

Intertemporal Pricing via Nonparametric Estimation: Integrating Reference Effects and Consumer Heterogeneity

Hansheng Jiang

Department of Industrial Engineering and Operations Research, University of California, Berkeley
hansheng_jiang@berkeley.edu

Junyu Cao

McCombs School of Business, The University of Texas at Austin
junyu.cao@mcombs.utexas.edu

Zuo-Jun Max Shen

Department of Industrial Engineering and Operations Research, University of California, Berkeley
maxshen@berkeley.edu

Problem definition: We consider intertemporal pricing in the presence of reference effects and consumer heterogeneity. Our research question encompasses how to estimate heterogeneous consumer reference effects from data and how to efficiently compute the optimal pricing policy. Understanding reference effects is essential for designing pricing policies in modern retailing. Our work contributes to this area by incorporating consumer heterogeneity under arbitrary distributions. **Methodology/results:** We propose a mixed logit demand model that allows arbitrary joint distributions of valuations, responsiveness to prices, and responsiveness to reference prices among consumers. We use a nonparametric estimation method to learn consumer heterogeneity from transaction data. Further, we formulate the pricing optimization as an infinite horizon dynamic programming problem and solve it by applying a modified policy iteration algorithm. Moreover, we investigate the structure of optimal pricing policies and prove the sub-optimality of constant pricing policies even when all consumers are loss-averse according to the classical definition. Our numerical studies show that our estimation and optimization framework improves the expected revenue of retailers via accounting for heterogeneity. We validate our model using real data from JD.com, a large E-commerce retailer, and find empirical evidence of consumer heterogeneity. **Managerial implications:** In practice, ignoring consumer heterogeneity may lead to a significant loss of revenue. Furthermore, heterogeneous reference effect offers a strong motive for promotions and price fluctuations.

Key words: reference effect, consumer heterogeneity, data-driven, intertemporal pricing, nonparametric estimation, online retailing

1. Introduction

Two prominent questions in pricing have been receiving considerable attention from both academia and industry over the years. The first question is how to estimate consumer demand accurately with

frequently changing prices. At traditional brick-and-mortar stores, predicting consumer behaviors is challenging partially because the collected data are often incomplete, and consumer identification is usually missing in the data. The total sales per store do not directly provide consumer information such as personal preference and browsing history. Nowadays, with the prevalence of online retailing, all sorts of individual consumer shopping activities, including clicks and purchases along with the timestamps of these activities, are reliably recorded by online platforms. According to the real-world dataset provided by JD.com (Shen et al. 2020), before making purchase decisions of a certain product, consumers may click the product page multiple times, during which the product price may undergo a few changes. The increased availability of detailed transaction data in online retailing calls for new and practical analytical tools to fully utilize such rich information (Ferreira et al. 2016). The second question is how to design an optimal pricing policy efficiently under a particular demand model. We use the term “optimal pricing policy” instead of “optimal price” because static pricing might be sub-optimal under special consumer behaviors. From an analytical perspective, understanding the structure of the optimal pricing policy can also provide insights into other retailing operations such as advertising and inventory control.

In retailing practice, it is commonly observed that consumer purchase behavior is influenced by both the current price and the price expectation. The reference-price-dependent demand model is one of the most well-established consumer behavioral demand models. We consider consumers who repeatedly visit a monopolist retailer, and their expectations toward price are influenced by historical prices. A consumer’s price expectation, used as a reference point when they make purchase decisions, is known as a *reference price*. Consumers perceive the discrepancy between the reference price and the current price as either *gains* if the reference price is higher than the current price, or *losses* if the reference price is lower than the current price. The demand change due to the presence of reference prices is known as the *reference effect*. The reference effect is the cornerstone of the prospect theory in economics (Kahneman and Tversky 1979). Kahneman and Tversky (1979) postulate that consumers are loss-averse — in the sense that they are more responsive to a loss than a gain of an equal amount. As a complement to loss-aversion, Greenleaf (1995) finds evidence of the other consumer type — gain-seeking consumers who are more responsive to gains than losses. Popescu and Wu (2007) study pricing under reference effects and establish structural properties of the optimal pricing policy, corresponding to the two consumer types respectively. Strong empirical evidences for reference effects is provided in the marketing literature (see Mazumdar et al. (2005) for a review).

The original prospect theory (Kahneman and Tversky 1979) is conducted at the individual level. However, previous studies on the characterization of optimal pricing policies under reference effects mostly consider demand models at the aggregate level, such as the widely adopted piece-wise linear model (Greenleaf 1995, Nasiry and Popescu 2011, Hu et al. 2016). To bridge the gap between the

individual consumer behavior and the market demand, we note that the aforementioned findings implicitly rely on at least the following assumptions. First, consumers homogeneously belong to one type (except some auxiliary discussions in the studies of Kopalle et al. (1996) and Popescu and Wu (2007)). Second, the aggregate demand from the whole market (or the aggregate demand from one consumer segment) naturally inherits the loss-averse or gain-seeking behavior from all the individual consumers. Third, consumer type (loss-averse or gain-seeking) is stationary — it is not affected by current prices and reference prices.

However, such assumptions can be prohibitive in practice. Gain-seeking and loss-averse consumers may coexist in the same market. Even when all consumers are of the same type, they may show diverse levels of responsiveness to reference effects. A few works have suggested that certain notions of consumer heterogeneity lead to different optimal pricing policies. Wang (2016) finds that the optimal pricing policy is cyclic if consumers follow heterogeneous arrival schedules. Chen and Nasiry (2020) consider a market comprising two consumer segments of different loss-aversion levels. They show that the optimal pricing policy is no longer constant, although neither consumer type is gain-seeking. Hu and Nasiry (2017) show that loss/gain preferences are inherited by the aggregate demand model if and only if consumer valuations are uniformly distributed.

To circumvent such practical limitations, we directly model individual consumer demand under reference effects. We allow arbitrary joint distributions of valuations, responsiveness to prices, and responsiveness to reference prices among consumers. Based on the observations from empirical data, consumers behave differently at different prices. Therefore we further relax the assumption on the stationary consumer types. In our model, a consumer could be either loss-averse or gain-seeking depending on the current price. Such non-stationary consumer types have not been captured by other studies. We elaborate more on the comparisons in Section EC.5.1 of the e-companion.

Main contributions. We address both model estimation and pricing optimization under reference effects. Our main contributions are summarized as follows.

1. *Demand model with reference price under heterogeneous consumer behaviors.* We propose a demand model to incorporate heterogeneous consumer behaviors into reference effects. In particular, our model has strong expressive power that can capture arbitrary distributions of consumer behaviors. Based on this individual-level demand model, we revisit and generalize notions of reference effects. Specifically, we introduce new concepts of gain-seeking, loss-averse, and neutral behaviors under mixed logit model and study its properties.
2. *Optimal pricing policies.* The pricing problem is considerably challenging in our setting, since it involves a trajectory of prices. We propose an efficient ϵ -approximation algorithm to solve the infinite horizon optimal pricing policy problem. Subsequently, we theoretically prove that constant pricing policies may not be optimal even when all consumers are locally loss-averse. Additionally, we give a performance guarantee of a computationally easy myopic pricing policy.

3. *Nonparametric estimation of consumer heterogeneity.* We use a nonparametric estimation method to learn the arbitrary joint distributions of responsiveness parameters from transaction data and characterize the structure of estimators. Numerical experiments show that integrating consumer heterogeneity improves the prediction accuracy significantly.
4. *Managerial insights for pricing practices.* We use transaction data from JD.com, a large E-commerce company, to validate our model and algorithms. We find that consumers are heterogeneous and cyclic pricing policies achieve the optimum for selected products in the dataset. Further, we observe over 40% revenue improvement for frequently purchased products when compared to the linear demand model with reference effects.

The remainder of the paper is organized as follows. In Section 2, we present a review of related literature on reference effects with a focus on consumer heterogeneity. In Section 3, we describe our modeling framework for individual consumer behavior. In Section 4, we address the pricing optimization problem. In Section 5, we explain a nonparametric estimation method for learning heterogeneous consumer reference effects. In Section 6, we present our numerical studies, including simulation studies and a case study on real data from JD.com. In Section 7, we conclude the paper with a discussion of future work.

2. Literature Review

Modeling of reference effects. Following Greenleaf (1995), reference effects have been primarily considered at the aggregate level (Popescu and Wu 2007, Aviv and Pazgal 2008, Nasiry and Popescu 2011, Hu et al. 2016, Chen et al. 2017, Chen and Nasiry 2020). Hu and Nasiry (2017) point out that individual level and aggregate level reference effects are not always identical. Specifically, the authors model a market of consumers with heterogeneous intrinsic valuations for one product. The authors show that the individual loss-averse or gain-seeking behaviors are not necessarily inherited by the aggregate market demand for most types of distributions of consumer valuations. The only exception is when consumers' valuations are uniformly distributed. From the perspective of demand modeling at the individual level, we share the same spirit as in the work of Hu and Nasiry (2017). However, we generalize the consumer heterogeneity considered by Hu and Nasiry (2017) from various perspectives, including intrinsic valuations, responsiveness to prices, and responsiveness to reference effects. We will discuss the study of Hu and Nasiry (2017) more comprehensively when we explain our modeling details in Section 3. Incidentally, modeling reference effects with individual choice-based models is favored in several empirical studies in the marketing literature (Chang et al. 1999, Bell and Lattin 2000, Kopalle et al. 2012). These works also provide strong validations for our individual-level demand model. More broadly, there have been abundant research interests in reference effects in OR/MS (see, e.g., Wu et al. (2015), Hedgcock et al. (2016), Hariss et al. (2020), Cohen et al. (2020),

Zhang et al. (2020), and see also Yu et al. (2021) on reference effects concerning time instead of price).

Empirical study of reference effects. Strong empirical evidence in the marketing literature indicates that consumer heterogeneity plays an important role when estimating reference effects (Chang et al. (1999), Bell and Lattin (2000), Klapper et al. (2005)). Kopalle et al. (2012) conduct an empirical analysis of pricing in the presence of household-level heterogeneity using retailing data for cola sales in a supermarket. However, the datasets used in these aforementioned works are mostly scanner-panel data collected in brick-and-mortar stores from more than a decade ago. Exploiting the JD.com dataset, we enrich this line of empirical works by examining reference effects on large-scale data from a modern online learning business.

Nonparametric estimation. Traditional estimation methods of consumer heterogeneity often involve additional assumptions, such as finite segments and hierarchical Bayesian priors (see McFadden and Train (2000) and Train (2009) for a comprehensive review). While these additional assumptions bring a higher risk of model misspecification, nonparametric methods refrain from such parametric assumptions and improve prediction accuracy. Specifically, our approach is based on a principled statistical methodology known as nonparametric maximum likelihood estimation (Kiefer and Wolfowitz 1956, Lindsay 1995). Several works explore other nonparametric estimation methods for pricing under settings that are quite different from ours (see, e.g., Chen and Gallego (2021), Wang et al. (2021)). In general, nonparametric modeling approaches of consumer choices have been an interesting and important topic (see, e.g., Farias et al. (2013), Chen and Mišić (2022), Jagabathula et al. (2020)).

