. (2016), Chen et al. (2017), Chen and Nasiry (2020), Jiang et al. (2022), and all these works focus on price optimization in a single-product setting. The classical results by Popescu and Wu (2007) are based on general demand functions, but their analyses require that the demand satisfies certain assumptions (Popescu and Wu 2007, Assumptions 1, 2, and 3), and importantly, these assumptions are all violated by the logit demand model with asymmetric reference effects. As a result, the analyses on pricing policies in Popescu and Wu (2007) are not applicable to our problem based on logit demand models, not even in the single-product case. Recent revenue management works on dynamic pricing under reference effects are mostly based on the piecewise linear demand, with new results on more comprehensive characterizations of myopic and optimal pricing policies as well as computations of them in a discretized price space (Hu et al. 2016, Chen et al. 2017, Chen and Nasiry 2020). Jiang et al. (2022) use the logit demand in a single-product setting, with a special focus on consumer heterogeneity and model estimation.

In addition, we note that there are interesting works on dynamic pricing under reference effects in a continuous time setting, which is very different from this paper's discrete-time setting in terms of analysis techniques and applications. Fibich et al. (2003) explicitly solve the single-product optimal pricing policy under a continuous time framework by considering a system of differential equations and prove that the pricing policy converges to a steady state in an infinite horizon. Hu (2015, Chapter 5) further generalizes the piecewise linear demand model of Fibich et al. (2003) to a multi-product setting. However, in the context of asymmetric reference effects, they limit the discussion to a two-product case due to the challenge of non-smoothness.

More broadly, besides reference effects, there has been active research on dynamic pricing under other consumer behaviors, where consumers may be patient, forward-looking, or strategic (see, e.g., Su (2007), Aviv and Pazgal (2008), Liu and Cooper (2015), Besbes and Lobel (2015), Lobel (2020), Araman and Fayad (2021), Zhang and Jasin (2022)).

1.2.2. Modeling of Reference Effects Under MNL Model. Logit demand models that incorporate reference effects have been considered in the marketing literature (Briesch et al. 1997, Chang et al. 1999, Kopalle et al. 2012, Bell and Lattin 2000). These works provide strong empirical evidence and motivating observations, but they do not theoretically establish any structural results

for myopic and optimal pricing policies. In the multi-product setting, Briesch et al. (1997) conduct a comparative analysis on the scanner panel data and conclude that the model depending on a product's own price history provides the best fit and strongest predictive power. Accordingly, our work is built upon this type of reference price model, known as the "brand-specific past prices" formulation (Briesch et al. 1997), which falls within the memory-based reference price category, a formulation widely adopted in the relevant literature (see, e.g., Kopalle et al. (1996), Mazumdar et al. (2005), Popescu and Wu (2007), Nasiry and Popescu (2011), Hu et al. (2016), Chen et al. (2017), and Hu and Nasiry (2018)).

A notable work by Wang (2018) incorporates reference effects into the MNL demand model differently. Instead of the memory-based reference price models, Wang (2018) defines the reference price as the lowest/highest price in the assortment, and it is subject to changes if the assortment differs. Wang (2018) focuses on assortment planning and pricing in a single-period setting rather than the pricing policies in an intertemporal setting. Our model also differs from Wang (2018) in that we allow different products to have different reference prices instead of sharing one reference price across the whole product set, and that we allow product-differentiated sensitivities to the price and the reference price while these sensitivities are assumed to be identical across products in Wang (2018). Lastly, although our work focuses on the memory-based reference price model, we believe it is interesting to study pricing problems under other reference price models as well, especially if they better portray reality in some situations.

1.2.3. Single-Period Price Optimization Under Logit Demand. In terms of modeling and methodology, our analysis of the myopic pricing policy is related to the rich literature on single-period price optimization under the MNL model and its variants, of which we list a few here, Hopp and Xu (2005), Li and Huh (2011), Wang (2012), Gallego and Wang (2014), Du et al. (2016), Zhang et al. (2018), Wang (2018), and Li et al. (2020). The classical result from Hopp and Xu (2005) concludes that when a uniform price sensitivity is assumed, the optimal single-period price with the MNL demand leads to the same markup for all products, where the markup here is defined as the price minus the cost. Subsequently, Li and Huh (2011) extend the result to the nested multinomial logit demand, proving that products within a nest admit the same markup. The follow-up work by Wang (2012) accommodates product-differentiated price sensitivities and shows that a so-called adjusted markup is the same for all products. Our work contributes to this stream of research by incorporating the non-smooth product-differentiated reference effects into the MNL model, where we propose a novel approach for solving its single-period pricing problem.

1.3. Organization

The rest of the paper is organized as follows: In Section 2, we introduce the problem formulation and necessary concepts. We then present a comprehensive analysis of the myopic pricing policy in Section 3, including its solution and long-run convergence behaviors. Section 4 is devoted to the optimal pricing policy, where we examine its performance guarantee and explore its long-run behaviors. We further discuss some intriguing observations via numerical experiments in Section 5 and conclude the paper in Section 6. All proofs are documented in the supplemental materials.

2. Preliminaries

2.1. Multi-Product Demand Model Under Reference Effects

We consider the scenario where a monopolist sells a set of n substitutable products, denoted by $N := \{1, 2, \dots, n\}$, with unlimited inventory to a market of homogeneous consumers over an infinite horizon. We employ the brand-specific past prices (*PASTBRSP*) reference price model from Briesch et al. (1997), which assumes that the reference price is memory-based and product-specific. In particular, the reference price for product i is modeled by exponentially smoothing its own historical prices, as formally presented in Definition 1.

DEFINITION 1 (EXPONENTIAL SMOOTHING UPDATE OF REFERENCE PRICE). At period t , we use $\mathbf{p}^t = (p_1^t, p_2^t, \dots, p_n^t)$ and $\mathbf{r}^t = (r_1^t, r_2^t, \dots, r_n^t)$ to denote the prices and reference prices of products in set N , respectively. The reference price update mechanism is defined as

$$\mathbf{r}^{t+1} = \alpha \mathbf{r}^t + (1 - \alpha) \mathbf{p}^t, \quad \forall t \geq 0, \quad (1)$$

where $\alpha \in [0, 1]$ is the memory parameter.

This update mechanism implies that the reference price is more strongly influenced by recent prices compared to older ones. Meanwhile, when the memory parameter α increases, consumers tend to be more conservative in absorbing new prices to adjust their reference price.

Existing papers on the single-product scenario (Greenleaf 1995, Hu et al. 2016, Chen et al. 2017, Chen and Nasiry 2020) mostly employ the linear demand model, partly due to its favorable analytical properties. However, the linear demand does not naturally extend to the multi-product scenario unless additional pairwise cross-effect terms are introduced. In our multi-product framework, we utilize the MNL demand that naturally captures the interactions between products through the random utility maximization theory (McFadden 1974).

The utility of purchase is characterized by a set of product-specific responsiveness parameters $(\mathbf{a}, \mathbf{b}, \mathbf{c}^+, \mathbf{c}^-)$, where $\mathbf{a} = (a_1, a_2, \dots, a_n)$ refers to the intrinsic utility, and $\mathbf{b} = (b_1, b_2, \dots, b_n)$ represents consumers' responsiveness to price, also known as price sensitivity. Consumers' responsiveness to perceived gains and losses through reference effects is characterized by reference price sensitivities $\mathbf{c}^+ = (c_1^+, c_2^+, \dots, c_n^+)$ and $\mathbf{c}^- = (c_1^-, c_2^-, \dots, c_n^-)$. We assume that $\mathbf{b} > \mathbf{0}$ and $\mathbf{c}^+, \mathbf{c}^- \geq \mathbf{0}$, meaning

that consumer utilities depend negatively on the prices and depend positively on the reference effects. This assumption aligns with consumer behaviors toward substitutes and has also been commonly adopted in the similar pricing literature (see, e.g., [Hu et al. \(2016\)](#), [Chen et al. \(2017\)](#), [Golrezaei et al. \(2020\)](#)). Formally, we define the utility for purchasing product $i \in N$ at period t as

$$U_i^t = a_i - b_i \cdot p_i^t + c_i^+ \cdot (r_i^t - p_i^t)_+ - c_i^- \cdot (p_i^t - r_i^t)_+ + \epsilon_i^t, \quad (2)$$

where $(\cdot)_+ := \max\{\cdot, 0\}$ and ϵ_i^t is the random effect that follows i.i.d. standard Gumbel distribution. Without loss of generality, let the no-purchase utility at period t be $U_0^t = \epsilon_0^t$. Based on the random utility maximization theory ([McFadden 1974](#)), the purchase probability (or market share) of product $i \in N$ at period t has the explicit form as

$$q_i(\mathbf{p}^t, \mathbf{r}^t) = \frac{\exp[a_i - b_i \cdot p_i^t + c_i^+ \cdot (r_i^t - p_i^t)_+ - c_i^- \cdot (p_i^t - r_i^t)_+]}{1 + \sum_{k \in N} \exp[a_k - b_k \cdot p_k^t + c_k^+ \cdot (r_k^t - p_k^t)_+ - c_k^- \cdot (p_k^t - r_k^t)_+]}. \quad (3)$$

Consequently, the total revenue in the period t can be expressed as

$$\Pi(\mathbf{p}^t, \mathbf{r}^t) = \sum_{i \in N} p_i^t \cdot q_i(\mathbf{p}^t, \mathbf{r}^t) = \sum_{i \in N} \frac{p_i^t \cdot \exp[a_i - b_i \cdot p_i^t + c_i^+ \cdot (r_i^t - p_i^t)_+ - c_i^- \cdot (p_i^t - r_i^t)_+]}{1 + \sum_{k \in N} \exp[a_k - b_k \cdot p_k^t + c_k^+ \cdot (r_k^t - p_k^t)_+ - c_k^- \cdot (p_k^t - r_k^t)_+]}. \quad (4)$$

A favorable consequence of the asymmetric reference effect is that the MNL model in Eq. (3) no longer suffers from *IIA property* (see [Debreu \(1952\)](#)) and *simple scalability* (see [Tversky \(1972\)](#)), which are the restrictions for general MNL models ([Wang 2018](#)).

2.2. Product-Specific Reference Effects

For the single-product setting, different definitions of reference effects are proposed in the existing literature, e.g., global, local, and regional reference effects ([Hu and Nasiry 2018](#), [Jiang et al. 2022](#)). In this work, we generalize the local reference effect definition to the multi-product setting, where the consumer behavior is categorized in a product-specific manner. Based on potentially asymmetric reactions consumers hold toward discounts and surcharges, the reference effect can be categorized into three types: *loss-averse*, *loss-neutral*, and *gain-seeking*, formally introduced in Definition 2.

DEFINITION 2 (TYPES OF REFERENCE EFFECTS). For product $i \in N$, consumers are more responsive to losses (i.e., loss-averse) if $c_i^+ < c_i^-$, equally responsive to losses and gains (i.e., loss-neutral) if $c_i^+ = c_i^-$, or more responsive to gains (i.e., gain-seeking) if $c_i^+ > c_i^-$.

Based on Definition 2, the products in N can be divided into three subsets N_l , N_n , and N_g , where N_l (resp. N_n or N_g) is constituted of all products i with $c_i^+ < c_i^-$ (resp. $c_i^+ = c_i^-$ or $c_i^+ > c_i^-$), i.e., consumers are more (resp. equally or less) sensitive to losses than gains for these products. Indeed, our Definition 2 complies with the concept of the local reference effect in [Hu and Nasiry \(2018, Definition 2\)](#). We defer the proof to Appendix A.1, where we also discuss the reason why we prefer the local reference effect over the global counterpart ([Hu and Nasiry 2018, Definition 1](#)).