Dynamic pricing and pricing optimization. Classical works (Kopalle et al. 1996, Popescu and Wu 2007) establish fundamental structural properties of the optimal pricing policy under reference effects (for results under a continuous time control framework, refer to Fibich et al. (2003)). Particularly, the constant pricing policy is optimal when facing loss-averse consumers and the high-low pricing policy is optimal when facing gain-seeking consumers. Popescu and Wu (2007) prove similar structural results for general demand functions; however, the proof relies on structural assumptions on the demand functions. Some other works deal with intertemporal pricing under aggregate-level reference effects (Nasiry and Popescu 2011, Wang 2016, Hu et al. 2016, Chen et al. 2017, Chen and Nasiry 2020).

Compared to more traditional pricing problems, a distinguishable characteristic of pricing optimization under reference effects is that the demand is not only affected by the current price but also prices from other periods. Therefore, the optimal pricing policy might not be static, but rather it could be composed of a trajectory of prices, probably cyclic over time. From this perspective, our study is related to the line of works on pricing under particular consumer behaviors such as being patient, strategic, forward-looking, etc. There is a vast and deep literature in each of the consumer behaviors mentioned above (see, e.g., Su (2007), Aviv and Pazgal (2008), Liu and Cooper (2015),

Besbes and Lobel (2015), Zhang and Jasin (2022)). Despite different problem settings, the optimal pricing policies are also shown to be cyclic under certain scenarios (Liu and Cooper 2015). Most importantly, we would like to highlight that consumer heterogeneity also draws significant attention. For example, Lobel (2020) considers cyclic pricing policies for consumers with heterogeneous patience levels.

Dynamic pricing has been extensively studied in the learning setting (see, e.g., Broder and Rusmevichientong (2012), Keskin and Zeevi (2014), Chen and Gallego (2021), den Boer and Keskin (2021)). Closely related to our work, den Boer and Keskin (2021) consider demand learning and pricing optimization under reference effects. Both theirs and our work assume that the retailer has little or no information on the price-demand relationship and must learn it from data. den Boer and Keskin (2021) deal with an online and sequential learning setting that needs to balance between exploration and exploitation. They aim to minimize the revenue gap between the optimal pricing policy and the online pricing policy. In contrast, we learn the demand from historical data in one shot, and design and analyze the optimal pricing policy that maximizes the long-term revenue. Moreover, our framework, which incorporates individual consumer heterogeneity, is not captured by the piece-wise linear aggregate demand model in den Boer and Keskin (2021). We believe that both their work and ours, as data-driven methodologies, are of independent interest in practice.

3. Demand Model and Heterogeneous Consumer Behaviors

In this section, we present a demand model that integrates reference effects and consumer heterogeneity. We analyze the fundamental properties of consumer behaviors under this demand model and systematically introduce different notions of reference effects.

3.1. Model

We consider a monopolist retailer selling a product with unlimited inventory to consumers over multiple periods. Consumers arrive with unit demand and form reference prices based on historical prices. At time t , the retailer sets price p_t for the product. Consumers make decisions based on the current price p_t and the reference price r_t . Each consumer is characterized by a set of parameters $\theta = (a, b, c_+, c_-)$, which we collectively refer to as *responsiveness parameters*. The parameter a represents the intrinsic valuation, b represents consumer responsiveness to prices, and (c_+, c_-) represents consumer responsiveness to reference prices. Consumers are heterogeneous in that their θ differ, and the feasible region of θ is denoted by Θ . At time t , the utility u_t for a consumer with parameters $\theta = (a, b, c_+, c_-)$ is defined as

$$u_t(\theta) = a - bp_t + c_+ \cdot (r_t - p_t)_+ + c_- \cdot (r_t - p_t)_- + \epsilon_t, \quad (1)$$

where we adopt the shorthand notation $(r_t - p_t)_+ = \max\{r_t - p_t, 0\}$, $(r_t - p_t)_- = \min\{r_t - p_t, 0\}$. ϵ_t represents the random fluctuation of consumer utility, following the standard Gumbel distribution.

The responsiveness parameter vector θ varies among consumers. Statistically speaking, θ is random and follows a probability distribution G . Notably, we do not impose any parametric assumptions on G . In the classical context of mixed logit models, it is common to assume that the distribution G is discrete with a finite number of components or that G follows a Gaussian distribution. However, as the distribution G directly affects the pricing strategy, and a misspecified G will lead to a significant revenue loss, as explained in Section 4 and Section 6.2, it is important to express the heterogeneity fully at the individual level. We still refer to our demand model as the mixed logit model under a broader context. Nevertheless, we would like to emphasize that the demand model is nonparametric, and it refrains from any parametric or prior assumptions on G .

For simplicity, we define function $\mathcal{C}(r - p)$ as a function of price discrepancy $r - p$ as

$$\mathcal{C}(r - p) = c_+ \cdot (r - p)_+ + c_- \cdot (r - p)_- = \begin{cases} c_+ \cdot (r - p) & \text{if } r - p \geq 0, \\ c_- \cdot (r - p) & \text{if } r - p \leq 0. \end{cases} \quad (2)$$

Therefore, the utility $u_t(\theta)$ in (1) can be written as $u_t(\theta) = a - bp_t + \mathcal{C}(r_t - p_t) + \epsilon_t$. According to the random utility theory (McFadden 1973), when ϵ_t follows the Gumbel distribution, the purchase probability of a customer under price p_t and reference price r_t with parameter θ is

$$\mathbb{P}(r_t, p_t | \theta) = \frac{\exp\{u_t(\theta)\}}{1 + \exp\{u_t(\theta)\}}, \quad (3)$$

where the utility for the no-purchase option is normalized to 0 without loss of generality.

The logit form of purchase probability under reference effects in Equation (3) is adopted in the marketing literature (see, e.g., Hardie et al. (1993), Bell and Lattin (2000), Kopalle et al. (2012)) and also in economics (Kallio et al. 2009). In OR/MS literature, Wang (2018) adopts a logit type demand model in studying assortment and pricing under reference effects. However, Wang (2018) considers a single-period multi-product setting while we study a multi-period single-product setting. Our model can be naturally extended to multiple products by introducing one utility function for every product (see discussion in Section 7). We denote the expected purchase probability among the population by $\mathbb{P}(r_t, p_t | G)$, which is calculated by taking the expectation of $\mathbb{P}(r_t, p_t | \theta)$ in (3) with respect to θ , i.e.,

$$\mathbb{P}(r_t, p_t | G) = \mathbb{E}_{\theta \sim G}[\mathbb{P}(r_t, p_t | \theta)] = \int_{\theta \in \Theta} \frac{\exp\{u_t(\theta)\}}{1 + \exp\{u_t(\theta)\}} dG(\theta). \quad (4)$$

Let $D^G(r_t, p_t)$ represent the expected demand from one consumer at time t when the price is p_t and reference price is r_t . Given any probability distribution G , $D^G(r_t, p_t)$ is

$$D^G(r_t, p_t) = \mathbb{P}(r_t, p_t | G). \quad (5)$$

The demand notation $D^G(r, p)$ is inclusive. As a special case, demand from one consumer with responsiveness parameter vector θ is represented by $D^{\delta_\theta}(r, p)$, where δ_θ is the point mass probability distribution concentrated at θ . Suppose the consumers in the market can be segmented into k groups according to responsiveness vectors, say, $\theta_1, \dots, \theta_k$. Let $\omega_1, \dots, \omega_k$ be the proportions of k groups respectively, so $\omega_1, \dots, \omega_k \geq 0$ and $\omega_1 + \dots + \omega_k = 1$. Subsequently, the demand is represented as $D^{G^{(k)}}$ with $G^{(k)} = \sum_{i=1}^k \omega_i \delta_{\theta_i}$, and $D^{G^{(k)}} = \sum_{i=1}^k \omega_i D^{\delta_{\theta_i}}$. We abbreviate the superscript G from D^G when the specification of G is clear from the context.

From a revenue maximization perspective, the retailer's goal is to determine an optimal trajectory of prices $(p_1, p_2, \dots)$ to maximize the expected long-term discounted revenue. We will formally define the pricing optimization problem in Section 4.

3.2. Loss-Averse and Gain-Seeking Behaviors

We define the *reference discrepancy* as the reference price r minus the current price p , which is denoted by $x = r - p$. As in Popescu and Wu (2007), we define the *reference effect* $R(x, p)$ as the difference incurred in demand, which is

$$R(x, p) = D(x + p, p) - D(p, p). \quad (6)$$

The discrepancy $x > 0$ corresponds to gains and $x < 0$ corresponds to losses. Generally speaking, consumers who respond stronger to losses are referred to as *loss-averse*; those who respond stronger to gains are referred to as *gain-seeking*; and those who do not have an explicit inclination to either side are called *loss-neutral*. The effect of x on demand is asymmetric when the same magnitude of x incurs different demand changes, depending on whether x is positive or negative.

Following Hu and Nasiry (2017), we use “loss-averse”, “gain/loss-neutral”, and “gain-seeking” to describe individual-level consumer behavior. We use “more responsive” to describe aggregate level behavior, i.e., the market demand. We give the global definition of asymmetric reference effects of the market demand in Definition 1. As we will see, the global definition imposes strict conditions.

DEFINITION 1 (GLOBAL). Given price p , we say the demand is (i) more responsive to losses (i.e., loss-averse) if $R(x, p) < -R(-x, p)$, for all $x > 0$; (ii) more responsive to gains (i.e., gain-seeking) if $R(x, p) > -R(-x, p)$, for all $x > 0$; (iii) equally responsive to losses and gains (i.e., loss-neutral) if $R(x, p) = -R(-x, p)$, for all $x > 0$.

Contrary to Definition 1, a more predominant local definition used by Kopalle et al. (1996) and Popescu and Wu (2007) is based on left and right partial derivatives of $R(x, p)$ in x . More specifically, they define $\lambda_-(p)$ and $\lambda_+(p)$ to be the left and right derivatives of $R(x, p)$ at $x = 0$, and the ratio of two derivatives valued at 0 is denoted as $\rho = \lambda_+(p)/\lambda_-(p)$. Three scenarios where $\rho < 1$, $\rho > 1$ and $\rho = 1$ are defined as gain-seeking, loss-averse and loss-neutral respectively. Following Hu and Nasiry

(2017), we shall refer to their definition of asymmetric reference effects as the local definition because the derivatives $\lambda_+(p)$, $\lambda_-(p)$ are defined in the infinitesimal region around 0. We formally introduce the local definition in the following Definition 2.

DEFINITION 2 (LOCAL). The demand is locally more responsive to losses if $\rho > 1$, locally more responsive to gains if $\rho < 1$, or locally responsive to gains and losses equally if $\rho = 1$.

Definition 2 coincides with Definition 1 when the demand model is linear in x , which is also the main parametric model adopted by Kopalle et al. (1996), Nasiry and Popescu (2011), and Chen et al. (2017). However, the local definition of reference effects crucially depends on the non-smoothness of the demand function, and it classifies any demand functions with continuous x -derivatives as loss-neutral. Moreover, requiring non-smoothness in demand functions brings greater difficulty in solving the subsequent pricing optimization problems (Chen et al. 2017).

In practice, consumer behaviors may be affected by the absolute value of the price discrepancy x . Motivated by such considerations, we introduce Definition 3 on *regional* reference effect.

DEFINITION 3 (REGIONAL). Given price p and $0 \leq w_1 \leq w_2$, we say the demand is (i) more responsive to losses (i.e., loss-averse) within region $[w_1, w_2]$ if $R(x, p) < -R(-x, p)$, for all $w_1 \leq x \leq w_2$; (ii) more responsive to gains (i.e., gain-seeking) within region $[w_1, w_2]$ if $R(x, p) > -R(-x, p)$, for all $w_1 \leq x \leq w_2$; (iii) equally responsive to losses and gains (i.e., loss-neutral) within region $[w_1, w_2]$ if $R(x, p) = -R(-x, p)$, for all $w_1 \leq x \leq w_2$.

To link the regional definition with the local definition, we provide the following lemma.

LEMMA 1. *If the demand is regionally loss-averse with $w_1 = 0$ and $w_2 > 0$ (Definition 3), the demand is locally loss-averse or neutral (Definition 2). Similarly, if the demand is regionally gain-seeking with $w_1 = 0$ and $w_2 > 0$, the demand is locally gain-seeking or neutral.*

As discussed by Hu and Nasiry (2017), Definitions 1 and 2 are related to some extent but one does not imply the other. In Hu and Nasiry (2017, Definition 3), they define “more responsive to losses/gains at large stakes”. In Definition 3, we introduce a regional definition that segments the price support into different regions according to the consumer’s different responses to reference discrepancy. If the demand is more responsive to losses at large stakes, the demand is more responsive to losses within the region $[w, \infty)$ for some $w > 0$, which corresponds to setting $w_1 = w$, and $w_2 = \infty$ in Definition 3. Deviating from linear demand models, we argue that Definition 1 and Definition 3 are more appropriate than Definition 2 in the logit demand model. As is shown in Proposition 2, varying parameters $\theta = (a, b, c_+, c_-)$ are sufficient to demonstrate both types of asymmetric reference effects. This is contrary to linear demand models, where non-smoothness is necessary to reflect the asymmetry in reference effects.

In conclusion, we have discussed three definitions: global, local and regional reference effect. Among the three definitions, the local definition (Definition 2), is predominantly used in the literature (Hu et al. 2016, Chen et al. 2017, Chen and Nasiry 2020); the global definition is recently considered in Hu and Nasiry (2017). In particular, we propose and study the regional definition (Definition 3), arguing that consumer behaviors may change across different regimes of x and p , without changing responsiveness parameters. As we will see in Section 4, the regional definition better describes individual consumer behaviors and sheds light on pricing optimization.

3.3. Properties of Regional Reference Effects

We discuss the properties of regional reference effects under our model. Proposition 1 (i) and (ii) echo with diminishing sensitivity and the decreasing curvature in Popescu and Wu (2007), respectively.

PROPOSITION 1. *For a consumer with responsiveness parameters θ satisfying $a, c_+, c_- \geq 0$, the reference effect satisfies the following two properties:*

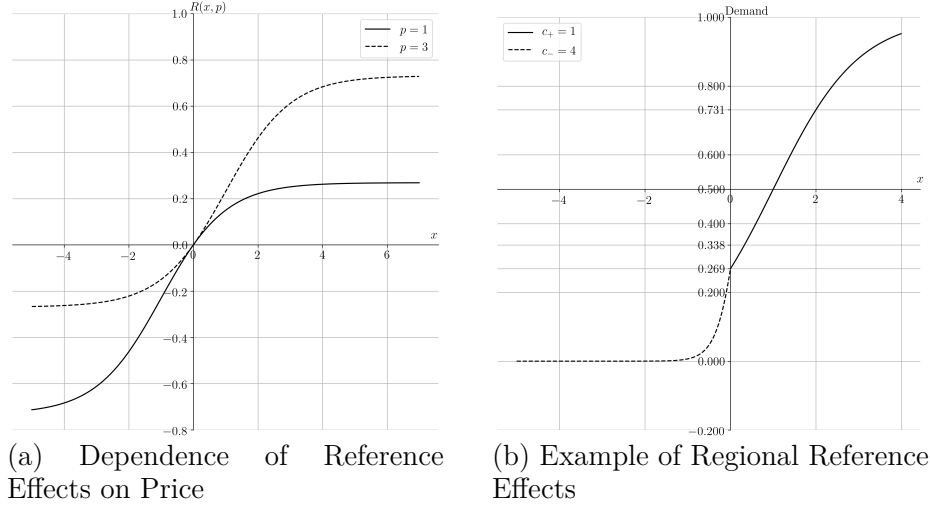
- (i) $R(x, p)$ is concave in $x \geq 0$ if $a - bp \geq 0$; $R(x, p)$ is convex in $x \leq 0$ if $a - bp \leq 0$.
- (ii) When $c_+ = c_-$, $R(x_1, p) - R(x_2, p)$ decreases in p if $x_2 < x_1 < (bp - a)/c_+$.

These two properties are not reflected by the linear demand models, as demand changes at a constant rate due to linearity. Diminishing sensitivity is originally stated in the prospect theory, and is reasonably aligned with intuition. As quoted from (Hardie et al. 1993, Page 380), “the marginal value of both gains and losses decreases with their size; the first sip of beer tastes the best, and the first dollar lost hurts the most.” To provide some intuition on the decreasing curvature, for example, most consumers may be willing to spend \$3 on a bottled juice when receiving a \$1 discount, whereas they will be less likely to respond to a \$1 discount on a \$50 steak. Although the absolute amount of price change is the same, consumers may respond less when the price is higher.

We illustrate the reference effects under the logit demand model with two concrete examples.

EXAMPLE 1. Suppose $a = 2, b = 1, c_+ = c_- = 1$, Figure 1(a) shows $R(x, p)$ against x for fixed $p = 1$ and fixed $p = 3$ respectively. We can verify the decreasing curvature property because the curve with $p = 3$ is flatter than that with $p = 1$ when the condition in Proposition 1 $x < \frac{bp-a}{c_+}|_{p=1} = -1$ holds.

EXAMPLE 2. Suppose $a = 0, b = 1, c_+ = 1, c_- = 4$ and $p = 1$, Figure 1(b) shows the demand against the reference discrepancy x . When comparing the demand change incurred by gains and losses, the demand is more responsive to losses when $|x|$ is small and more responsive to gains when $|x|$ is large. This can be concluded from Figure 1(b) by comparing changes in x when the demand increases or decreases by the same quantity. Starting from $x = 0$ with a demand of 0.269, it takes a larger change on x for the solid line (demand when there is gain) to increase by 0.069 compared to the amount of change on x for the dashed line to decrease by 0.069. We can also reach the same conclusion by comparing anchor points (2, 0.731) and (−2, 0.0) in Figure 1(b).

Figure 1 Illustration of Logit Demand

PROPOSITION 2. For a consumer with responsiveness parameters $\theta = (a, b, c_+, c_-)$, the reference effect at the individual level is defined as follows,

$$R(x, p) = \begin{cases} \frac{\exp(a-bp+c_+x)}{1+\exp(a-bp+c_+x)} - \frac{\exp(a-bp)}{1+\exp(a-bp)} & \text{if } x > 0, \\ \frac{\exp(a-bp+c_-x)}{1+\exp(a-bp+c_-x)} - \frac{\exp(a-bp)}{1+\exp(a-bp)} & \text{if } x < 0. \end{cases} \quad (7)$$

There are the following scenarios of individual consumer behavior based on different relations of responsiveness parameters θ .

- (i) When either of the conditions 1) $a - bp < 0, c_+ - c_- \geq 0$ or 2) $a - bp \leq 0, c_+ - c_- > 0$ holds, the consumer is regionally gain-seeking.
- (ii) When either of the conditions 1) $a - bp > 0, c_+ - c_- \leq 0$ or 2) $a - bp \geq 0, c_+ - c_- < 0$ holds, the consumer is loss-averse.
- (iii) When $a - bp = 0$ and $c_+ - c_- = 0$, the consumer is loss-neutral.
- (iv) When $a - bp < 0, c_+ < c_-$ or $(a - bp) > 0, c_+ > c_-$, whether the consumer is gain-seeking or loss-averse depends on x . To be more specific:
 - (a) When $a - bp < 0, 0 \leq c_+ < c_-$, there exists a point $x^* > 0$ such that the consumer is loss-averse when $|x| \leq x^*$ and gain-seeking when $|x| > x^*$.
 - (b) When $a - bp > 0, c_+ > c_- \geq 0, c_- > 0$, there exists a transition point $x^{**} > 0$ such that the consumer is gain-seeking when $|x| \leq x^{**}$ and loss-averse when $|x| > x^{**}$.

Proposition 2 shows that the logit demand, even at the individual level, naturally introduces more complex consumer behaviors compared to the linear demand model. In particular, Proposition 2 (iv) reflects the regional reference effect as we defined it in Definition 3. For all four statements in Proposition 2, the gain-seeking or loss-averse behaviors are retained even when $c_+ - c_-$ is set as

0. This indicates that the non-smooth demand assumption is not absolutely necessary for allowing consumers to deviate from loss-neutral behaviors in the logit demand model. On the other hand, these findings are related to Example 2 of Hu and Nasiry (2017, page 1387). The authors derive that the market demand is of logit demand form when consumer valuations follow the logistic distribution. Our logit demand model aligns with that of Hu and Nasiry (2017); our parameter a , which reflects intrinsic valuation, effectively corresponds to the valuation v in Hu and Nasiry (2017). However, our demand model (4) represents the purchase probability for one single consumer or the aggregate demand from a normalized homogeneous market, while the logit demand model in Hu and Nasiry (2017) refers to the market demand only. Additionally, we introduce randomness in the consumer valuation term a with a nonparametric distribution in our model. Consequently, our demand model has a stronger expressive power than the logit model example in Hu and Nasiry (2017). Notably, the x^* in statement (iv)(a) (similarly x^{**} in statement (iv)(b)) depends on responsiveness parameters (a, b, c_+, c_-) . Therefore, when (a, b, c_+, c_-) vary among consumers, consumers may encounter behavior change at different x^* , and thus whether the aggregate demand is more responsive to gains or losses may change multiple times as x changes. This demonstrates the necessity of introducing Definition 3 where consumers could behave very differently when the reference discrepancy $x = r - p$ varies.

4. Pricing Optimization

In this section, we focus on the pricing optimization problem based on the model in Section 3. Given an initial reference price and estimated demand model parameters, the goal is to find an optimal pricing policy that maximizes the long-term revenue.

4.1. Problem Formulation

We use $\Pi(r, p) = pD(r, p)$ to denote the retailer's single-period revenue. Given reference price r_t and price p_t at period t , the single-period revenue is

$$\Pi(r_t, p_t) = p_t D(r_t, p_t) = \int_{\theta \in \Theta} p_t \frac{\exp\{a - bp_t + \mathcal{C}(r_t - p_t)\}}{1 + \exp\{a - bp_t + \mathcal{C}(r_t - p_t)\}} dG(\theta).$$

We adopt the widely used memory-based reference price model, where reference price at time t is completely determined by the reference price from the last period $t - 1$ and the price during period $t - 1$. The updating rule of the reference price is captured by a bivariate function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, and $r_t = f(r_{t-1}, p_{t-1})$ (Baucells et al. 2011). Here $f(x, y)$ is non-decreasing in both x and y . Following Greenleaf (1995), Popescu and Wu (2007) and other literature, we take f as a linear function f_α defined as $f_\alpha(x, y) = \alpha x + (1 - \alpha)y$ for some $\alpha \in [0, 1]$. With f_α , the reference price is updated as $r_t = \alpha r_{t-1} + (1 - \alpha)p_{t-1}$, where α represents a memory parameter that describes how the reference price in the current period depends on the reference price in the last period. By this recursive

definition of reference prices, r_t is an exponentially weighted sum of the initial reference price and all the past prices before time t . Thus, this updating scheme is also known as exponential smoothing.