2.3. Dynamic Price Optimization Problem

From a revenue maximization perspective, the intertemporal nature of reference effects leads to a sequential price optimization problem, which can be formulated as a dynamic program. We consider the price and reference price to be drawn from the same admissible price set, defined as a continuous interval $\mathcal{P} = [0, p_H]$, with the price ceiling $p_H > 0$ representing the maximum allowable price. Denote $\mathbf{p}^\pi : \mathcal{P}^n \rightarrow \mathcal{P}^n$ as a stationary pricing policy that maps reference price $\mathbf{r}$ to price $\mathbf{p}$. Then, for any initial reference price $\mathbf{r}^0$, the optimization problem can be written as

$$\max_{\mathbf{p}^\pi} V^\pi(\mathbf{r}^0) = \sum_{t=0}^{\infty} \beta^t \Pi(\mathbf{p}^\pi(\mathbf{r}^t), \mathbf{r}^t), \quad \text{s.t. } \mathbf{r}^{t+1} = \alpha \mathbf{r}^t + (1 - \alpha) \cdot \mathbf{p}^\pi(\mathbf{r}^t), \quad \forall t \geq 0, \quad (5)$$

where V^π is the long-term discounted revenue (or value function) under policy $\mathbf{p}^\pi$, and $\beta \in [0, 1]$ is the discount factor. We use V^* and $\mathbf{p}^*$ to denote the optimal value function and the optimal pricing policy, respectively. As $\mathbf{p}^t, \mathbf{r}^t \in \mathcal{P}^n$, it follows that the single-period revenue in Eq. (4) is bounded, thus the problem in Eq. (5) amounts to solving the Bellman equation:

$$V^*(\mathbf{r}^0) = \Pi(\mathbf{p}^*(\mathbf{r}^0), \mathbf{r}^0) + \beta V^*(\alpha \mathbf{r}^0 + (1 - \alpha) \cdot \mathbf{p}^*(\mathbf{r}^0)). \quad (6)$$

There exist two main obstacles in tackling this intertemporal price optimization problem. Firstly, from the analytical perspective, the MNL demand does not retain many of the favorable properties present in the linear demand found in the prior literature (Chen et al. 2017). For example, the revenue function under the single-product linear demand is a piecewise quadratic function, which is strictly supermodular in $(\mathbf{p}, \mathbf{r})$, and jointly concave in $\mathbf{p}$ and $\mathbf{r}$ for loss-averse and loss-neutral reference effects. Regrettably, such useful characteristics break down in the MNL model, complicating the analysis of pricing policies in our context. Specifically, $\Pi(\mathbf{p}, \mathbf{r})$ does not exhibit concavity in either $\mathbf{p}$ or $\mathbf{r}$ irrespective of the reference effect type. The second difficulty comes from a computational standpoint. In the single-product case, one can approximate the optimal policy by discretizing the continuous price space and then applying the policy iteration algorithm. The error of such approximation is bounded linearly by the granularity of the discretization (Jiang et al. 2022). However, the problem becomes intractable when moving to the multi-product case, as state and action spaces scale exponentially with the number of products. This makes the policy iteration computationally prohibitive when high accuracy is expected (Howard 1960, Bertsekas 2012).

In light of the difficulties tied to the optimal policy, our interest extends to heuristic policies that can provide performance guarantees relative to the optimal long-term discounted revenue. One such heuristic is the myopic pricing policy, which maximizes the single-period revenue at the current time and ignores the eroding consequence on future revenues. The myopic pricing policy can be formulated as the following single-period price optimization problem:

$$\mathbf{p}^m(\mathbf{r}) = \arg \max_{\mathbf{p} \in \mathcal{P}^n} \{\Pi(\mathbf{p}, \mathbf{r})\}, \quad \Pi^m(\mathbf{r}) = \max_{\mathbf{p} \in \mathcal{P}^n} \{\Pi(\mathbf{p}, \mathbf{r})\} = \Pi(\mathbf{p}^m(\mathbf{r}), \mathbf{r}), \quad (7)$$

where $\mathbf{p}^m : \mathcal{P}^n \rightarrow \mathcal{P}^n$ denotes the myopic policy, and $\Pi^m(\mathbf{r})$ represents the optimal single-period revenue given current reference price $\mathbf{r}$. The long-term discounted revenue of the myopic policy is

$$V^m(\mathbf{r}^0) = \sum_{t=0}^{\infty} \beta^t \Pi(\mathbf{p}^m(\mathbf{r}^t), \mathbf{r}^t), \quad (8)$$

where the reference price update follows $\mathbf{r}^{t+1} = \alpha \mathbf{r}^t + (1 - \alpha) \cdot \mathbf{p}^m(\mathbf{r}^t)$ for every $t \geq 0$.

2.4. Discussion on Price Ceiling

Both optimal and myopic pricing policies, as defined in Eq. (5) and Eq. (7), are constrained by the feasible price set $\mathcal{P} = [0, p_H]$. In the following proposition, we formalize the impact of the price ceiling p_H on these price optimization problems, with its proof deferred to Appendix A.2.

PROPOSITION 1 (Role of Price Ceiling). *The role of price ceiling is summarized as follows:*

1. Suppose that $\alpha < 1$, $\beta > 0$, and $\max_{i \in N} \{c_i^+\} > 0$. Then, for any initial reference price $\mathbf{r}^0$, the optimal value function $V^*(\mathbf{r}^0)$ is non-decreasing in the price ceiling p_H and $\lim_{p_H \rightarrow \infty} V^*(\mathbf{r}^0) = \infty$.
2. Let $\hat{p}_H$ be the unique solution p to Eq. (9).

$$p = \chi + \sum_{i \in N} \frac{1}{b_i + c_i^+} \cdot \exp(a_i - b_i p + (b_i + c_i^+) \chi - 1), \text{ where } \chi := \max_{i \in N} \left\{ \frac{1}{b_i + \min\{c_i^+, c_i^-\}} \right\}. \quad (9)$$

Then, for any $p_H \geq \hat{p}_H$, the price ceiling p_H is a non-binding constraint for the myopic optimization problem in Eq. (7) across all reference price $\mathbf{r} \in \mathcal{P}^n$, i.e.,

$$\mathbf{p}^m(\mathbf{r}) = \arg \max_{\mathbf{p} \in \mathcal{P}^n} \{\Pi(\mathbf{p}, \mathbf{r})\} = \arg \max_{\mathbf{p} \in \mathbb{R}^n} \{\Pi(\mathbf{p}, \mathbf{r})\}, \forall \mathbf{r} \in \mathcal{P}^n. \quad (10)$$

REMARK 1 (SOLVING FOR $\hat{p}_H$). In Eq. (9), we observe that the right-hand side decreases exponentially fast with respect to p . Hence, we can reasonably anticipate that $\hat{p}_H = \mathcal{O}(\max_{i,j \in N} a_i/b_j)$. Numerically, since the right-hand side of Eq. (9) is monotone decreasing in p , solving Eq. (9) is equivalent to finding the root of a univariate function. Therefore, the value of $\hat{p}_H$ can be explicitly determined by a one-dimensional bisection method.

Existing literature on price optimization with the intertemporal reference effect commonly assumes that the problem is constrained by a feasible price range (see, e.g., Popescu and Wu (2007), Hu (2015), Chen et al. (2017), Chen and Nasiry (2020), Jiang et al. (2022)). As we prove in Proposition 1, without a price ceiling, the optimal policy could explode and approach infinity. Intuitively, this is because a higher p_H leaves more flexibility for the optimal pricing policy to boost its revenue through a cyclic pattern, i.e., raising the reference price in one period to realize higher revenues for the following periods. Consequently, imposing a price ceiling is essential to ensure the problem is well-defined. Furthermore, Proposition 1 shows that the myopic value function $V^m(\mathbf{r})$ is

independent of the price ceiling p_H as long as $p_H \geq \hat{p}_H$, as in this case, p_H is an inactive constraint for the myopic optimization problem $\mathbf{p}^m(\mathbf{r})$ given any $\mathbf{r} \in \mathcal{P}^n$.

In fact, our myopic pricing policy has a very similar problem structure as the single-period price optimization under the MNL demand (see, e.g., Hopp and Xu (2005), Gallego and Wang (2014), Du et al. (2016), Wang (2018)). These studies typically formulate an unconstrained pricing optimization problem, since restricting the price ceiling might lead to pathological boundary solutions. On a high level, the requirement for setting a price range hinges on the nature of an optimization problem—if a problem remains well-defined without a price ceiling, then this constraint can be either omitted or made non-binding. Thereby, we assume, without loss of generality, that the price ceiling p_H is greater than or equal to $\hat{p}_H$. This setup ensures that the optimal policy remains well-defined and the myopic policy is unconstrained for any $\mathbf{r} \in \mathcal{P}^n$.

3. Myopic Pricing Policy

This section is dedicated to the investigation of the myopic pricing policy, as defined in Eq. (7). We emphasize that, without further specification, the choice set N in this section refers to a set of products with arbitrary combinations of reference effect types.

3.1. Structure and Computation of Myopic Pricing Policy

In the absence of reference effects, the myopic pricing policy can be obtained by solving a standard single-period multi-product price optimization problem—an area that has attracted substantial research interests (see, e.g., Li and Huh (2011), Gallego and Wang (2014), and the references therein). The incorporation of asymmetric reference effects, however, fundamentally changes the original problem by introducing the non-smoothness to the revenue function. Further, the complexities are intensified by the product-differentiated sensitivity parameters associated with the price and the reference price. There are two primary obstacles when attempting to solve the myopic policy in Eq. (7). Firstly, the non-smoothness in the objective hinders the application of a common technique—the market share transformation, i.e., using market share as the optimization variable (Li and Huh 2011, Du et al. 2016, Li et al. 2019). Secondly, the non-concavity of single-period revenue in price creates an extra challenge in discerning between global and local maxima.

In response to these technical challenges, we develop a novel approach that leverages the sub-gradient and a transformed version of the first-order condition. Particularly, we incorporate the sub-gradient and then apply the inverse Jacobian matrix $(\partial \mathbf{q} / \partial \mathbf{p})^{-1}$ to the first-order condition, which allows us to decouple the structure of myopic price $\mathbf{p}^m(\mathbf{r})$ and relate it to the associated single-period revenue $\Pi^m(\mathbf{r})$. With this relation in Eq. (B.11), we discover that a product with either the loss-averse or the loss-neutral reference effect admits a unique stationary point, facilitating easy identification of the optimum. By contrast, a gain-seeking product yields three stationary points,

as outlined in Eq. (B.14). Fortunately, we are able to pinpoint the optimum among these three candidates by exploiting the convexity of an intermediate function defined in Eq. (B.15).

As the explicit structure of the myopic pricing policy $\mathbf{p}^m(\mathbf{r})$ relies on the corresponding optimal single-period revenue $\Pi^m(\mathbf{r})$. We begin with characterizing $\Pi^m(\mathbf{r})$ for any given $\mathbf{r}$. To facilitate clarity in the presentation of Theorem 1, we introduce two auxiliary functions in Definition 3.

DEFINITION 3. For each loss-averse product i ($c_i^+ < c_i^-$), i.e., $i \in N_l$, define $\lambda_i(\Pi, r_i)$ as

$$\begin{aligned} \lambda_i(\Pi, r_i) := & \mathbb{1} \left\{ r_i < \Pi + \frac{1}{b_i + c_i^-} \right\} \frac{1}{b_i + c_i^-} \cdot \exp(a_i - (b_i + c_i^-)\Pi + c_i^- r_i - 1) \\ & + \mathbb{1} \left\{ \Pi + \frac{1}{b_i + c_i^-} \leq r_i \leq \Pi + \frac{1}{b_i + c_i^+} \right\} (r_i - \Pi) \cdot \exp(a_i - b_i r_i) \\ & + \mathbb{1} \left\{ r_i > \Pi + \frac{1}{b_i + c_i^+} \right\} \frac{1}{b_i + c_i^+} \cdot \exp(a_i - (b_i + c_i^+)\Pi + c_i^+ r_i - 1). \end{aligned} \quad (11)$$

For each gain-seeking product i ($c_i^+ > c_i^-$), i.e., $i \in N_g$, define $\gamma(\Pi, r_i)$ as

$$\begin{aligned} \gamma_i(\Pi, r_i) := & \mathbb{1} \{ r_i < \Pi + Z_i \} \frac{1}{b_i + c_i^-} \cdot \exp(a_i - (b_i + c_i^-)\Pi + c_i^- r_i - 1) \\ & + \mathbb{1} \{ r_i \geq \Pi + Z_i \} \frac{1}{b_i + c_i^+} \cdot \exp(a_i - (b_i + c_i^+)\Pi + c_i^+ r_i - 1), \end{aligned} \quad (12)$$

where $Z_i := \frac{\log(b_i + c_i^+) - \log(b_i + c_i^-)}{c_i^+ - c_i^-} \in \left[\frac{1}{b_i + c_i^+}, \frac{1}{b_i + c_i^-} \right]$.

THEOREM 1 (**Optimal Single-Period Revenue**). *The optimal single-period revenue $\Pi^m(\mathbf{r})$ for any given reference price $\mathbf{r}$ is equal to the unique solution Π of the equation:*

$$\Pi = \sum_{i \in N_l} \lambda_i(\Pi, r_i) + \sum_{i \in N_n} \frac{1}{b_i + c_i} \cdot \exp(a_i - (b_i + c_i)\Pi + c_i r_i - 1) + \sum_{i \in N_g} \gamma_i(\Pi, r_i), \quad (13)$$

where $c_i := c_i^+ = c_i^-$, $\forall i \in N_n$; $\lambda_i(\Pi, r_i)$ and $\gamma_i(\Pi, r_i)$ are introduced in Definition 3.

Despite the uniqueness of $\Pi^m(\mathbf{r})$ as shown in Theorem 1, solving a nonlinear equation is typically computationally intensive. Nevertheless, we circumvent this by utilizing the monotonicity of Eq. (13), as stated in the next corollary. The proofs for both results can be found in Appendix B.2.

COROLLARY 1 (**Computation of Optimal Single-Period Revenue**). *For a fixed reference price $\mathbf{r}$, the right-hand side of Eq. (13) is monotone decreasing with respect to Π . Hence, the unique value of $\Pi^m(\mathbf{r})$ can be determined by a one-dimensional bisection method.*

Corollary 1 conveys that the numerical computation of $\Pi^m(\mathbf{r})$ enjoys a linear convergence rate. Next, in Theorem 2, we explicitly show the relation between $\mathbf{p}^m(\mathbf{r})$ and $\Pi^m(\mathbf{r})$, thereby enabling the direct derivation of the myopic policy from its corresponding optimal single-period revenue. Particularly, for any product $i \in N$, we demonstrate that the difference of $p_i^m(\mathbf{r}) - \Pi^m(\mathbf{r})$ depends first on the type of reference effects and then on the relative magnitude between r_i and corresponding thresholds within that type. We relegate the proof of Theorem 2 to Appendix B.1.

THEOREM 2 (Myopic Pricing Policy). *For any reference price $\mathbf{r}$, the relation between myopic price $\mathbf{p}^m = \mathbf{p}^m(\mathbf{r})$ and its associated optimal single-period revenue $\Pi^m = \Pi^m(\mathbf{r})$ can be described as*

1. *For each loss-averse product i ($c_i^+ < c_i^-$), i.e., $i \in N_l$, the myopic price satisfies*

$$p_i^m = \begin{cases} \Pi^m + \frac{1}{b_i + c_i^-}, & \text{if } r_i < \Pi^m + \frac{1}{b_i + c_i^-}, \\ \Pi^m + \frac{1}{b_i + c_i^+}, & \text{if } r_i > \Pi^m + \frac{1}{b_i + c_i^+}, \\ r_i, & \text{otherwise.} \end{cases} \quad (14)$$

2. *For each loss-neutral product i ($c_i^+ := c_i^- = c_i$), i.e., $i \in N_n$, the myopic price satisfies*

$$p_i^m = \Pi^m + \frac{1}{b_i + c_i}. \quad (15)$$

3. *For each gain-seeking product i ($c_i^+ > c_i^-$), i.e., $i \in N_g$, the myopic price satisfies*

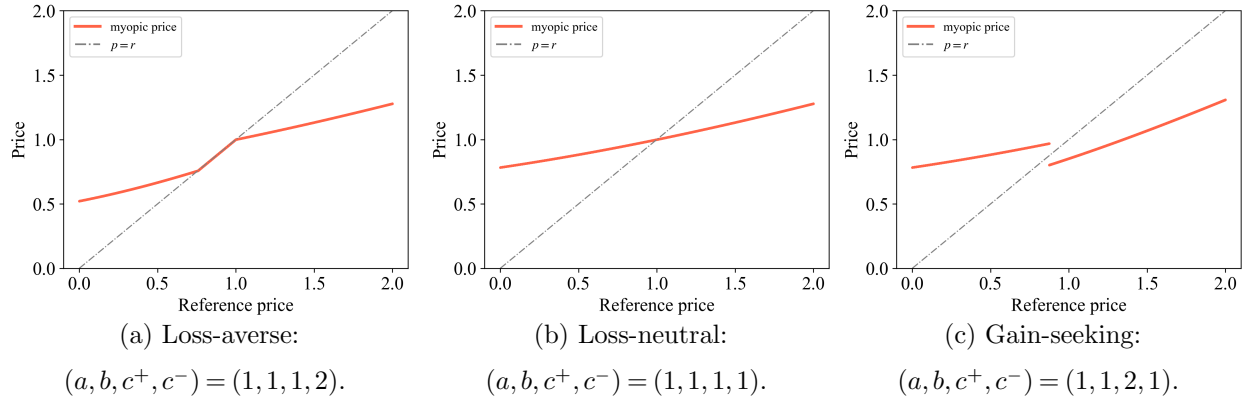
$$p_i^m = \begin{cases} \Pi^m + \frac{1}{b_i + c_i^-}, & \text{if } r_i < \Pi^m + Z_i, \\ \Pi^m + \frac{1}{b_i + c_i^+}, & \text{if } r_i \geq \Pi^m + Z_i, \end{cases} \quad (16)$$

where $Z_i = \frac{\log(b_i + c_i^+) - \log(b_i + c_i^-)}{c_i^+ - c_i^-} \in \left[\frac{1}{b_i + c_i^+}, \frac{1}{b_i + c_i^-} \right]$.

The myopic price for every product $i \in N$ equals $\Pi^m(\mathbf{r})$ plus a product-differentiated constant that only depends on the product's responsiveness parameters. Combining Theorems 1 and 2, we can solve the myopic pricing policy by calculating the associated single-period revenue $\Pi^m(\mathbf{r})$. This implies that the myopic policy, initially formulated as a multi-dimensional optimization problem in Eq. (7), is simplified to a one-dimensional equation in Eq. (13) that can be efficiently found by the bisection method. Therefore, an ϵ -approximation to the myopic policy can be computed within $\mathcal{O}(\log(1/\epsilon))$ number of floating-point operations. Additionally, $\Pi^m(\mathbf{r})$ is monotone increasing with respect to $\mathbf{r}$, meaning that a higher reference price yields a larger optimal single-period revenue.

REMARK 2. Our price optimization problem Eq. (7) is formulated in terms of revenue instead of profit. Maximizing the single-period profit (see, e.g., Hopp and Xu (2005), Li and Huh (2011)), defined as revenue minus total cost, can be reduced to the same type of optimization problem as maximizing the single-period revenue, and similar conclusions in Theorem 2 follow (see the proof in Appendix B.3). We comment that the *price* in a revenue maximization problem is essentially equivalent to the concept of *markup* in a profit maximization problem.

We visualize the myopic pricing policy in Figure 2, where the policy is illustrated using a single product merely for presentation clarity. In accordance with Theorem 2, a product with the loss-averse reference effect has its myopic pricing policy broken down into three segments (see Figure 2a), where the myopic price matches the reference price in the middle part; a product with the

Figure 2 Myopic Pricing Policy Illustration

loss-neutral reference effect has a smooth myopic pricing policy (see Figure 2b); and a product with the gain-seeking reference effect has its myopic pricing policy split into two disjoint segments (see Figure 2c). Different from loss-averse and loss-neutral scenarios, where there exists a reference price r such that $p^m(r) = r$, the disjunction in the gain-seeking scenario prohibits the existence of such a fixed point. This feature helps to explain the long-run convergence behavior of the myopic price path, which is discussed in Section 3.2.

Our myopic pricing policy is closely related to a stream of research on single-period price optimization (see, e.g., Hopp and Xu (2005), Li and Huh (2011), Wang (2012)), which shows that the optimal single-period price with the MNL demand leads to the same (adjusted) markup for all products, where the markup here is defined as the price minus the cost. When contrasting our myopic policy to the classical results for the logit demand without reference effects, a key distinction emerges: the markup (i.e., price, see Remark 2) in our problem depends on types of reference effects. Moreover, in loss-averse and gain-seeking scenarios, the myopic policy deviates from the classical markup rule due to the non-smoothness caused by asymmetric reference effects.

In addition, we compare our result with that of Wang (2018), who also investigates the optimal single-period profit with reference effects. Nevertheless, their reference price formulation differs from ours in that they use a common reference price, proportional to the cardinality of set N , for all products. Besides, Wang (2018) assumes identical reference price sensitivity across products, which then forces all products in the choice set to have the same type of reference effect. In contrast, our model allows for product-differentiated sensitivities to both price and reference price, thereby accommodating different types of reference effects in the choice set. From a technical perspective, we adopt a different line of analyses that performs a linear transformation to the first-order condition. By doing so, we achieve the improved results that can adapt to both asymmetric reference effects and product-differentiated sensitivity parameters.

3.2. Long-Run Behavior of Myopic Pricing Policy

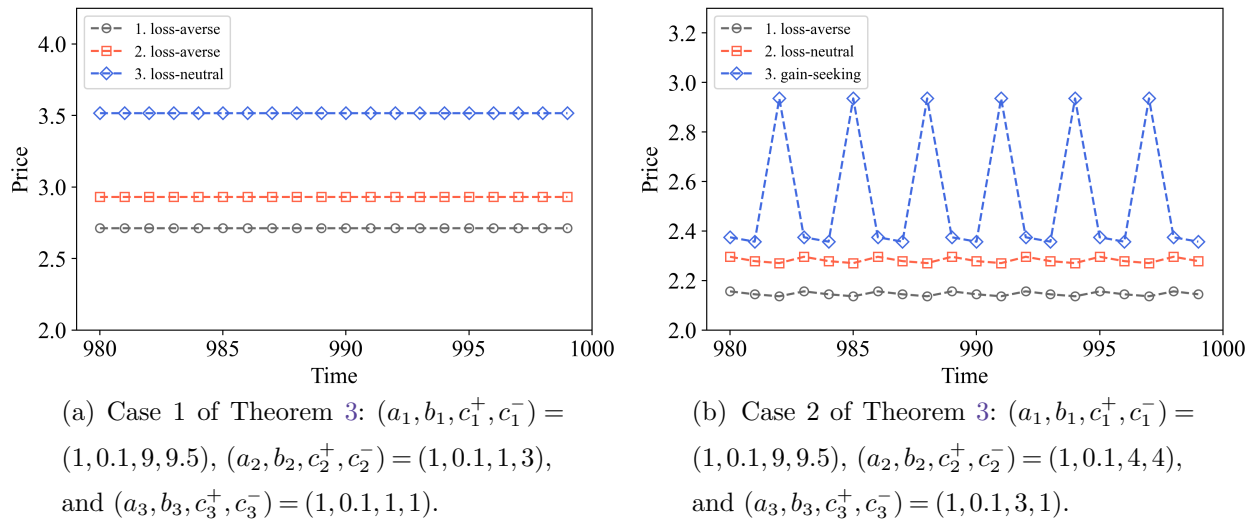
In this section, we investigate the long-run behavior of the price path under the myopic pricing policy and provide conditions under which it converges to a steady state. Formally, the long-run behavior of a price path $\{\mathbf{p}^t\}_{t=0}^\infty$ is defined as the limiting pattern that the path exhibits when t approaches infinity, and a pricing policy $\mathbf{p}^\pi(\cdot)$ has a steady state if there exists $\mathbf{r} \in \mathcal{P}^n$ such that $\mathbf{p}^\pi(\mathbf{r}) = \mathbf{r}$. We note that if the price path under $\mathbf{p}^\pi(\cdot)$ converges, $\mathbf{p}^\pi(\cdot)$ must admit a steady state.