We consider the discounted setting where the revenue at time t is discounted by factor β^t for some $\beta \in (0, 1)$. Similar to Popescu and Wu (2007), Hu and Nasiry (2017), Hu et al. (2016), and Chen and Nasiry (2020), we focus on an infinite horizon setting and aim to maximize the long-term revenue. Given an initial reference price r_0 , the retailer's objective is to determine a price sequence $\{p_t\}_{t=0}^\infty$ to maximize $\sum_{t=1}^\infty \beta^t \Pi(r_t, p_t)$.

We now formulate this pricing optimization problem as an infinite horizon dynamic programming problem and formally define the scope of pricing policies considered here. Let $\mathcal{P}$ be the set of feasible prices and $\mathcal{P} \subset [0, p_H]$ where p_H is an upper bound. At time t , reference price r_t is given based on the historical price, and a pricing policy determines a price $p_t \in \mathcal{P}$ according to r_t . Using the language of dynamic programming, we view r_t as the state at time t and p_t as the control at time t . A new reference price r_{t+1} is formulated according to the updating rule $f_\alpha(r_t, p_t)$. Thus a pricing policy can be equivalently viewed as a mapping from the reference price (state) to the price (control). We introduce notation p_π , a function that maps a reference price to a price. Here π indexes a pricing policy. Such policies are also known as stationary policies in the dynamic programming and control area since the function p_π does not change over time. It is important not to confuse the concept of the stationary policy here with the steady/constant pricing policy. Under stationary policies, the prices may change but the function p_π remains unchanged over time.

The long-term discounted revenue by pricing policy p_π is

$$\begin{aligned} V_\pi(r_0) &= \sum_{t=1}^\infty \beta^t \Pi(r_t, p_\pi(r_t)) \\ \text{s.t. } r_t &= \alpha r_{t-1} + (1 - \alpha) p_\pi(r_{t-1}) \quad t = 1, 2, \dots \end{aligned}$$

Optimal pricing policy. The optimal pricing policy refers to any pricing policy that achieves the maximum long-term revenue $V^*(r_0)$ given any initial reference price r_0 , i.e.,

$$\begin{aligned} V^*(r_0) &= \max_{p_t \in \mathcal{P}} \sum_{t=0}^\infty \beta^t \Pi(r_t, p_t) \\ \text{s.t. } r_t &= \alpha r_{t-1} + (1 - \alpha) p_{t-1} \quad t = 1, 2, \dots \end{aligned} \tag{8}$$

We use p_{π^*} to denote the optimal pricing policy, then according to the optimality condition, p_{π^*} satisfies the Bellman equation,

$$V^*(r) = \Pi(r, p_{\pi^*}(r)) + \beta V^*(\alpha r + (1 - \alpha) p_{\pi^*}(r)). \tag{9}$$

In addition, we study a popular heuristic policy — the myopic pricing policy. In Proposition 3, we establish a performance guarantee of the myopic pricing policy by comparing its long-term revenue denoted by V_m to V^* . We summarize the proof of Proposition 3, a simulation study, and detailed discussion of the myopic pricing policy into Section EC.2 of the e-companion.

PROPOSITION 3. Let V^* be the long-term revenue of the optimal pricing policy and let V_m be the long-term revenue of the myopic pricing policy. It holds for any initial reference price r that

$$0 \leq V^*(r) - V_m(r) \leq \frac{\beta(1-\alpha)}{(1-\alpha\beta)(1-\beta)} \eta p_H,$$

where $\eta := \min \left\{ 1, \sup_{(a,b,c_+,c_-) \in \Theta} \frac{\max(c_+,c_-)}{b+c_-} \right\}$.

4.2. Properties of Optimal Pricing Policy

We consider the optimal pricing policy under our demand model. The pricing optimization problem (8) is challenging in general, since it is both non-convex and non-smooth (Hu et al. 2016). Proposition 4 gives the monotonicity property of both single-period revenue and long-term revenue under certain conditions.

PROPOSITION 4. If c_+ and c_- are both non-negative over Θ , $\Pi(r, p)$ and $V^*(r)$ are monotonically non-decreasing in r .

In practice, the responsiveness parameter corresponding to $r - p$ could be negative, in which case the monotonicity does not hold. This motivates us to design a computationally efficient algorithm that works without the nonnegative assumption, and the algorithm is described later in this section. Before doing computation, a natural question to ask is whether a constant pricing policy is optimal. Our Proposition 5 suggests that a constant pricing policy may not be optimal when parameters are in certain regions. This is in contrast with the structural result in Popescu and Wu (2007), which is based on a general demand function under some technical assumptions.

PROPOSITION 5. In the logit demand model, a constant pricing policy is not optimal if $c_+ \leq c_-$ and utility parameters a, b satisfy that $a < \log \left(\frac{1-\beta}{\beta} \right)$ for all consumers, where c_- is sufficiently large.

REMARK 1. For condition $a < \log \left(\frac{1-\beta}{\beta} \right)$ in Proposition 5, the term a can be viewed as the initial value of consumer utility in the absence of reference effects. When the discount parameter β is 0.95, $\log \left(\frac{1-\beta}{\beta} \right)$ is about -2.94 . Therefore the condition holds when the initial purchase probability is roughly smaller than 5%. This fact aligns with our observation in the JD.com dataset, where the purchase probabilities for most products are lower than 5%.

We interpret Proposition 5 based on our Proposition 2 in Section 3. According to the local definition (Definition 2) of reference effects, consumers satisfying assumptions $(c_+ \leq c_-)$ in Proposition 5 are loss-averse/neutral, which would make the optimal pricing policy static according to the classic result. On the other hand, since $a - bp$ is negative given that $a < \log \left(\frac{1-\beta}{\beta} \right)$ and the usual choice of β is larger than $\frac{1}{2}$, the scenario in Proposition 5 falls into the category of statement (iv)(a) in Proposition 2. By invoking the conclusion in Proposition 2 (iv)(a), consumers are not necessarily loss-averse here, and a consumer is gain-seeking when $x > x^*$ for some x^* dependent on parameters

a, b, c_+ and c_- . Therefore, it is reasonable that the optimal pricing policy is not constant, and this argument necessitates the introduction of Definition 3 on regional reference effects.

We compare Proposition 5 with two relevant results from Hu and Nasiry (2017) and Chen and Nasiry (2020), respectively. Our Proposition 5 indicates that the constant pricing policy may not be optimal even if consumers are locally loss-averse in the logit demand model. In Proposition 6 of Hu and Nasiry (2017), the authors consider the exponential demand model and show that a steady state does not exist for a certain range of demand parameters under a simplified assumption that the memory parameter α is 0. Note that Proposition 5 in our paper does not rely on this assumption. Our Proposition 5 is also related to Proposition 5 of Chen and Nasiry (2020), which states that the constant pricing policy is not optimal when a certain demand model parameter is sufficiently large. However, our demand model is considerably different from theirs since Chen and Nasiry (2020) consider a piece-wise linear demand model with two market segments. In contrast, our model can potentially have arbitrarily large number of market segments.

4.3. Computation of Optimal Pricing Policy

Now we show how to compute the optimal pricing policy and illustrate the structure of the optimal pricing policy. For computational feasibility, we use price discretization, which is also common in practice because of the natural granularity in prices. Let $\mathcal{R}_\epsilon$ represent an ϵ -approximation of the continuous price range $\mathcal{P}$ so that for any $r \in \mathcal{P}$, $\min_{\bar{r} \in \mathcal{R}_\epsilon} |\bar{r} - r| \leq \epsilon$. The feasible set of price p_t is $\mathcal{P}_\epsilon(r_t)$ and is defined as $\mathcal{P}_\epsilon(r_t) := \{p \in \mathcal{P} : \alpha r_t + (1 - \alpha)p \in \mathcal{R}_\epsilon\}$, or equivalently $\mathcal{P}_\epsilon(r_t) = \{(r - \alpha r_t)/(1 - \alpha) | r \in \mathcal{R}_\epsilon\} \cap \mathcal{P}$. Therefore, the choice of $\mathcal{P}_\epsilon(r_t)$ ensures that any r_{t+1} satisfying the equality constraint $r_{t+1} = \alpha r_t + (1 - \alpha)p_t$ must lie in $\mathcal{R}_\epsilon$. In summary, we consider a discretized version of the pricing optimization problem as follows:

$$\begin{aligned} V_\epsilon^*(r_0) &= \max_{p_t \in \mathcal{P}} \sum_{t=0}^{\infty} \beta^t \Pi(r_t, p_t) \\ \text{s.t. } r_{t+1} &= \alpha r_t + (1 - \alpha)p_t, \quad t = 0, 1, 2, \dots \\ p_t &\in \mathcal{P}_\epsilon(r_t), \quad t = 0, 1, 2, \dots \end{aligned} \tag{10}$$

Note that we consider a discretized price space only for the computational purpose. Our previous theoretical analysis applies to the general continuous price space. We have the following accuracy guarantee for discretization, in terms of the long-term revenue V_ϵ^* .

PROPOSITION 6. *For any $r \in \mathcal{R}_\epsilon$,*

$$0 \leq V^*(r) - V_\epsilon^*(r) \leq \left[\frac{C_1}{4(1 - \alpha)(1 - \beta)} + \frac{\beta C_2}{4(1 - \alpha)(1 - \alpha\beta)} \right] \epsilon, \tag{11}$$

where $C_1 = \int_{\theta \in \Theta} \max\{b + c_+, b + c_-\} dG(\theta)$ and $C_2 = \int_{\theta \in \Theta} \max\{c_+, c_-\} dG(\theta)$.

REMARK 2. Notably, the bound on the right hand side of (11) depends linearly on ϵ . Moreover, when the first moment of coefficients over the mixing probability distribution G exists, it is guaranteed that both C_1 and C_2 are finite. Since this bound applies to any $r \in \mathcal{R}_\epsilon$, it follows that

$$\limsup_{\epsilon \rightarrow 0} \sup_{r \in \mathcal{R}_\epsilon} |V^*(r) - V_\epsilon^*(r)| = 0.$$

We compute optimal pricing policies with a modified policy iteration algorithm described in Algorithm 1. The modified policy iteration algorithm is known to converge to optimality (Bertsekas 2019). Compared to the value iteration algorithm, policy iteration is computationally more efficient in general (Bertsekas 2019). In Algorithm 1, we generate a sequence of long-term revenue function and pricing policy $\{V^{(k)}(r), p_{\pi^k}(r)\}$ for all r following an iterative procedure that alternates between policy improvement and approximate policy evaluation. At iteration k , policy p_{π^k} is obtained by improving the previous policy $p_{\pi^{k-1}}$ in line 5-7, and then cumulative revenue from the first s_k periods is used to approximate the long-term value function of policy p_{π^k} in line 8-10. The constraint in line 7 of Algorithm 1 restricts the choices of price at time t so that the updated reference price of the next period $t+1$ still falls into the discretized set $\mathcal{R}_\epsilon$. The choice of s_k in line 9 of Algorithm 1 does not affect the final convergence. For a more accurate policy evaluation at each iteration, we recommend a relatively larger s_k so that α^{s_k} is smaller.

Algorithm 1 Modified Policy Iteration Algorithm for Finding Optimal Pricing Policy

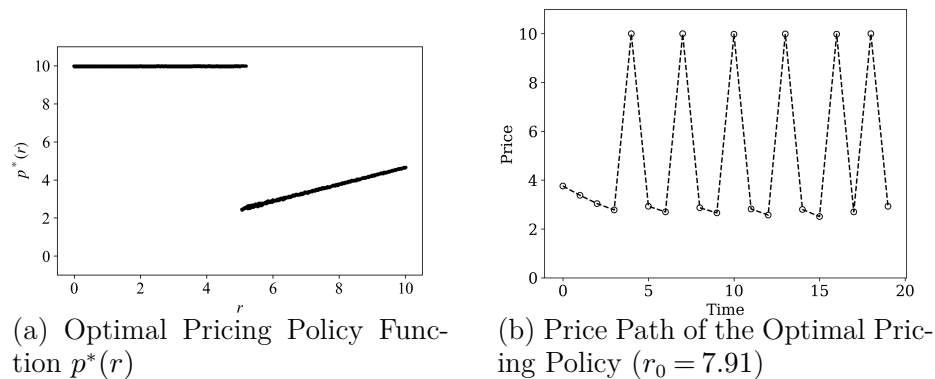
- 1: Input: $\delta > 0$
 - 2: Initialize: $k = 0$, $V^{(0)}(r) = 0$, $\forall r \in \mathcal{R}_\epsilon$
 - 3: **repeat**
 - 4: Increase k by 1
 - 5: Policy improvement: Generate new pricing policy p_{π^k} according to
 - 6:
$$p_{\pi^k}(r) \in \arg \max_{p \in \mathcal{P}} \{\Pi(r, p) + \beta V^{(k-1)}(\alpha r + (1 - \alpha)p)\},$$
 - 7: s.t. $\alpha r + (1 - \alpha)p_{\pi^k}(r) \in \mathcal{R}_\epsilon$, $\forall r \in \mathcal{R}_\epsilon$
 - 8: Approximate policy evaluation: Calculate the long-term revenue of pricing policy p_{π^k} for any initial reference price $r \in \mathcal{R}_\epsilon$ by
 - 9:
$$V^{(k)}(r) = \beta^{s_k} V^{(k-1)}(r_{s_k}) + \sum_{j=0}^{s_k-1} \beta^j \Pi(r_j, p_{\pi^k}(r_j)),$$
 - 10: s.t. $r_0 = r$, $r_j = \alpha r_{j-1} + (1 - \alpha)p_{\pi^k}(r_{j-1})$, $j = 1, \dots, s_k$
 - 11: **until** $|V^{(k)}(r) - V^{(k-1)}(r)| < \delta$, $\forall r \in \mathcal{R}_\epsilon$
-

4.4. Numerical Examples of Pricing Policies

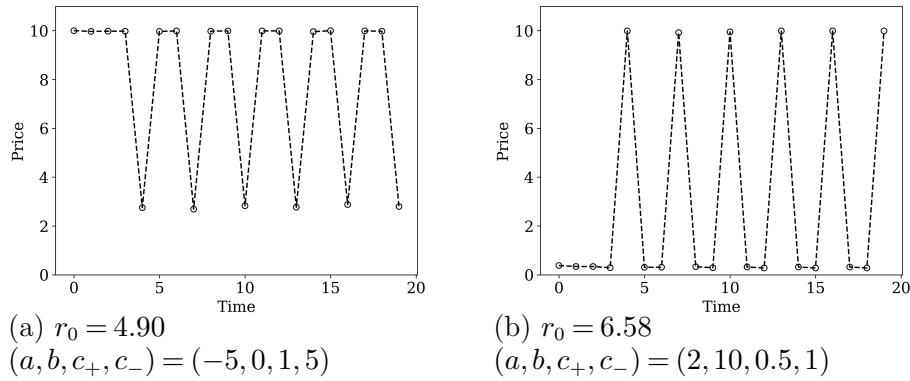
In all numerical examples, we fix the price range $\mathcal{P} = [0, 10]$ and the discretization accuracy $\epsilon = 0.01$. The price memory parameter is $\alpha = 0.80$, and the discount rate is $\beta = 0.95$. Parameters in Algorithm 3 are chosen as $s_k = 10^3$ for all k and $\delta = 10^{-100}$. Given a discretization accuracy ϵ and a price range $\mathcal{P}$ of length L , the size of the discretized price space $\mathcal{R}_\epsilon$ is $L\epsilon^{-1}$. Using one core on a 2.6 GHz 6-Core Intel Core i7 processor, the average running time of Algorithm 1 is 10 minutes if $L = 10$, $\epsilon = 0.01$ and $|\mathcal{R}_\epsilon| = 10^3$, and is less than one minute if $\epsilon = 0.1$ and $|\mathcal{R}_\epsilon| = 10^2$. In Figures 2 and 3, initial reference prices are randomly selected by default, since they do not affect the optimal pricing policy and in most cases only affect the first few periods of optimal price paths.

Cyclic pricing policy can be optimal for smooth demand functions. In Figure 2, we consider one consumer with responsiveness parameters $(a, b, c_+, c_-) = (1, 1, 1, 1)$. The demand function of this consumer is smooth since $c_+ = c_-$, and this consumer is locally loss-neutral according to Definition 2. at the aggregate demand counterpart, it is generally believed that such a consumer responds equally to gains and losses so that the optimal pricing policy is constant (Kopalle et al. 1996, Popescu and Wu 2007). However, this simulation firmly shows that such general rule does not apply to our demand model. Figure 2 shows the optimal price path, which is cyclic. From the theoretical perspective, the sub-optimality of the constant pricing policy is captured by Proposition 5.

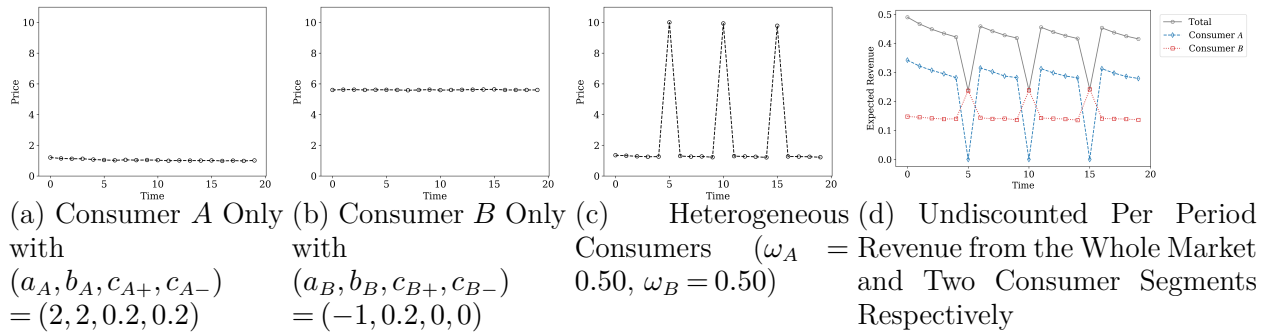
Figure 2 Pricing Optimization for Locally Loss-neutral Consumers $(a, b, c_+, c_-) = (1, 1, 1, 1)$



Local loss-aversion does not preclude price variations. We show two numerical examples of pricing for locally loss-averse consumers. In Figure 3(a) and Figure 3(b), the responsiveness parameters are $(-5, 0, 1, 5)$ and $(2, 10, 0.5, 1)$ respectively. In both cases, the condition $c_- > c_+$ holds so consumers are locally loss-averse according to Definition 2. However, the optimal pricing policy is cyclic. Two examples in Figure 3 illustrate the difference between our model and the linear demand model, and further validate Proposition 5.

Figure 3 Cyclic Optimal Pricing Policy for Loss-averse Consumers

Constant optimal pricing + constant optimal pricing $\neq$ constant optimal pricing. We consider two types of consumers coexisting in the same market and refer to them as consumer *A* and consumer *B*. The responsiveness parameters of consumer *A* and *B* are $(2, 2, 0.2, 0.2)$ and $(-1, 0.2, 0, 0)$, respectively. As shown in Figure 4(a) and Figure 4(b), a constant pricing policy is optimal for a market of consumer *A* only or consumer *B* only. However, the constant values are different. The optimal constant price p_A^* for consumer *A* lies in the interval $[0, 2]$, while p_B^* lies in $[4, 6]$. We assume that two types of consumers coexist in the market with equal proportions, and the price is uniform for all consumers. Figure 4(c) shows the optimal pricing path, which is cyclic. For Figure 4 (a)(b)(c), we fix r_0 as the average price 5.0 for better alignment in comparisons. Figure 4(d) shows the per period revenue obtained from consumers of type *A* and type *B* under the pricing policy in Figure 4(c), and we leave some discussion of Figure 4(d) to Section EC.3.4 of the e-companion.

Figure 4 Optimal Price Paths for Homogeneous versus Heterogeneous Consumers

5. Learning Consumer Heterogeneity

In this section, we discuss our computational procedure for recovering heterogeneous consumer reference effects, which is a preceding step of pricing optimization. To be precise, the estimation goal is

to retrieve the joint mixing distribution G of consumer responsiveness parameters $\boldsymbol{\theta} = (a, b, c_+, c_-)$ from transaction data, without imposing any parametric assumptions on the mixing distribution G .

5.1. Likelihood Function and Optimization Problem

Suppose N data points $\{(r_n, p_n, d_n)\}_{n=1}^N$ have been collected from multiple consumers across multiple periods, where $d_n \in \{0, 1\}$ represents a purchase decision, and r_n, p_n represent a reference price and a price respectively. We define the covariate vector $\mathbf{x}_n = (1, -p_n, (r_n - p_n)_+, (r_n - p_n)_-)$ for $1 \leq n \leq N$. The utility of purchase is $\mathbf{x}_n^\top \boldsymbol{\theta}$, and that of no-purchase is 0. The likelihood of (r_n, p_n, d_n) given mixing distribution G is then

$$\mathcal{L}(\mathbf{x}_n, d_n | G) = d_n \int_{\boldsymbol{\theta} \in \Theta} \frac{\exp\{\mathbf{x}_n^\top \boldsymbol{\theta}\}}{1 + \exp\{\mathbf{x}_n^\top \boldsymbol{\theta}\}} dG(\boldsymbol{\theta}) + (1 - d_n) \int_{\boldsymbol{\theta} \in \Theta} \frac{1}{1 + \exp\{\mathbf{x}_n^\top \boldsymbol{\theta}\}} dG(\boldsymbol{\theta}), \quad (12)$$

where Θ is some feasible region of $\boldsymbol{\theta}$. Given G , the likelihood vector $\mathbf{f}^G$ is composed of N choice probabilities, i.e., $\mathbf{f}^G(n) = \mathcal{L}(\mathbf{x}_n, d_n | G)$ for $1 \leq n \leq N$. Corresponding to the case where $G = \delta_{\boldsymbol{\theta}}$, $\mathbf{f}^{\delta_{\boldsymbol{\theta}}}$ is called an atomic likelihood vector and denoted by $\mathbf{f}^{\boldsymbol{\theta}}$ for simplicity, i.e., $\mathbf{f}^{\boldsymbol{\theta}}(n) = \mathcal{L}(\mathbf{x}_n, d_n | \delta_{\boldsymbol{\theta}})$ for $1 \leq n \leq N$.