THEOREM 3 (Long-Run Behavior of Myopic Pricing Policy). *The long-run convergence behavior of the myopic pricing policy depends on whether there exists a product with the gain-seeking reference effect:*

1. *If there does not exist any gain-seeking product, i.e., $N_g = \emptyset$, the myopic price path $\{\mathbf{p}^m(\mathbf{r}^t)\}_{t=0}^\infty$ always converges to a steady state, whose value may depend on the initial reference price $\mathbf{r}^0 \in \mathcal{P}^n$.*
2. *If there exists any gain-seeking product, i.e., $N_g \neq \emptyset$, the myopic price path $\{\mathbf{p}^m(\mathbf{r}^t)\}_{t=0}^\infty$ admits no steady state regardless of the initial reference price $\mathbf{r}^0 \in \mathcal{P}^n$.*

We defer the proof of Theorem 3 to Appendix B.4. Figure 3 illustrates the convergence behavior described in Theorem 3, where we plot the long-run myopic price paths of three substitutes initialized at $\mathbf{r}^0 = (r_1^0, r_2^0, r_3^0) = (3, 2, 1)$ with $\alpha = 0.8$. As shown in Figure 3a, the myopic policy converges to a constant if no product is associated with gain-seeking reference effects. In contrast, as shown in Figure 3b, once the set contains a gain-seeking product, the myopic price path becomes cyclic.

Figure 3 Long-Run Behaviors of Myopic Price Paths Corresponding to the Two Cases in Theorem 3



Several insights can be drawn from Theorem 3. In the first case where the myopic policy converges, if the choice set contains any product with the loss-averse reference effect, the myopic steady

state is not unique and depends on the initial reference prices. The reason is that for a product with an asymmetric reference effect, different initial reference prices will give rise to different effective reference price sensitivities. In the second case of Theorem 3, the cyclic path originates from the discontinuity of the myopic pricing policy for gain-seeking products (see Eq. (16)), where it has two candidate myopic prices and results in the higher one if its reference price is below the threshold. As the threshold is between the two candidate myopic prices, the price path oscillates when it approaches the threshold. Further, a larger gap between c_i^+ and c_i^- enhances the intensity of such price variations, i.e., a more significant gain-seeking effect contributes to a wider price fluctuation.

In the literature, Popescu and Wu (2007) and Hu et al. (2016) also study the long-run behaviors of the myopic policy, focusing exclusively on single-product frameworks. Despite the difference in demand models, the result of Theorem 3 in the loss-averse and loss-neutral scenarios coincides with Proposition 1 and Proposition 5 in Popescu and Wu (2007), respectively. Besides, for the single-product setting, it is straightforward to show that the convergence of the myopic price path in our work also enjoys the monotone property exhibited in Popescu and Wu (2007). In the gain-seeking scenario, our analysis suggests that the myopic price path always admits a cyclic pattern. This differs from the finding in Hu et al. (2016, Proposition 1), where certain conditions need to be met in their linear demand for the myopic price path to be cyclic.

Moreover, in a special case that only considers symmetric reference effects, i.e., all products belong to the loss-neutral subset, we study the steady state and the convergence pattern of the myopic pricing policy, where the proof of Corollary 2 can be found in Appendix B.5.

COROLLARY 2 (Uniqueness of Myopic Steady State with Loss-Neutrality). *In the loss-neutral scenario, i.e., $N = N_n$, regardless of the initial reference price $\mathbf{r}^0 \in \mathcal{P}^n$, the myopic policy admits a unique steady state, denoted as $\lim_{t \rightarrow \infty} \mathbf{p}^m(\mathbf{r}^t) =: \mathbf{p}^{m*} = (p_1^{m*}, p_2^{m*}, \dots, p_n^{m*})$, such that*

$$p_i^{m*} = \Pi^{m*} + \frac{1}{b_i + c_i}, \quad \forall i \in N, \quad (17)$$

where $c_i := c_i^+ = c_i^-$, $\forall i \in N$, and $\Pi^{m*} := \lim_{t \rightarrow \infty} \Pi^m(\mathbf{r}^t)$ is the unique solution to the equation

$$\Pi = \sum_{i \in N} \frac{1}{b_i + c_i} \cdot \exp \left(a_i - b_i \Pi - \frac{b_i}{b_i + c_i} \right). \quad (18)$$

Furthermore, the convergence patterns for all products are synchronous and must belong to one of the following three cases:

1. $\{p_i^m(\mathbf{r}^t)\}_{t=0}^\infty$ is monotone increasing for every $i \in N$;
2. $\{p_i^m(\mathbf{r}^t)\}_{t=0}^\infty$ is a monotone decreasing sequence for every $i \in N$;
3. $\{p_i^m(\mathbf{r}^t)\}_{t=0}^\infty$ first monotonically decreases and then monotonically increases for every $i \in N$.

Since Eq. (18) can be viewed as a special case of Eq. (13), the solution of Π^{m*} can also be obtained via a simple bisection search on the admissible price range. The intuition behind the unique steady state in the loss-neutral case lies in the fact that sensitivities to the reference price are insusceptible to the initial reference price. Moreover, the convergence pattern of the myopic price path shows that it has at most one nadir and no peak.

4. Optimal Dynamic Pricing

In this section, we study the dynamic optimization problem defined in Eq. (5). To the best of our knowledge, in a single-product discrete-time framework, an explicit solution to the optimal policy remains elusive. The relevant research on single-product scenarios with logit demand has been conducted by Jiang et al. (2022), where they concentrate on the computational aspect, implementing the standard policy iteration algorithm. Nevertheless, this approach becomes intractable when extended to the multi-product scenario due to the curse of dimensionality, i.e., exponential growth of the state space. Other scalable alternatives, such as policy gradient and fitted Q-learning algorithms, often lack theoretical guarantees on the optimality gap due to the commonly used neural networks for function approximations. Given the challenges, our study of the optimal policy mainly focuses on its structural properties, including the performance guarantee and long-run behaviors.

4.1. Optimality Gap

By definitions in Eq. (5) and Eq. (8), it always holds that $V^m(\mathbf{r}^0) \leq V^*(\mathbf{r}^0)$ for any $\mathbf{r}^0 \in \mathcal{P}^n$. Furthermore, according to Theorem 1, computing $V^m(\mathbf{r}^0)$ to an ϵ precision only requires $\mathcal{O}(\log^2(1/\epsilon))$ number of floating-point operations. Thus, it is natural to use the myopic value function as a performance lower bound for the optimal policy with the optimality gap formalized in Proposition 2. The proof is documented in Appendix C.1.

PROPOSITION 2 (Optimality Gap). *Let V^* be the optimal value function defined in Eq. (6) and V^m be the myopic value function defined in Eq. (8). Then, for any initial reference price $\mathbf{r}^0 \in \mathcal{P}^n$, it holds that*

$$0 \leq V^*(\mathbf{r}^0) - V^m(\mathbf{r}^0) \leq \frac{\beta(1-\alpha)}{(1-\alpha\beta)(1-\beta)} \cdot \omega p_H, \quad (19)$$

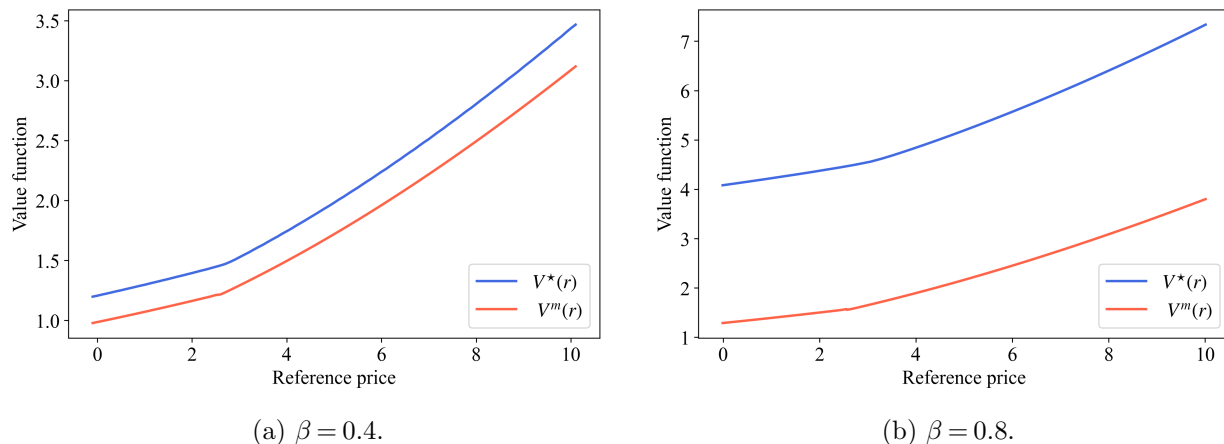
where $\omega := \max_{i \in N} \left\{ \frac{c_i^+}{b_i + c_i^+}, \frac{c_i^-}{b_i + c_i^-} \right\}$.

Proposition 2 provides a uniform performance bound for the myopic pricing policy across all initial reference prices. This bound is tight and shrinks to zero in the three extreme cases: (i) $\alpha = 1$, (ii) $\beta = 0$, or (iii) $\omega = 0$. Firstly, a larger α indicates that the reference price relies more on historical prices and is less likely to evolve over time. When $\alpha = 1$, the reference price remains unchanged throughout the entire planning horizon. Then, the dynamic price optimization in Eq. (5) simplifies

to a series of identical single-period pricing problems, rendering the myopic pricing policy optimal. As pointed out by den Boer and Keskin (2021), this case of a fixed reference price (i.e., $\alpha = 1$) could occur in real-world situations when a retailer has maintained an incumbent price over a prolonged period, and customers have settled an ingrained reference price that is immune to new prices. Secondly, a reduction in β places a stronger emphasis on short-term revenues and distributes less weight to the future. Hence, the myopic policy yields a better performance as β decreases, and it coincides with the optimal policy in the case of $\beta = 0$, under which the long-term cumulative revenue reduces to the revenue of the initial period. We note that a low discount factor β might be observed in practice when firms face high opportunity costs, operate in volatile markets (Bansal et al. 2014), or sell time-sensitive goods that have minimal salvage values, such as perishables (Elmaghraby and Keskinocak 2003). Lastly, the parameter ω captures the relative sensitivity between reference price and price, where a smaller ω indicates a relatively weaker reference effect. When ω approaches zero, the intertemporal nature of the dynamic pricing problem becomes less pronounced, thus narrowing the performance gap between the optimal and myopic pricing policies.

We also note that the optimality gap depends linearly on the price ceiling p_H . The underlying reason can be traced back to Proposition 1, which implies that while the performance of the myopic pricing policy stays unaffected regardless of p_H , the optimal value function follows a non-decreasing trend with respect to p_H . This phenomenon can be further understood by noting that when $p_H \geq \hat{p}_H$ is assumed (see Eq. (9) for the definition of $\hat{p}_H$), p_H becomes a non-binding constraint for the single-period optimization problem, implying that a larger p_H would not improve the myopic policy. Conversely, for the optimal policy, an increase in the price ceiling p_H creates additional space for revenue enhancement, making the optimal policy outperform the myopic policy even further.

Figure 4 Value Functions of Optimal and Myopic Pricing Policies with Parameters from Briesch et al. (1997): $(a, b, c^+, c^-) = (0.32, 0.39, 0.27, 0.13)$, $\alpha = 0.58$, and $p_H = 10$



In Figure 4, we assess the performance of the myopic pricing policy using realistic model specifications from an empirical study by Briesch et al. (1997) (for illustrative simplicity, the example uses a single product, as plotting a multi-product value function requires at least three dimensions). We adopt their estimated sensitivity and memory parameters, which are calibrated from retailing datasets for ground coffee. We set $p_H = 10$, a reasonable price ceiling for coffee back in 1997. Given that Briesch et al. (1997) do not provide a discount factor, we experiment with two values $\beta = 0.4$ and 0.8 . The performance gap between the myopic and optimal policies in Figure 4 aligns with the theoretical result from Proposition 2. Specifically, with a small discount factor $\beta = 0.4$, the myopic value function is merely 10% – 18% lower than that of its optimal counterpart. Whereas an increase in β to 0.8 results in the myopic policy falling 48% – 68% below the optimality.