Following the nonparametric maximum likelihood estimation (MLE) (Kiefer and Wolfowitz 1956) method, the estimator $\hat{G}$ is obtained by maximizing the total log likelihood function,

$$\max_G \ell(\mathbf{f}^G) = \sum_{n=1}^N \log\{\mathbf{f}^G(n)\} \text{ s.t. } G \text{ is supported on } \Theta. \quad (13)$$

In (13), G can be an arbitrary (either continuous or discrete) probability distribution, which is different from traditional MLE where G is usually assumed to follow a certain parametric form (Meng and Rubin 1993).

The maximization objective $\ell(\mathbf{f}^G)$ in (13) is concave in $\mathbf{f}^G$, and the constraint space of $\mathbf{f}^G$, $\mathcal{G} = \{\mathbf{f}^G : G \text{ is supported on } \Theta\}$, is convex. Therefore, (13) is a convex optimization problem with respect to $\mathbf{f}^G$. The constraint space $\mathcal{G}$ is quite large since Θ is usually uncountable and even unbounded. To make $\mathcal{G}$ more tractable, we invoke the fact that (Lindsay 1983)

$$\mathcal{G} = \text{conv}(\text{cl}(\{\mathbf{f}^{\boldsymbol{\theta}} | \boldsymbol{\theta} \in \Theta\})), \quad (14)$$

where $\text{cl}(\{\mathbf{f}^{\boldsymbol{\theta}} | \boldsymbol{\theta} \in \Theta\})$ represents the closure of set $\{\mathbf{f}^{\boldsymbol{\theta}} | \boldsymbol{\theta} \in \Theta\}$ and conv represents the convex hull. The closure $\text{cl}(\{\mathbf{f}^{\boldsymbol{\theta}} | \boldsymbol{\theta} \in \Theta\})$ is often strictly larger than $\{\mathbf{f}^{\boldsymbol{\theta}} | \boldsymbol{\theta} \in \Theta\}$ when Θ is unbounded because $\mathbf{f}^{\boldsymbol{\theta}}(n) \neq 0, 1$ for any $\mathbf{f}^{\boldsymbol{\theta}} \in \{\mathbf{f}^{\boldsymbol{\theta}} | \boldsymbol{\theta} \in \Theta\}$ while $\text{cl}(\{\mathbf{f}^{\boldsymbol{\theta}} | \boldsymbol{\theta} \in \Theta\})$ may include an extreme vector that contains limiting value 0 or 1 of the choice probabilities. Combining (14) with (13), we rewrite the likelihood maximization problem as follows,

$$\max_G \ell(\mathbf{f}^G) = \sum_{n=1}^N \log\{\mathbf{f}^G(n)\} \text{ s.t. } \mathbf{f}^G \in \text{conv}(\text{cl}(\{\mathbf{f}^{\boldsymbol{\theta}} | \boldsymbol{\theta} \in \Theta\})), \quad (15)$$

which is the basis of our computational procedure introduced later in Section 5.2.

5.2. Computation of Mixing Distribution

Our computational procedure summarized in Algorithm 2 is built upon the framework of the conditional gradient (CG) algorithm, also known as the Frank-Wolfe algorithm (Frank et al. 1956). Jagabathula et al. (2020) apply the CG algorithm to estimate general mixed logit models and show that it has superior performances compared to the popular expectation-maximization (EM) algorithm (Train 2008). In Algorithm 2, we apply CG to the optimization problem in (15) with necessary modifications.

Algorithm 2 Computing Nonparametric MLE with the Conditional Gradient Method

- 1: Input: data points $\{(r_n, p_n, d_n)\}_{n=1}^N$, number of iterations T
 - 2: Initialize: $\mathbf{f}_{(0)} = \mathbf{f}^{\theta_0}$, support set $\mathcal{S}_{(0)} = \{\mathbf{f}^{\theta_0}\}$
 - 3: **for** $t = 0, 1, \dots, T-1$ **do**
 - 4: $\tilde{\mathbf{g}}_{(t)} \in \arg \max_{\mathbf{g} \in \text{cl}(\{\mathbf{f}^\theta | \theta \in \Theta\})} \mathbf{g} \cdot \nabla \ell(\mathbf{f}_{(t)})$ $\triangleright$ Find consumer segment via solving subproblem
 - 5: $\mathcal{S}_{(t+1)} = \mathcal{S}_{(t)} \cup \{\tilde{\mathbf{g}}_{(t)}\}$ $\triangleright$ Update support set of consumer segments
 - 6: $\mathbf{f}_{(t+1)} = \arg \max_{\mathbf{f} \in \text{conv}(\mathcal{S}_{(t+1)})} \ell(\mathbf{f})$ $\triangleright$ Remaximize objective function over new support set
 - 7: **end for**
 - 8: Output: likelihood vector $\mathbf{f}_{(T)}$, support set $\mathcal{S}_{(T)}$ and corresponding responsiveness parameters
-

At each iteration, Algorithm 2 adaptively searches for a new consumer segment by solving a subproblem. Algorithm 2 finds a new support point $\tilde{\mathbf{g}}_{(t)}$ that solves the subproblem

$$\max_{\mathbf{g} \in \text{cl}(\{\mathbf{f}^\theta | \theta \in \Theta\})} \mathbf{g} \cdot \nabla \ell(\mathbf{f}_{(t)}), \quad (16)$$

where $\mathbf{f}_{(t)}$ is the likelihood vector at iteration t . The subproblem (16) is equivalent to maximizing the first-order Taylor approximation of ℓ around $\mathbf{f}_{(t)}$, which is

$$\max_{\mathbf{g} \in \text{cl}(\{\mathbf{f}^\theta | \theta \in \Theta\})} \ell(\mathbf{f}_{(t)}) + (\mathbf{g} - \mathbf{f}_{(t)}) \cdot \nabla \ell(\mathbf{f}_{(t)}).$$

A solution $\tilde{\mathbf{g}}_{(t)}$ to the subproblem (16) is a feasible direction (belonging to $\text{cl}(\{\mathbf{f}^\theta | \theta \in \Theta\})$) that is the closest to the direction of the gradient $\nabla \ell(\mathbf{f}_{(t)})$, in which the objective function increases at the fastest rate. The objective function is then reoptimized by maximizing $\ell(\mathbf{f}^G)$ over the convex hull of all support points collected from iteration $0, 1, \dots, t$,

$$\max \ell(\mathbf{f}^G) = \sum_{n=1}^N \log\{\mathbf{f}^G(n)\} \text{ s.t. } \mathbf{f}^G \in \text{conv}(\{\mathbf{f}^{\theta_0}, \tilde{\mathbf{g}}_{(0)}, \tilde{\mathbf{g}}_{(1)}, \dots, \tilde{\mathbf{g}}_{(t)}\}), \quad (17)$$

which is much easier to solve compared to the original problem (15) because the feasible region in (17) is a convex hull of finite points. The final output $\mathbf{f}_{(T)}$ is in the form

$$\pi_{\theta_0} \mathbf{f}^{\theta_0} + \pi_0 \tilde{\mathbf{g}}_{(0)} + \pi_1 \tilde{\mathbf{g}}_{(1)} + \dots + \pi_T \tilde{\mathbf{g}}_{(T-1)},$$

where $\pi_{\theta_0}, \pi_0, \dots, \pi_{T-1} \geq 0$, $\pi_{\theta_0} + \pi_0 + \dots + \pi_{T-1} = 1$, and π_t represents the corresponding proportion of the recovered consumer segment $\tilde{\mathbf{g}}_{(t)}$ for $1 \leq t \leq T-1$. Since only at most one new component is added to the mixture at each iteration of Algorithm 2, the final output of Algorithm 2 has at most T components. It is usually the case that most components only occupy very low probabilities ($< 10^{-3}$). Therefore, in practical implementation, the final output is trimmed to keep only those components with significant probabilities.

Provided that the subproblem is solved within certain error tolerance, the overall $O(1/T)$ convergence of the CG algorithm holds (Jaggi 2013). Moreover, the objective function is non-decreasing across iterations because of the reoptimization step at each iteration. Therefore, it is sufficient to approximately solve the subproblem (16) in practice.

While Algorithm 2 enjoys a theoretical guarantee due to the fact that its outer framework is based on a convex problem, the subproblem step of Algorithm 2 proves to be non-trivial. Even constrained in a simpler set $\{\mathbf{f}^\theta | \theta \in \Theta\}$, the subproblem (16) is a non-convex problem in θ (Udell and Boyd 2013, Akçakus and Mišić 2021). Moreover, the feasible region $\text{cl}(\{\mathbf{f}^\theta | \theta \in \Theta\})$ contains limiting points of $\{\mathbf{f}^\theta | \theta \in \Theta\}$. Some of these limiting points might not belong to $\{\mathbf{f}^\theta | \theta \in \Theta\}$ and represent some “extreme” consumer types corresponding to “always purchase” or “always no-purchase”. The solution to the subproblem may be one of these limiting points, which cannot be captured by a single set of parameter θ . We prove that all solutions to the subproblem can be rigorously captured by two sets of parameters (θ, ϑ) (Proposition EC.1 in Section EC.4.1 of the e-companion). Jagabathula et al. (2020) extensively study the single-period scenario of the subproblem. However, our reference effects model is different from the single-period scenario of Jagabathula et al. (2020) because it necessarily includes data from multiple periods where product features (prices) might not be stationary. Motivated by the two-parameter structure of the optimal solution, we propose a natural alternating maximization procedure for the subproblem step of Algorithm 2. We have found that this alternating maximization procedure combined with commercial non-convex solvers are effective in finding an subproblem solution that improves the objective function consistently (see details in Section EC.4.1 of the e-companion).

6. Numerical Results of Estimation and Optimization

In this section, we first show a simulation study and then use a real-world dataset from the online retailing company JD.com (Shen et al. 2020) to demonstrate the performance and practicality of our proposed framework.

We focus on the following four approaches in the numerical study:

- (i) Nonparametric MLE: This is the main approach considered in this work as previously described in Section 3 and computed by the algorithm in Section 5.

- (ii) Finite mixed logit model via EM: If the number of components is fixed and given before parameter estimation, the purchase probability is essentially modeled by a finite mixed logit model and can be estimated by the expectation-maximization (EM) algorithm (Train 2008). In contrast, the number of components in applying nonparametric MLE is not fixed but selected in a data-driven fashion.
- (iii) Single logit model: The purchase probability is specified by one fixed set of responsiveness parameter θ according to the logit model (3). The main distinction of this approach compared to (i)(ii) is that it ignores the consumer heterogeneity at the individual level.
- (iv) Aggregate linear model: The last approach is not a logit-based individual-level model but an aggregate-level demand model. In terms of aggregate demand models for reference effects, the linear demand model (18) is widely adopted in the literature (Greenleaf 1995, Chen et al. 2017, Hu et al. 2016, Chen and Nasiry 2020),

$$D_t = A - Bp_t + C_+(r_t - p_t)_+ + C_-(r_t - p_t)_-, \quad (18)$$

where A, B, C_+, C_- are four parameters that need to be estimated from the data.