4.2. Long-Run Behavioral Characterizations

Given the computational burden of solving the optimal policy, exploring its structural characterizations becomes a meaningful endeavor. Regrettably, deriving a complete structure for the optimal policy is also intractable. Even with the single-product linear demand model, the only known structural characterization is found by Hu et al. (2016) in a special gain-seeking scenario, where $\alpha = 0$ and $c^- = 0$. Yet, their analytic techniques do not readily generalize to the logit demand, let alone to the multi-product framework. In Appendix A, we further show the intractability by illustrating how a slight perturbation in parameters can dramatically alter the dynamics of optimal policy.

Hence, in this work, we primarily focus on the long-term dynamics of the optimal pricing policy, i.e., under what circumstances the optimal policy admits a steady state or exhibits a cyclic pattern. This topic is of particular interest since market stability—the consistency of both price and demand over time—significantly influences upstream operations in supply chain management, as stable and fluctuating markets require distinct production and inventory control strategies (Huang et al. 2002). Precisely, a stable market is highly desirable since it creates favorable conditions for implementing efficient planning strategies (Caves and Porter 1978). Conversely, fluctuating demands necessitate more carefully designed strategies for effective logistics management (see, e.g., Song and Zipkin (1993), Aviv and Federgruen (2001), Treharne and Sox (2002), Liu et al. (2022)). We start our discussion by formally defining the steady state for the optimal pricing policy in Definition 4.

DEFINITION 4 (OPTIMAL STEADY STATE). The optimal steady state exists when the optimal pricing policy has a fixed point, i.e., there exists $\mathbf{p}^{**}$ such that $\mathbf{p}^*(\mathbf{p}^{**}) = \mathbf{p}^{**}$.

Revisiting Proposition 1, we reaffirm the necessity of imposing a price ceiling p_H for the optimal policy. As interpreted in Section 2.4, a higher p_H tends to induce a cyclic optimal price pattern in the long run. We formalize this intuition in the next proposition.

PROPOSITION 3. *Suppose that $\alpha < 1$, $\beta > 0$, and $\max_{i \in N} \{c_i^+\} > 0$. Then, regardless of the types of reference effects, the optimal pricing policy admits no steady state if the price ceiling p_H is sufficiently large.*

In the proof of Proposition 3 (see Appendix C.2), we provide a sufficient condition that characterizes the size of p_H such that no steady state exists (see Eq. (C.13)). By Proposition 3, to retain the practical significance of the problem, we avoid setting p_H to be excessively large. This setup of the price ceiling aligns with common practice in the reference effects literature (see, e.g., Popescu and Wu (2007), Hu (2015), Chen et al. (2017), Chen and Nasiry (2020), Jiang et al. (2022)).

The rest of this section examines the long-run behaviors under different types of reference effects. We remark that the previous works commonly assume all products share identical sensitivity parameters, which naturally results in having the same type of reference effects (Briesch et al. 1997, Wang 2018). Our paper, however, is the first theoretical work that takes a more general perspective, where we not only allow different degrees of reference effects through product-differentiated sensitivity parameters but also cover all potential combinations of reference effect types among products. Specifically, in the loss-averse scenario, we only require at least one product to exhibit the loss-averse reference effect, while the rest can display either loss-averse or loss-neutral reference effects. For the gain-seeking scenario, it suffices to have just one gain-seeking product, regardless of the reference effects for all remaining products.

THEOREM 4 (Non-existence of Steady State with Gain-Seekingness). *If there exists any gain-seeking product, i.e., $N_g \neq \emptyset$, the optimal pricing policy admits no steady state.*

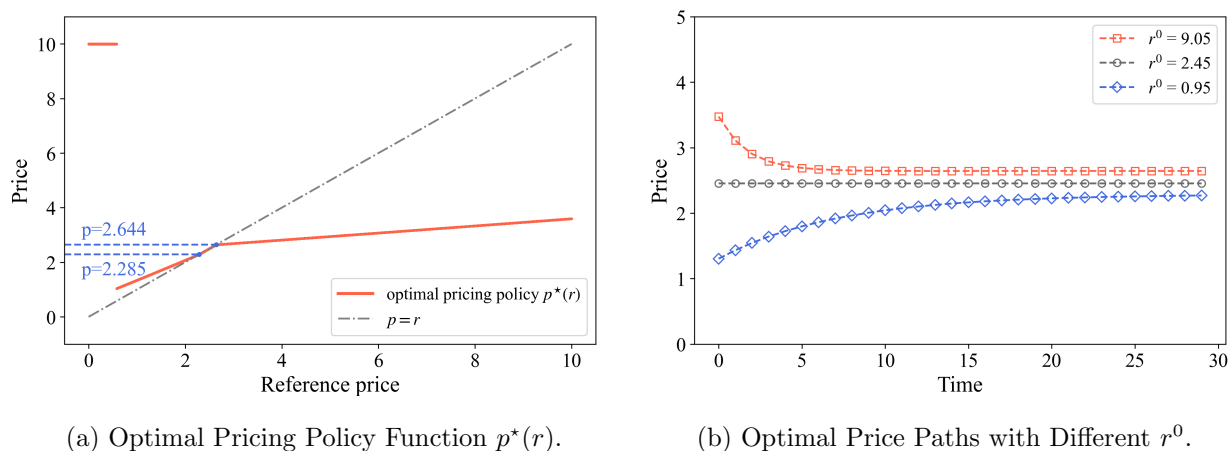
The comparison between Theorem 3 and Theorem 4 yields an interesting observation: with a gain-seeking product, both the myopic and optimal policies are cyclic in the long run. The rationale behind the superiority of cyclic pricing policies in the gain-seeking context arises from consumers' higher sensitivities toward gains. When prices for a product $i \in N$ oscillate between high and low, its effective reference price sensitivity fluctuates accordingly, i.e., being c_i^- during price surges and switching to c_i^+ during price dips. With $c_i^+ > c_i^-$ for a gain-seeking product i , the temporary adverse impact of a high price in one period (thereby raising the reference price) is counterbalanced by the subsequent rise in demand and revenue when prices are lowered. This principle is also employed in the proof of Theorem 4 (see Appendix C.3), where we show that constant prices are suboptimal since a slight price perturbation can always lead to a higher cumulative revenue.

Theorem 4 is consistent with the classical understanding in single-product frameworks: gain-seekingness precludes the optimality of a static policy and favors cyclic price patterns (Popescu and Wu 2007, Hu et al. 2016). However, previous findings do not readily extend to the logit demand, even in the single-product setting, as our model violates the technical assumptions (Popescu and Wu

2007, Assumption 1) or the linearity of demand (Hu et al. 2016) required by their analyses. It is also noteworthy that our findings in the multi-product setting adapt to more comprehensive scenarios, where only one product needs to be gain-seeking while the others can exhibit any reference effects.

Although Theorem 4 confirms the cyclic nature of the optimal policy in the presence of any gain-seeking product, providing a general characterization for the cyclic pattern proves to be quite complicated, even in the single-product setting with linear demand (Hu et al. 2016). It is observed that a minor modification in memory parameter α can lead to a dramatic change in the policy, a phenomenon called *border collision bifurcation* (Jain and Banerjee 2003). By contrast, without gain-seeking products, an optimal steady state can possibly exist. We illustrate the occurrence of steady states with loss-averse consumers in Figure 5, where a single product is employed for clarity since depicting the policy for multiple products would require a plot of four or more dimensions.

Figure 5 Occurrence of Optimal Steady States in Loss-Averse Scenario with Parameters $(a, b, c^+, c^-) = (10, 3, 0.5, 10)$, $p_H = 10$, $\alpha = 0.5$, and $\beta = 0.4$



In this example, we numerically compute the optimal pricing policy by implementing the policy iteration algorithm on the discretized price space, where the feasible price range $\mathcal{P} = [0, 10]$ and the discretization interval is 0.001. We observe that the optimal policy (see Figure 5a) overlaps with the dashed line $p = r$ in the interval $[2.644, 2.285]$, indicating that $p^*(r) = r$ for $r \in [2.644, 2.285]$. Consequently, if the initial reference price falls within this interval, the optimal policy instantaneously reaches a steady state, as depicted by the static price path displayed in Figure 5b. On the other hand, the downward and upward price paths in Figure 5b demonstrate situations where the reference price is initialized outside the interval of $p^*(r) = r$. In such convergent cases, we are able to provide a characterization of optimal steady states.

THEOREM 5 (Steady State Characterization). *In the loss-neutral and loss-averse scenarios, i.e., $N = N_l \cup N_n$, if there exists an optimal steady state, the price $\mathbf{p}^{**} = (p_1^{**}, p_2^{**}, \dots, p_n^{**})$ and its associated single-period revenue $\Pi(\mathbf{p}^{**}, \mathbf{p}^{**})$ at the steady state admit the following structure*

$$p_i^{**} - \Pi(\mathbf{p}^{**}, \mathbf{p}^{**}) \in \left[\frac{1}{b_i + c_i^- \kappa}, \frac{1}{b_i + c_i^+ \kappa} \right], \quad \forall i \in N, \quad (20)$$

where $\kappa := (1 - \beta)/(1 - \alpha\beta)$.

We remark that the relation described in Eq. (20) indicates that optimal steady states can be non-unique. Indeed, this relation is tight such that in certain cases, all solutions to Eq. (20) are steady states. One example is from Figure 5, where it is straightforward to verify that values of $p^{**} - \Pi(p^{**}, p^{**})$ at the two endpoints (i.e., $p^{**} = 2.285$ and 2.644) match the lower and upper bounds of the interval in Eq. (20), respectively. Building on Theorem 5, Corollary 3 explicitly depicts the range of optimal steady states, with its proof deferred to Appendix C.4.

COROLLARY 3. *If there exists an optimal steady state, it holds that $\Pi(\mathbf{p}^{**}, \mathbf{p}^{**}) \in [\Pi^L, \Pi^U]$ and*

$$p_i^{**} \in \left[\Pi^L + \frac{1}{b_i + c_i^- \kappa}, \Pi^U + \frac{1}{b_i + c_i^+ \kappa} \right], \quad \forall i \in N, \quad (21)$$

where Π^L and Π^U are the unique solutions to the following Eq. (22a) and Eq. (22b), respectively, and both of them can be computed through a one-dimensional bisection method

$$\Pi = \sum_{i \in N} \frac{1}{b_i + c_i^- \kappa} \cdot \exp \left(a_i - b_i \Pi - \frac{b_i}{b_i + c_i^+ \kappa} \right), \quad (22a)$$

$$\Pi = \sum_{i \in N} \frac{1}{b_i + c_i^+ \kappa} \cdot \exp \left(a_i - b_i \Pi - \frac{b_i}{b_i + c_i^- \kappa} \right). \quad (22b)$$

Corollary 3 allows a direct derivation of the lower and upper bounds for optimal steady states $\mathbf{p}^{**}$, using a bisection method similar to Corollary 2. Prices falling outside this interval in Eq. (21) are thereby disqualified as potential optimal steady states. Moreover, the corollary stipulates a uniform ceiling for the optimal value function across all steady states, i.e., $V(\mathbf{p}^{**}) \leq \Pi^U/(1 - \beta)$ for all $\mathbf{p}^{**}$. Hence, if a cyclic price pattern generates a cumulative revenue exceeding $\Pi^U/(1 - \beta)$ from any initial reference price $\mathbf{r}^0$, such an initial state can never be an optimal steady state.

While the optimal steady states can be non-unique with loss-averse products, pure loss-neutrality guarantees the uniqueness of the steady state. A closer examination of Theorem 5 shows that, when all products have loss-neutral reference effects, we have $c_i^+ = c_i^-$, $\forall i \in N$, and thereby the interval in Eq.(21) collapses to a single value. Furthermore, this value of $\mathbf{p}^{**}$ can be obtained by a one-dimensional bisection method, requiring only $\mathcal{O}(\log(1/\epsilon))$ number of floating-point operations to find $\mathbf{p}^{**}$ within ϵ accuracy.

COROLLARY 4 (Uniqueness of Steady State with Loss-Neutrality). *In the loss-neutral scenario, i.e., $N = N_n$, if an optimal steady state exists, it is unique and satisfies that*

$$p_i^{**} = \Pi^{**} + \frac{1}{b_i + c_i \kappa}, \quad \forall i \in N, \quad (23)$$

where $c_i := c_i^+ = c_i^-$, $\forall i \in N$, and $\Pi^{**} := \Pi(\mathbf{p}^{**}, \mathbf{p}^{**})$ is the unique solution to the equation

$$\Pi = \sum_{i \in N} \frac{1}{b_i + c_i \kappa} \cdot \exp \left(a_i - b_i \Pi - \frac{b_i}{b_i + c_i \kappa} \right). \quad (24)$$

The comparison of Corollary 4 and Corollary 2 reveals that structures of steady states in optimal and myopic policies are surprisingly similar, composed of their corresponding single-period revenues plus product-specific constants. Notably, the constants for the optimal steady state additionally depend on the memory parameter α and the discount factor β via the multiplier $\kappa \in (0, 1]$ in the denominator. This intuitively aligns with the fact that the optimal policy needs to account for the intertemporal nature of the problem and aims at maximizing the cumulative revenue.

While the optimal steady state can be characterized when it exists, we comment that providing a general sufficient condition for its existence remains a great challenge. In current literature, such conditions are only established in the single-product setting, where the demand function needs to meet certain technical assumptions and reference effects are incorporated at an aggregate level (Popescu and Wu 2007). For the multi-product logit model in our paper, the problem is no longer tractable for the following reasons. Firstly, Hu and Nasiry (2018) demonstrate that long-term dynamics are directly associated with the type of reference effects at the aggregate level rather than the individual level. However, in the multi-product framework where each product can be priced differently, the concept of aggregate-level or global reference effects lacks a proper generalization (see the discussion in Appendix A.1.2), making the existing approach based on the aggregate level inapplicable. Secondly, even for single-product logit demand models, no condition for steady states has been established (Jiang et al. 2022). This can be attributed to inconsistencies in the responsiveness to losses/gains at the individual-level and aggregate-level demands (Hu and Nasiry 2018), e.g., in the logit model, aggregate demand may be more responsive to gains even when all consumers exhibit loss-aversion individually. Unfortunately, Jiang et al. (2022) further show that for the MNL model, there does not exist a clear mapping from individual-level reference effects to their aggregate-level counterparts. Finally, without good properties on the responsiveness to losses/gains at the aggregate level, confirming the existence of the steady state from definition requires showing that a higher revenue is not achievable under any cyclic policy. This proves to be a daunting task, given the infinitely many possible cyclic patterns.

4.2.1. A Special Case: Promotion-Insensitive Consumers. Despite the difficulties in fully characterizing the steady-state existence, we demonstrate that, for a specific loss-averse case where $c_i^- \geq c_i^+ = 0$ for all $i \in N$, the optimal policy always admits a steady state. This seemingly restrictive scenario indeed offers a good approximation to markets with promotion-insensitive consumers. For instance, in the luxury industry, particularly with high-end watches, jewelry, or designer handbags, consumers tend to be extremely loss-averse and are less responsive to promotions. They may perceive a discount or promotion as a potential sign of declining prestige in the brand, which in turn reinforces their loss aversion (Kapferer and Bastien 2012, Wiedmann and Hennigs 2012). We note that this extreme reference-type assumption is not unique to our work. Hu et al. (2016) adopt a similar assumption with gain-seeking consumers in the single-product setting, presuming these individuals are insensitive to losses ($c^+ \geq c^- = 0$).

With these promotion-insensitive consumers, Theorem 6 suggests that constant pricing is always optimal in the long run regardless of parameters and initial reference prices (see its proof in Appendix C.5).

THEOREM 6. *Suppose the consumers are insensitive to gain effect, i.e., $c_i^- \geq c_i^+ = 0, \forall i \in N$, then the optimal pricing policy always admits a steady state, i.e., $\exists \mathbf{p}^{**} \in \mathcal{P}^n$ such that $\mathbf{p}^*(\mathbf{p}^{**}) = \mathbf{p}^{**}$.*

The classic results suggest that cyclic pricing antagonizes loss-averse consumers (Fibich et al. 2003, Popescu and Wu 2007, Nasiry and Popescu 2011). Nevertheless, recent findings by Hu and Nasiry (2018) challenge such beliefs, arguing that the loss-aversion does not necessarily preclude price variations. The key factor lies in the heterogeneity of consumer valuations. According to Hu and Nasiry (2018), the loss-aversion at an individual level is inherited by the aggregated-level demand function (see definition in Eq. (A.1)) if and only if consumer valuations are distributed uniformly. As a result, in the logit model where consumer valuations follow a logistic distribution, the aggregate demand may be more responsive to gains, despite individual loss-aversion. This possibility paves the way for the optimal policy to exhibit a cyclic pattern even though all consumers are loss-averse, as demonstrated by Jiang et al. (2022, Proposition 5) in the single-product setting. By contrast, under this special case where consumers are entirely insensitive to promotions, we conclude in Theorem 6 that an optimal steady state always exists. This finding extends beyond existing results, as the general guarantee of a constant policy being optimal is limited to situations where the aggregate demand is globally more sensitive to losses; however, in our case, the aggregate demand can still be more sensitive to gains in certain regions (Jiang et al. 2022, Proposition 2). The core of our proof relies on a critical observation: once consumers are insensitive to promotions, there constantly exists a fixed point $\mathbf{p}^{**}$ for the myopic pricing policy, i.e., $\mathbf{p}^m(\mathbf{p}^{**}) = \mathbf{p}^{**}$, that achieves the highest possible single-period revenue, i.e., $(\mathbf{p}^{**}, \mathbf{p}^{**}) \in \arg \max_{\mathbf{p}, \mathbf{r} \in \mathcal{P}^n} \Pi(\mathbf{p}, \mathbf{r})$. As a result, when the reference price $\mathbf{r}$ approaches $\mathbf{p}^{**}$, setting price $\mathbf{p}^{**}$ is optimal in both short-term and long-term perspectives (see Eq. (C.28)), which confirms $\mathbf{p}^{**}$ to be an optimal steady state.

4.3. Multi-Product Framework Facilitates Cyclic Pricing

In practice, a constant pricing policy over the long term (i.e., the convergence to an optimal steady state) is not frequently observed. This section sheds light on how our multi-product framework can partly explain this observation. We demonstrate that, in contrast to the single-product model where substitutes are priced separately, the optimal policy derived from our multi-product model is more likely to be cyclic in the long run. Especially, our theoretical finding is based on a loss-neutral scenario with $n = |N| \geq 2$ perfect substitutes, i.e., products having the same parameters $a = a_i, b = b_i$ and $c = c_i$ for all $i \in N$. This setting effectively approximates real-world scenarios where products are interchangeable in use and have similar intrinsic values and responsiveness parameters. The examples include clothing in the same style but different colors, different brands of gasoline, and daily necessities from various brands, among others.

We compare the long-run dynamics of the optimal policy under both multi-product and single-product frameworks. Notably, the single-product model essentially implies pricing products separately, and in the context of perfect substitutes, is equivalent to making a single price decision, i.e., setting the same price for all products. We use $p^s(r)$ to denote the optimal policy with a single price decision, named the *single-product-based policy*, which can be expressed by the following optimization problem

$$p^s = \arg \max_{p^\pi: \mathcal{P} \rightarrow \mathcal{P}} \sum_{t=0}^{\infty} \beta^t \Pi(p^\pi(r^t) \cdot \mathbf{1}, r^t \cdot \mathbf{1}), \quad \text{s.t. } r^{t+1} = \alpha r^t + (1 - \alpha) \cdot p^\pi(r^t), \quad \forall t \geq 0, \quad (25)$$

where $\mathbf{1} := (1, 1, \dots, 1)$ denotes the vector of ones. We note that the single-product-based policy always yields uniform pricing, whereas the multi-product optimal policy (as defined in Eq. (5)) permits differentiated prices among products, even when they have identical reference prices.

THEOREM 7. *Suppose consumers are loss-neutral and have the same parameters toward all products $i \in N$ with $n = |N| \geq 2$. Then, the optimal pricing policy $\mathbf{p}^*(\cdot)$ is more likely to be cyclic than the single-product-based policy $p^s(\cdot)$ in the following senses:*

1. *If the single-product-based policy is cyclic in the long run, the optimal policy is also cyclic.*
2. *When the single-product-based policy admits a steady state, the optimal policy can still be cyclic. Particularly, a sufficient condition is given by Eq. (26), where Eq. (26a) ensures the constant pricing of single-product-based policy, and Eq. (26b) ensures the cyclic pricing of optimal policy.*

$$p_H [1 - \zeta_1 \exp(-(b+c) \cdot p_H)] - \zeta_1 \zeta_2 \exp(-(b+c) \cdot p_H) \leq \zeta_2, \quad (26a)$$

$$[(1-\kappa)^2 + \beta(1-\alpha\kappa)^2]c^2 - (1+\beta\alpha^2) \cdot (b+c\kappa)^2 \geq \frac{(b+c)^2 + \beta(\alpha b+c)^2}{n-1}, \quad (26b)$$

where $\zeta_1 := \exp(a + \log n + cp^{m*})$, $\zeta_2 := [2\alpha b + c(\alpha + 1)] / [(ab + c)(b + c)]$, $\kappa := (1 - \beta) / (1 - \alpha\beta)$, and p^{m*} is the common myopic steady state for all products.

Below, we comment on Theorem 7, with its formal proof found in Appendix C.6. The first condition Eq. (26a) is equivalent to having an upper bound on p_H with its value determined by all other parameters: if $1 - \zeta_1 \exp(-(b+c) \cdot p_H) \leq 0$, then Eq. (26a) always holds; otherwise, the left-hand side of Eq. (26a) is monotone increasing in p_H . This aligns with previous findings in Proposition 3, where we have shown that an overly high price ceiling must lead to cyclic long-run behavior. Moreover, it is easy to verify that this upper bound on p_H is consistent with our assumption for p_H in Section 2.4, i.e., p_H is large enough so that it is an inactive constraint for the myopic optimization problem for any $\mathbf{r} \in \mathcal{P}^n$. The second condition Eq. (26b) implies that the optimal policy is more likely to be cyclic with strong reference effects (i.e. large c). The underlying intuition is that stronger reference effects enable the cyclic policy to achieve higher revenue through price variations but have little influence on the static pricing policy.

COROLLARY 5 (Asymptotic Analysis of Eq. (26)). *When the number of products $n \rightarrow \infty$, the single-product-based policy always admits a steady state, whereas the optimal policy is cyclic if*

$$1 - 2 \cdot \frac{1 - \beta}{1 - \alpha\beta} \geq \frac{b}{c}. \quad (27)$$

In the asymptotic case, Eq. (26a) naturally holds as $n \rightarrow \infty$, and Eq. (26b) can be reduced to Eq. (27). This reduction reaffirms that stronger reference effects (i.e., large c) facilitate cyclic pricing of the optimal policy. Further, the left-hand side of Eq. (27) suggests that cyclic policies are preferred when consumers have short-term memories (i.e., small α), and the retailer attaches importance to future revenues (i.e., large β). Meanwhile, according to Proposition 2, we observe that the above trends on c, α , and β would widen the performance gap between the myopic and optimal policies. This is indeed expected since the myopic policy, which always converges in the long run (see Theorem 3), exhibits a larger optimality gap when the optimal is actually cyclic.

REMARK 3. If both the single-product-based policy and the optimal policy admit steady states, their steady states must be the same. This can be observed from Corollary 4, where we show that the prices for all products are the same, i.e., uniform, at the optimal steady state.