6.1. Simulation Study of Prediction Accuracy

For the simulation study, we consider the canonical setting where consumers' responsiveness parameters are uniformly distributed over a region Θ , which is analogous to that of Hu and Nasiry (2017) where consumer valuations are uniformly distributed. We use k to index replications and use n to index data points within each replication. For each experiment replication, we generate a training dataset as follows. For any $k = 1, \dots, K$ and $n = 1, \dots, N$, price $p_{k,n}$ and reference discrepancy $r_{k,n} - p_{k,n}$ are independently generated from distributions μ_1 and μ_2 respectively, and a consumer responsiveness parameter vector $\theta_{k,n}$ is sampled uniformly from Θ . We take distribution μ_1 as $\max(0, \text{Normal}(1, 9))$, $\mu_2 = \text{Normal}(0, 9)$, and $\Theta = [0, 10]^4$. The purchase probability $\mathbb{P}_{k,n}$ is computed according to the logit form (3) introduced in Section 3. The purchase decision $d_{k,n}$ is then generated randomly so that $d_{k,n} = 1$ with probability $\mathbb{P}_{k,n}$, and $d_{k,n} = 0$ with probability $1 - \mathbb{P}_{k,n}$. The testing dataset is generated similarly during each replication. The sizes of the training set and the testing set are both 500 in each replication, and the number of replications is set as 100.

To evaluate the prediction performances, we use the root mean square error (RMSE) metric and the mean absolute error (MAE) metric. Given an estimated purchase probability $\hat{\mathbb{P}}_{k,n}$, the two metrics are defined as

$$\text{RMSE} = \sqrt{\frac{1}{K \times N} \sum_{k=1}^K \sum_{n=1}^N (\mathbb{P}_{k,n} - \hat{\mathbb{P}}_{k,n})^2}, \quad \text{MAE} = \frac{1}{K \times N} \sum_{k=1}^K \sum_{n=1}^N |\mathbb{P}_{k,n} - \hat{\mathbb{P}}_{k,n}|. \quad (19)$$

Different from the other three approaches (i)(ii)(iii), the aggregate linear model (iv) is defined and fitted with aggregate data. To facilitate the comparison, we aggregate the demand forecasts from the three individual-level models and compare them with the aggregate linear model under the aggregate version of the RMSE and MAE metrics. More details of data aggregation and metric definitions are described in Section EC.5.2 of the e-companion. The performances of all four approaches are summarized in Table 1, where the first three approaches are evaluated at both the individual level and the aggregate level while the linear model is only evaluated at the aggregate level. From Table 1, we conclude that nonparametric MLE approach can improve prediction accuracy significantly in the simulated settings. The performance of EM algorithm is the second-best among all approaches.

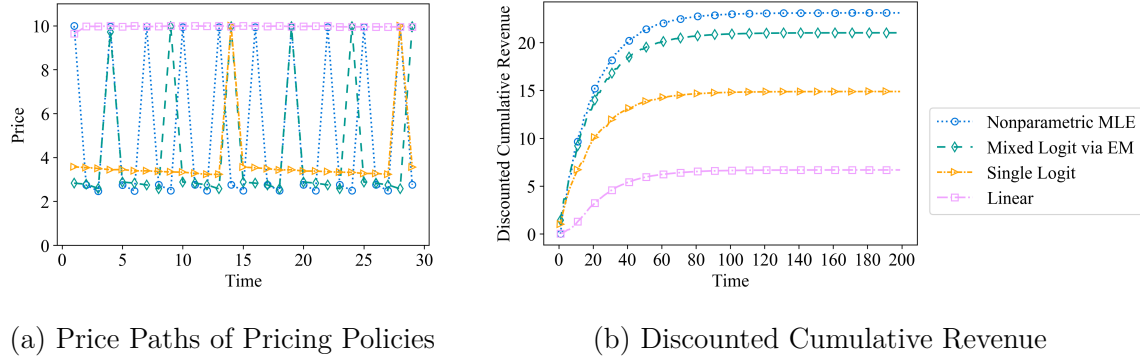
Table 1 In-sample and Out-of-sample Prediction Performances

Method	Individual level metric				Aggregate level metric			
	In-sample		Out-of-sample		In-sample		Out-of-sample	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
Nonparametric MLE	0.037	0.014	0.037	0.014	0.255	0.105	0.274	0.111
Mixed Logit via EM	0.050	0.018	0.052	0.018	0.307	0.127	0.335	0.132
Single Logit	0.074	0.031	0.073	0.031	0.468	0.196	0.544	0.214
Aggregate Linear	—	—	—	—	0.508	0.199	0.687	0.275

Subsequently, we investigate the impact of estimation accuracy on long-term revenue maximization. Given any demand estimator, we can plug in the demand estimator and calculate an optimal pricing policy using the policy iteration algorithm in Algorithm 1. Figure 5(a) shows the price trajectories of the four pricing policies, where the price range is fixed as $[0, 10]$, and the initial reference price is the average price 5.0. Figure 5(b) shows the expected discounted cumulative revenue of these pricing policies under the ground truth of uniform distribution. In Figure 5, three policies based on individual level logit models are cyclic while the pricing policy under the aggregate linear model is nearly static. Comparing three cyclic pricing policies, the pricing policies under the nonparametric MLE and EM have much shorter cycles than the single logit model, and therefore the price changes are more frequent and rapid under the former two models. This phenomenon suggests that frequent price changes are favored when more consumer heterogeneity is captured. The long-term revenue generated by the pricing policy based on our nonparametric MLE approach is higher than the others.

6.2. Case Study: Online Retailing Data from JD.com

In this section, we illustrate our approach on a real dataset provided by JD.com, a large E-commerce retailer. We estimate heterogeneous consumer responsiveness to reference effects from data and compare the predictive performances with other methods. Our numerical experiments on pricing optimization further indicate that there can be significant revenue gain by considering consumer heterogeneity.

Figure 5 Comparison of Pricing Policies

6.2.1. Data structure and data preprocessing. The JD.com dataset consists of transaction data of 31,868 products from 2.5 million consumers in March 2018 (Shen et al. 2020). For each product, there is a unique stock-keeping unit (SKU) associated with it, and we will refer to one product by its SKU directly. Two data tables in the JD.com dataset are essential to our analysis: the `clicks` table and the `orders` table. There are multiple columns and numerous rows in both tables (2 million clicks and 500k orders). To exemplify the data structure, we selectively show sample rows and relevant columns of the `clicks` and `orders` tables in Table 2 and Table 3 respectively. In addition, there is a `skus` table in the JD.com dataset containing the brand, type information (first-party or third-party product) of all SKUs.

Table 2 Sample Rows and Selected Columns from the `clicks` Table

SKU ID	user ID	request time
924eba6741	06102f7920	2018-03-01 23:23:54
6ae4031fc2	ac3e22068f	2018-03-31 10:04:02

Table 3 Sample Rows and Selected Columns from the `orders` Table

brand ID	SKU ID	user ID	order time	original price	selling price
198cec62a1	0abe9ef2ce	581d5b54c1	2018-03-01 17:14:25	89.00	79.00
9b0d3a5fc6	fb2d3d4da1	067b673f2b	2018-03-29 08:40:55	99.90	69.80

Each user corresponds to one consumer in this market. Potentially, one person may own multiple user accounts on this platform, or that multiple persons may share one user account. However, verifying such information is prohibitive, and we treat each user alone as one unique consumer. Thus, we will use the two terms, user and consumer, interchangeably as they are equivalent in our empirical study. As we can see from Table 2 and Table 3, a timestamp named “request time” is generated each time a consumer clicks one product; similarly, a timestamp named “order time” is generated each time a consumer places an order. Given any consumer, we can obtain all click and order activities

in March 2018 by querying all the rows with the corresponding user ID. Then, we can group these activities separately by SKU and arrange all the activities in the order of time for any given pair of consumer and SKU. In this way, we recover the complete shopping schedule of each consumer.

We explain how the choice information and price information are extracted. For simplicity, suppose we are looking at one consumer labeled as c . Whenever c clicks on a product labeled as s , c is deciding on whether to purchase. If c orders product s , one row is generated in the `orders` table containing the price information at that time. Otherwise, the consumer c may click s multiple times but no order is made. Since there is no price column in the `clicks` table, the price of s at that time is not directly available from the click data from consumer c . We adopt two reasonable procedures to deal with missing or redundant information in the data. First, for the price information, we query the activities from all other consumers who ordered s during a short time window containing the click time of consumer c . We then use the average price among all prices from other orders as the price that c encounters. Second, consumer c might click s multiple times during a short period when consumer c is hesitating or going back and forth to check for more detailed product information. Therefore, for any consumer and product pair, we merge all the clicks that happened within one hour.

We follow the popular exponential smoothing scheme to calculate the reference price. In doing so, a memory parameter $\alpha \in [0, 1]$ is needed. We choose a range of potential values for α and use cross-validation to choose α with the best prediction performances. For the initial reference price, we may estimate it together with the heterogeneous responsiveness parameters in our original demand model and nonparametric MLE approach. However, the statistical model may not be identifiable anymore with the extra unknown parameter (Dufour and Hsiao 2010). Instead, we impute the initial reference price as a weighted average of the original unit price and the unit selling price from the consumer's first click of this product. The unit selling price refers to the actual price that the user needs to pay for the product after all applicable price deduction due to coupons or other promotional activities. In practice, the original unit price is always shown to the consumers when they are viewing the product, so it is included in consumers' price histories.

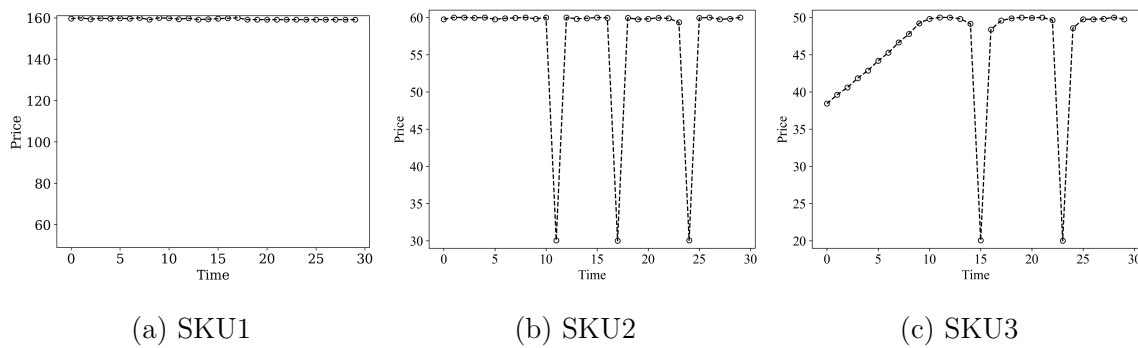
6.2.2. Estimation of heterogeneous reference effects. We focus the analysis on fast-moving products and analyze them individually. Fast-moving products are those with large demand and frequent buyers. By these criteria, we focus on SKUs $\{\text{'adfedb6893'}, \text{'3c79df1d80'}, \text{'b5cb479f7d'}\}$, and we number them as SKU1, SKU2 and SKU3. All these SKUs are first-party products, which means that they are sold by JD.com rather than third parties. For a specific product, we collect clicking and ordering information from all consumers who have made at least one order of the products. This restriction is due to the popularity of the online platform. As a lot of users click product view pages multiple times but never make a purchase, data from the entire customer group leads

to large noise. Moreover, as far as the reference effect is concerned, the retailer's main target is the consumers who visit frequently. Using the computational procedure introduced in Section 5, we compute nonparametric MLE and document the estimation results in Section EC.5.4 of the e-companion. We observe that the estimation results commonly contain multiple market segments with varying proportions, which exhibits consumer heterogeneity in reference effects. Later in Section 6.2.4, we then evaluate the prediction accuracy and revenue of the nonparametric MLE approach with other three approaches, EM algorithm for fitting finite mixed logit, single logit model, and the aggregate linear model.