While pricing multiple substitutes with the single-product-based policy is obviously not optimal, Theorem 7 further reveals that such a pricing strategy is more likely to result in a steady state. The underlying reason is that according to Remark 3, the steady state revenues are the same for both optimal and single-product-based policies. Yet, when being cyclic, the single-product-based policy is limited to synchronous cyclic patterns, whereas the true optimal policy under the multi-product framework allows arbitrary cyclic patterns, which offers greater potential for revenue growth when transitioning from a constant to a cyclic pattern. In fact, we demonstrate in Section 5.2 that asynchronous cyclic patterns generally outperform synchronous ones. The advantage of an

asynchronous cycle lies in its ability to maintain a high purchase probability consistently, unlike its synchronous counterpart, which tends to lose revenue during high-price periods. Hence, compared to the single-product-based policy, the optimal policy is prone to exhibit cyclic behaviors. Given this discrepancy in long-run dynamics, retailers should be cautious about using the single-product-based policy for pricing multiple similar substitutes. Such an approach could misguide the understanding of the long-run pricing behavior, ending up with more constant policies than it should be.

REMARK 4. For products with possibly different parameters, a more general sufficient condition for the cyclic long-run behavior is given by

$$\exists i \in N, \text{ such that } \kappa_{i,1} - \kappa_{i,2} > (2\kappa_{i,1} - \kappa_{i,2}) \cdot q_i(\mathbf{p}^{**}, \mathbf{p}^{**}), \quad (28)$$

where $\kappa_{i,1} := (b_i + c_i)^2 + \beta(\alpha b_i + c_i)^2$, $\kappa_{i,2} := 2(b_i + c_i\kappa) \cdot [b_i(1 + \alpha^2\beta) + c_i(1 + \alpha\beta)]$, and $q_i(\mathbf{p}, \mathbf{r})$ is the market share of product i at $(\mathbf{p}, \mathbf{r})$. When all n products have the same parameters, we must have $q_i(\mathbf{p}^{**}, \mathbf{p}^{**}) < 1/n$ for all $i \in N$. With this relation, it is straightforward to derive Eq. (26b) from Eq. (28). In Appendix C.6, we prove this remark together with Theorem 7.

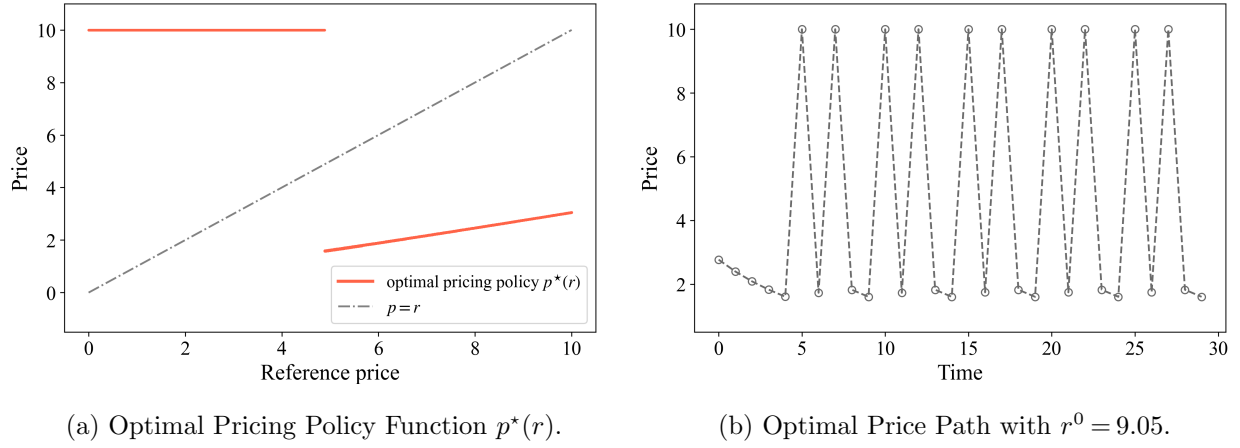
5. Numerical Studies

This section presents numerical experiments illustrating key features of the optimal pricing policy, with the theoretical justifications in Appendices D.1 and D.2. For all experiments, we apply the policy iteration on the price space $\mathcal{P} = [0, 10]$ discretized at 0.01 intervals, with the memory parameter $\alpha = 0.8$ and the discount factor $\beta = 0.95$. Due to the computational intractability, examples are carried out on two or fewer products.

5.1. Periodical Peaks at Price Ceiling of Cyclic Optimal Policy

Our experiments consistently show that when the optimal policy is cyclic, the price trajectory periodically hits the price ceiling, as illustrated in Figure 6, where a single product is used for ease of presentation. The optimal policy depicted in Figure 6a indicates that the price reaches the ceiling for $r \in [0, 4.89]$, a pattern further demonstrated in the optimal price path of Figure 6b. While earlier studies on the single-product setting have explored cyclic optimal policies such as cyclic skimming (Hu et al. 2016) and high-low pricing strategies (Popescu and Wu 2007, Hu et al. 2016), they do not examine how these cyclic patterns relate to the likelihood of hitting the price ceiling. Our research is the first to investigate this particular aspect of the cyclic policy.

In Appendix D.1, we offer a theoretical interpretation of this phenomenon using a single-product example. Here, we briefly state the rationale behind it. When the cyclic pattern occurs, there must exist some reference price $\hat{r}$ such that $p^*(\hat{r}) > \hat{r}$. Then, we are left to show that it is highly probable that $p^*(\hat{r}) = p_H$ given that $p^*(\hat{r})$ is reasonably larger than $\hat{r}$. From the Bellman equation $V^*(\hat{r}) = \Pi(p^*(\hat{r}), \hat{r}) + \beta V^*(\alpha\hat{r} + (1 - \alpha)p^*(\hat{r}))$, we find that setting the optimal price $p^*(\hat{r})$ is mainly

Figure 6 Illustration of Periodically Striking Price Ceiling with Parameters $(a, b, c^+, c^-) = (2, 1, 1, 1)$ 

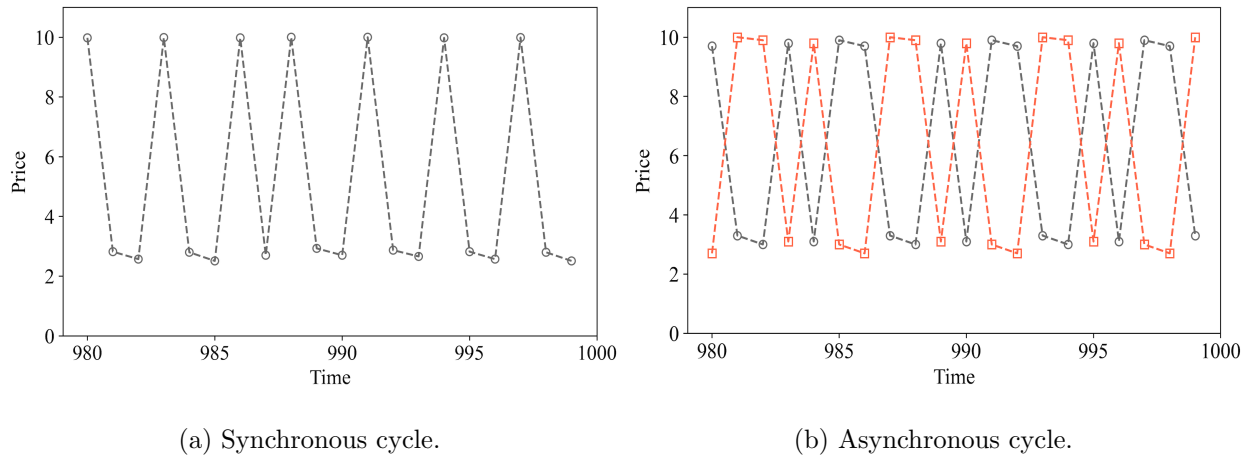
concerned with balancing the short-term revenue $\Pi(p^*(\hat{r}), \hat{r})$ and long-term revenue $V^*(\alpha\hat{r} + (1 - \alpha)p^*(\hat{r}))$: a higher $p^*(\hat{r})$ may decrease the short-term gain, but is beneficial for the long-term cumulative reward. The core idea hinges on the discrepancy in the rates of change between short-term and long-term revenues. On one side, it is straightforward to verify that short-term reward $\Pi(p^*(\hat{r}), \hat{r})$ is exponentially sensitive to the price $p^*(\hat{r})$, as per its definition in Eq. (4). On the other side, by exploiting structural properties of the optimal pricing policy, we prove that long-term reward $V^*(\alpha\hat{r} + (1 - \alpha)p^*(\hat{r}))$ at least grows at a constant rate with respect to $p^*(\hat{r})$. Given that $p^*(\hat{r})$ is reasonably larger than $\hat{r}$, $\Pi(p^*(\hat{r}), \hat{r})$ is already exponentially small and close to zero. Therefore, further increasing the value of $p^*(\hat{r})$ is beneficial for the overall cumulative revenue, as there is minimal scope for $\Pi(p^*(\hat{r}), \hat{r})$ to decrease, while $V^*(\alpha\hat{r} + (1 - \alpha)p^*(\hat{r}))$ can continue rising at a constant rate. This explains why the cyclic pattern tends to reach the price ceiling.

5.2. Asynchronous Fluctuations of Cyclic Optimal Policy

In Section 4.3, we demonstrate that applying the single-product model to price similar substitutes, i.e., the uniform pricing policy defined in Eq. (25), is likely to have different long-run behavior than the optimal policy. Here, we dig further into the reason why the optimal pricing policy outperforms the uniform pricing policy, especially when both policies exhibit cyclic behaviors over the long term. By definition, when being cyclic, the uniform pricing policy always yields *synchronous* fluctuations. Conversely, in our numerical experiments, it is frequently observed that the true optimal policy exhibits *asynchronous* fluctuations (see a two-product example in Figure 7). The intuition behind the superior performance of the asynchronous cycle is that it keeps a consistently high purchase probability in each period, whereas the synchronous counterpart dissipates the revenue during high-price periods. From a managerial perspective, this phenomenon implies that simultaneous discounts on similar substitutes result in suboptimal pricing policies. For further theoretical insight,

an example with two products is provided in Appendix D.2. Finally, we comment that according to our experiments, even for products with different parameters, asynchronous patterns are more likely to occur when the optimal policy appears cyclic.

Figure 7 Long-Run Behaviors of Synchronous and Asynchronous Cycles for Two Identical Products with Parameters $(a_1, b_1, c_1^+, c_1^-) = (a_2, b_2, c_2^+, c_2^-) = (1, 1, 1, 1)$ and Initial Reference Prices $(r_1^0, r_2^0) = (5, 5)$



6. Concluding Remarks

This paper studies a multi-product dynamic pricing problem under reference effects. In the first part of our research, we derive a comprehensive solution for the myopic pricing policy, which advances the current state-of-the-art in single-period price optimization under the MNL demand. Additionally, we conduct a convergence analysis of its long-term behaviors and quantify the optimality gap between the myopic and the optimal policies. We then explore the optimal pricing policy, with the main focus on its long-term behaviors. We demonstrate that in a multi-product framework, the existence of the optimal steady state predominately hinges on the presence of gain-seeking reference effects. Notably, we find that having even a single gain-seeking product would prohibit the optimality of a constant policy under any model specification. In loss-neutral and loss-averse scenarios where the optimal steady state can exist, we provide characterizations and computational methods for the steady state. Then, we demonstrate that the optimal policy always admits the steady state in a special loss-averse scenario where consumers are promotion-insensitive.

Existing OR literature predominantly concentrates on the single-product framework. However, our findings suggest the need for caution when applying the single-product model to separately price multiple substitutes, as this approach could potentially misrepresent the long-term dynamics. Specifically, this single-product-based strategy tends to result in a constant long-run policy, whereas the true underlying optimal pricing (under a multi-product framework) is cyclic. Compared to the single-product setting, the multi-product framework better reflects the real-market conditions, and

the above discrepancy in long-run behaviors further confirms that the multi-product-based policy aligns more closely with the widespread practice of offering periodic discounts in reality.

Beyond the scope of this work, one promising future direction is to incorporate other types of reference formulations, notably the stimulus-based (or external) reference price. This perspective suggests that the price judgment is formed at the point of purchase by using the current external information such as the price of other products, rather than relying on past memory (Lynch Jr and Srull 1982, Biehal and Chakravarti 1983). The stimulus-based reference formulation could be, for example, the average, lowest, and highest price of among all products in the set, or simply the price of some specifically featured product (Ji et al. 2023). Moreover, as consumers interact repeatedly with the retailer, they may learn the cyclic pricing pattern and respond strategically. It is interesting to consider these strategic behaviors and design more robust pricing policies.

Appendix A: Illustration of Optimal Pricing Policy

In the first example (see Figure EC.1), we illustrate the optimal pricing policy under different levels of sensitivities to gains (i.e., c^+). The red lines depict the optimal policies $\mathbf{p}^*(\mathbf{r})$, which map

Figure EC.1 Optimal Pricing Policies Under Different c^+ with Same Parameters $\alpha = 0.8$ and $\beta = 0.95$

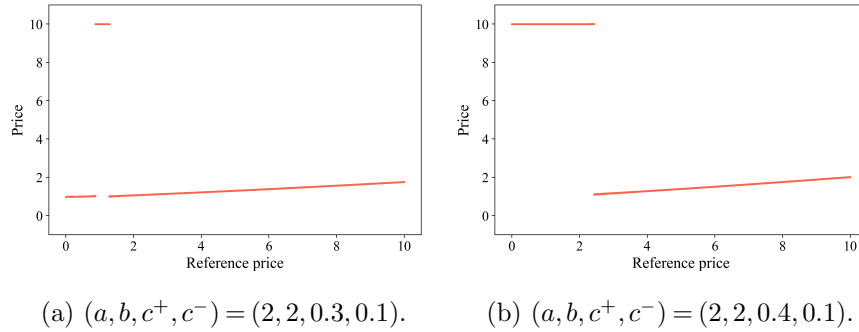
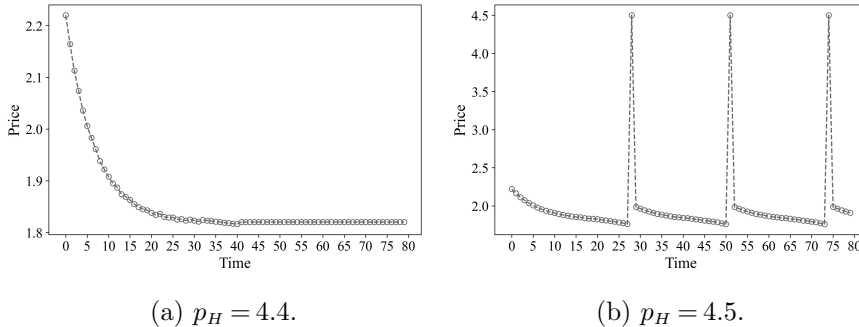


Figure EC.2 Optimal Price Paths Under Different Price Ceiling p_H with Same Parameters $(a, b, c^+, c^-) = (2, 1, 1, 1)$, $\alpha = 0.8$, $\beta = 0.95$, and Same Initial Reference Price $r^0 = 3$



reference price to price. Remarkably, even in this single-product case, a minimal increase of 0.1 in c^+ , while keeping other parameters constant, can induce a significant change in the optimal policy.

The second example illustrates two possible dynamics under different price ceilings (i.e., p_H). The price paths shown in Figure EC.2 demonstrate that merely increasing p_H from 4.4 to 4.5 can cause the long-run behaviors to change drastically. This example further emphasizes the role of the price ceiling and highlights the complexity of explicitly characterizing the optimal pricing policy.

References

- Araman VF, Fayad B (2021) Intertemporal price discrimination with time-varying valuations. *Operations Research* 69(1):245–265.
- Aviv Y, Federgruen A (2001) Capacitated multi-item inventory systems with random and seasonally fluctuating demands: implications for postponement strategies. *Management science* 47(4):512–531.
- Aviv Y, Pazgal A (2008) Optimal pricing of seasonal products in the presence of forward-looking consumers. *Manufacturing & Service Operations Management* 10(3):339–359.
- Bansal R, Kiku D, Shaliastovich I, Yaron A (2014) Volatility, the macroeconomy, and asset prices. *The Journal of Finance* 69(6):2471–2511.
- Bell DR, Lattin JM (2000) Looking for loss aversion in scanner panel data: The confounding effect of price response heterogeneity. *Marketing Science* 19(2):185–200.
- Bertsekas D (2012) *Dynamic programming and optimal control: Volume I*, volume 1 (Athena scientific).
- Besbes O, Lobel I (2015) Intertemporal price discrimination: Structure and computation of optimal policies. *Management Science* 61(1):92–110.
- Biehal G, Chakravarti D (1983) Information accessibility as a moderator of consumer choice. *Journal of Consumer Research* 10(1):1–14.
- Briesch RA, Krishnamurthi L, Mazumdar T, Raj SP (1997) A comparative analysis of reference price models. *Journal of consumer research* 24(2):202–214.
- Caves RE, Porter ME (1978) Market structure, oligopoly, and stability of market shares. *The Journal of Industrial Economics* 289–313.
- Chang K, Siddarth S, Weinberg CB (1999) The impact of heterogeneity in purchase timing and price responsiveness on estimates of sticker shock effects. *Marketing Science* 18(2):178–192.
- Chen N, Nasiry J (2020) Does loss aversion preclude price variation? *Manufacturing & Service Operations Management* 22(2):383–395.
- Chen X, Hu P, Hu Z (2017) Efficient algorithms for the dynamic pricing problem with reference price effect. *Management Science* 63(12):4389–4408.
- Debreu G (1952) A social equilibrium existence theorem. *Proceedings of the National Academy of Sciences* 38(10):886–893.
- den Boer AV, Keskin NB (2021) Dynamic pricing with demand learning and reference effects. *Management Science* Forthcoming.

- Du C, Cooper WL, Wang Z (2016) Optimal pricing for a multinomial logit choice model with network effects. *Operations Research* 64(2):441–455.
- Elmaghraby W, Keskinocak P (2003) Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Management science* 49(10):1287–1309.
- Fibich G, Gavious A, Lowengart O (2003) Explicit solutions of optimization models and differential games with nonsmooth (asymmetric) reference-price effects. *Operations Research* 51(5):721–734.
- Gallego G, Wang R (2014) Multiproduct price optimization and competition under the nested logit model with product-differentiated price sensitivities. *Operations Research* 62(2):450–461.
- Golrezaei N, Jaillet P, Liang JCN (2020) No-regret learning in price competitions under consumer reference effects. *Advances in Neural Information Processing Systems* 33:21416–21427.
- Greenleaf EA (1995) The impact of reference price effects on the profitability of price promotions. *Marketing science* 14(1):82–104.
- Hopp WJ, Xu X (2005) Product line selection and pricing with modularity in design. *Manufacturing & Service Operations Management* 7(3):172–187.
- Howard RA (1960) *Dynamic programming and markov processes*. (John Wiley).
- Hu Z (2015) *Dynamic pricing with reference price effects* (University of Illinois at Urbana-Champaign).
- Hu Z, Chen X, Hu P (2016) Dynamic pricing with gain-seeking reference price effects. *Operations Research* 64(1):150–157.
- Hu Z, Nasiry J (2018) Are markets with loss-averse consumers more sensitive to losses? *Management Science* 64(3):1384–1395.
- Huang SH, Uppal M, Shi J (2002) A product driven approach to manufacturing supply chain selection. *Supply Chain Management: An International Journal* .
- Jain P, Banerjee S (2003) Border-collision bifurcations in one-dimensional discontinuous maps. *International Journal of Bifurcation and Chaos* 13(11):3341–3351.
- Ji S, Yang Y, Shi C (2023) Online learning and pricing for multiple products with reference price effects. *Available at SSRN 4349904* .
- Jiang H, Cao J, Shen ZJM (2022) Intertemporal pricing via nonparametric estimation: Integrating reference effects and consumer heterogeneity. *Manufacturing & Service Operations Management* .
- Kahneman D, Tversky A (1979) Prospect theory: An analysis of decision under risk. *Econometrica* 47(2):263–292.
- Kapferer JN, Bastien V (2012) *The luxury strategy: Break the rules of marketing to build luxury brands* (Kogan page publishers).
- Kopalle PK, Kannan P, Boldt LB, Arora N (2012) The impact of household level heterogeneity in reference price effects on optimal retailer pricing policies. *Journal of Retailing* 88(1):102–114.
- Kopalle PK, Rao AG, Assuncao JL (1996) Asymmetric reference price effects and dynamic pricing policies. *Marketing Science* 15(1):60–85.
- Li H, Huh WT (2011) Pricing multiple products with the multinomial logit and nested logit models: Concavity and implications. *Manufacturing & Service Operations Management* 13(4):549–563.

- Li H, Webster S, Mason N, Kempf K (2019) Product-line pricing under discrete mixed multinomial logit demand. *Manufacturing & Service Operations Management* 21(1):14–28.
- Li H, Webster S, Yu G (2020) Product design under multinomial logit choices: Optimization of quality and prices in an evolving product line. *Manufacturing & Service Operations Management* 22(5):1011–1025.
- Liu H, Wang Y, Lee LH, Chew EP (2022) An approximate dynamic programming approach for production-delivery scheduling under non-stationary demand. *Naval Research Logistics (NRL)* 69(4):511–528.
- Liu Y, Cooper WL (2015) Optimal dynamic pricing with patient customers. *Operations Research* 63(6):1307–1319.
- Lobel I (2020) Dynamic pricing with heterogeneous patience levels. *Operations Research* 68(4):1038–1046.
- Lynch Jr JG, Srull TK (1982) Memory and attentional factors in consumer choice: Concepts and research methods. *Journal of consumer research* 9(1):18–37.
- Mazumdar T, Raj SP, Sinha I (2005) Reference price research: Review and propositions. *Journal of marketing* 69(4):84–102.
- McFadden D (1974) Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics* 105–142.
- Nasiry J, Popescu I (2011) Dynamic pricing with loss-averse consumers and peak-end anchoring. *Operations research* 59(6):1361–1368.
- Popescu I, Wu Y (2007) Dynamic pricing strategies with reference effects. *Operations research* 55(3):413–429.
- Song JS, Zipkin P (1993) Inventory control in a fluctuating demand environment. *Operations Research* 41(2):351–370.
- Su X (2007) Intertemporal pricing with strategic customer behavior. *Management Science* 53(5):726–741.
- Trehan JT, Sox CR (2002) Adaptive inventory control for nonstationary demand and partial information. *Management Science* 48(5):607–624.
- Tversky A (1972) Elimination by aspects: A theory of choice. *Psychological review* 79(4):281.
- Wang R (2012) Capacitated assortment and price optimization under the multinomial logit model. *Operations Research Letters* 40(6):492–497.
- Wang R (2018) When prospect theory meets consumer choice models: Assortment and pricing management with reference prices. *Manufacturing & Service Operations Management* 20(3):583–600.
- Wang Z (2016) Intertemporal price discrimination via reference price effects. *Operations research* 64(2):290–296.
- Wiedmann KP, Hennigs N (2012) *Luxury marketing: A challenge for theory and practice* (Springer).
- Wu S, Liu Q, Zhang RQ (2015) The reference effects on a retailer's dynamic pricing and inventory strategies with strategic consumers. *Operations Research* 63(6):1320–1335.
- Zhang H, Jasin S (2022) Online learning and optimization of (some) cyclic pricing policies in the presence of patient customers. *Manufacturing & Service Operations Management* 24(2):1165–1182.
- Zhang H, Rusmevichientong P, Topaloglu H (2018) Multiproduct pricing under the generalized extreme value models with homogeneous price sensitivity parameters. *Operations Research* 66(6):1559–1570.