6.2.3. Pricing optimization under heterogeneous reference effects. We conduct pricing optimization with heterogeneous reference effects for all three SKUs based on our estimation in Section 6.2.2. As before, the optimal pricing policies are computed by the modified policy iteration algorithm in Section 4. When computing the optimal pricing policies, the price range for each SKU is taken as the smallest interval containing the highest and the lowest unit selling price observed in the sales data of that SKU, and the discretized price range is of size 500.

Figure 6 shows the pricing policies for three SKUs. The pricing policy of SKU1 is constantly high, while the pricing policies for SKU2 and SKU3 are both cyclic. All three policies contain prices that are boundary points of the selected range. This is not uncommon for pricing optimization under reference effects, as previously shown in simulations (see Figure 3 and 4). Setting prices at the highest or lowest prices means larger price variations and may incur more demand from consumers under reference effects.

Figure 6 Optimal Price Paths for Sample SKUs



6.2.4. Evaluations of prediction accuracy and revenue. Similar to Section 6.1, we compare the performance of our nonparametric model with the other three approaches on prediction accuracy. Previously in (19), the metrics are defined using the ground truth of mean unit demand, which are not observable in practice. As a proxy, we compare the prediction result with purchase ratios,

which is the number of consumers purchasing this product divided by the number of consumers clicking this product within a certain time frame. The purchase ratios in evaluation metrics are analogous to market shares in the multi-product setting in Jagabathula et al. (2020). Aligning with the estimation procedure of the linear models, we compute the actual purchase ratios on a daily basis. The evaluation metrics are then defined as $\text{RMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^T (\hat{d}_t - \bar{d})^2}$ and $\text{MAE} = \frac{1}{T} \sum_{t=1}^T |\hat{d}_t - \bar{d}|$ where $\hat{d}_t$ represents a predicted purchase ratio and $\bar{d}$ is the average market share.

Table 4 Comparison of Prediction Accuracy

Method\Evaluation Metric	SKU1		SKU2		SKU3	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
Nonparametric MLE	0.0013	0.0011	0.0009	0.0005	0.0060	0.0046
Mixed Logit via EM	0.0016	0.0012	0.0150	0.0132	0.0213	0.0164
Single Logit	0.1348	0.1346	0.1431	0.1430	0.1494	0.1475
Aggregate Linear	0.1420	0.0401	0.3028	0.0584	0.3433	0.0948

Table 4 shows the performances of four approaches in these two metrics for all three example SKUs. It is clear from Table 4 that the nonparametric MLE approach has the best performance in all the metrics. In our implementation, EM is initiated with the same number of components as the nonparametric MLE in each setting. Therefore the model expressive power of the mixed logit model in nonparametric MLE and EM are technically the same. Yet, Table 4 indicates that nonparametric MLE brings at least 9% decrease for all six metrics of three SKUs, in particular more than 100% decrease for SKU2. On the other hand, the two approaches (single logit and aggregate linear model) that do not account for consumer heterogeneity have significantly larger errors compared to the first two approaches in Table 4.

Table 5 Long-term Revenue in Simulations

Policy\Revenue	SKU1		SKU2		SKU3	
	Mean	Median	Mean	Median	Mean	Median
Nonparametric MLE	145.117	145.119	817.221	823.848	70.617	71.497
Mixed Logit via EM	145.117	145.119	640.566	663.708	54.355	54.197
Single Logit	133.015	133.795	662.508	669.424	59.197	57.945
Aggregate Linear	145.071	145.074	391.507	436.808	1.236	1.227

We then investigate the performances of four “optimal” pricing policies based on each of the four estimation methods. The “optimal” term is relative and it means that the pricing policy is optimal with respect to the demand response environment simulated by one of the estimation results. As before, given the demand response model, the optimal pricing policies are computed using the

modified policy iteration algorithm in Section 4. Since the nonparametric MLE approach gives the best overall performance in RMSE and MAE in Table 4, it is used to simulate the true environment of market response when evaluating the pricing policies. For each SKU, we compute long-term revenues given different initial reference prices across the price range observed in the dataset, and summarize the mean and median of these long-term revenues in Table 5. The numerical results in Table 5 suggest that an increase in prediction accuracy can potentially lead to an increase in long-term revenue though the increase in revenue might not be necessarily proportional to that in prediction accuracy. The reason why the performances of all four policies for SKU1 are very similar is that the optimal pricing policy has relatively simple structure as shown in Figure 6(a). On the other hand, the increases in long-term revenue for SKU2 and SKU3 are rather significant in Table 5 with around 30% even compared with the second highest revenue. As a robustness check, we evaluate the performance of these pricing policies under the linear model environments and find that the pricing policies by nonparametric MLE are still relatively good compared to the optimal benchmark pricing policy and outperform the other two individual level approaches (see Section EC.5.4 of the e-companion).

7. Conclusion and Future Research

In this paper, we propose a framework for demand estimation and pricing under reference effects, with a novel way of incorporating consumer heterogeneity. We carefully revisit the notion of reference effects from the perspective of individual consumers, whose preferences are heterogeneous in nature. We present a nonparametric statistical methodology to simultaneously recover the heterogeneity from mixed consumer data and estimate the responsiveness parameters under reference effects. Furthermore, we draw valuable managerial insights from our theoretical analysis and numerical experiments. Moreover, we demonstrate the practicality of our proposed approach by applying it to real-world data from the E-commerce retailer JD.com. The case study on JD.com provides empirical evidence for the value of incorporating consumer heterogeneity in reference effects.

Our work points to a few interesting directions. As an application, our work may be used for personalized pricing and retargeting. Based on the estimated mixing probability distribution and consumers' past purchase behaviors, we can use Bayesian posterior to estimate a single consumer's preference under different prices. After collecting such information, the online retailer can execute personalized retargeting (e.g., by sending coupons and promotion information to selected consumers) to boost revenue. In contrast, a commonly used method for personalized retargeting is clustering. Clustering is computationally very challenging when considering all the frequent online activities. Moreover, the number of clusters needs to be predetermined. Our nonparametric modeling approach provides potential advantages in avoiding such issues.

It is also interesting to consider reference effects in the more challenging multi-product setting. In practice, consumers often compare the prices of similar products before making decisions. The

price of one product may affect reference prices of other related products (Briesch et al. 1997, Wang 2018). Under the multi-product setting, two types of reference effects need to be accounted for simultaneously – the “horizontal” effect by comparing with prices of similar products offered at the same time and the “vertical” effect by comparing with historical prices. From a data-driven perspective, the learning process could also be more computationally difficult since the number of parameters regarding cross-product effect increases exponentially with the number of products.

Our logit demand model naturally extends to the multi-product setting, which we term as “the random coefficient multinomial logit model with reference effects” and briefly discuss it as below. Assuming that there is a set S of substitutable products, the utility function of product i from set S at time t is

$$u_{i,t}(\boldsymbol{\theta}) = a - bp_{i,t} + c_+ \cdot (r_{i,t} - p_{i,t})_+ + c_- \cdot (r_{i,t} - p_{i,t})_- + \epsilon_{i,t},$$

where the random coefficient $\boldsymbol{\theta} = (a, b, c_+, c_-)$ follows the probability distribution G . The expected purchase probability of product i from one consumer is

$$\int_{\boldsymbol{\theta} \in \Theta} \frac{\exp\{u_{i,t}(\boldsymbol{\theta})\}}{1 + \sum_{j \in S} \exp\{u_{j,t}(\boldsymbol{\theta})\}} dG(\boldsymbol{\theta}).$$

In general, the reference price $r_{i,t}$ can be modeled as a function of historical and current prices of all products,

$$r_{i,t} = f(\{p_{j,\tau}\}_{j \in S, 1 \leq \tau \leq t}).$$

The choice of f may go beyond exponential smoothing in this case and should be carefully designed in a domain-specific fashion. In the multi-product and single-period setting, Wang (2018) considers several different ways of modeling reference prices, such as the highest price $\max_{j \in S} p_{j,t}$ or the lowest price $\min_{j \in S} p_{j,t}$ of all products in S . Finally, in addition to pricing optimization, one may also consider assortment optimization or joint assortment and pricing optimization.

Acknowledgments

The authors gratefully thank the editor-in-chief, the associate editor, and two anonymous referees for constructive comments that helped improve this work significantly. This research is partially supported by the National Natural Science Foundation of China (Grant 71991462) (to Z.-J. M. Shen).

References

- Akçakuş İ, Mišić V (2021) Exact logit-based product design. *Available at SSRN 3875986*.
- 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.
- Baucells M, Weber M, Welfens F (2011) Reference-point formation and updating. *Management Science* 57(3):506–519.

- 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 DP (2019) *Reinforcement learning and optimal control* (Athena Scientific Belmont, MA).
- Besbes O, Lobel I (2015) Intertemporal price discrimination: Structure and computation of optimal policies. *Management Science* 61(1):92–110.
- Briesch RA, Krishnamurthi L, Mazumdar T, Raj SP (1997) A comparative analysis of reference price models. *Journal of Consumer Research* 24(2):202–214.
- Broder J, Rusmevichientong P (2012) Dynamic pricing under a general parametric choice model. *Operations Research* 60(4):965–980.
- 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, Gallego G (2021) Nonparametric pricing analytics with customer covariates. *Operations Research* 69(3):974–984.
- 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.
- Chen YC, Mišić VV (2022) Decision forest: A nonparametric approach to modeling irrational choice. *Management Science* .
- Cohen MC, Gupta S, Kalas JJ, Perakis G (2020) An efficient algorithm for dynamic pricing using a graphical representation. *Production and Operations Management* 29(10):2326–2349.
- den Boer A, Keskin NB (2021) Dynamic pricing with demand learning and reference effects. *Management Science* Forthcoming.
- Dufour JM, Hsiao C (2010) Identification. *Microeconomics*, 65–77 (Springer).
- Farias VF, Jagabathula S, Shah D (2013) A nonparametric approach to modeling choice with limited data. *Management science* 59(2):305–322.
- Ferreira KJ, Lee BHA, Simchi-Levi D (2016) Analytics for an online retailer: Demand forecasting and price optimization. *Manufacturing & Service Operations Management* 18(1):69–88.
- 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.
- Frank M, Wolfe P, et al. (1956) An algorithm for quadratic programming. *Naval research logistics quarterly* 3(1-2):95–110.
- Greenleaf EA (1995) The impact of reference price effects on the profitability of price promotions. *Marketing Science* 14(1):82–104.

- Hardie BG, Johnson EJ, Fader PS (1993) Modeling loss aversion and reference dependence effects on brand choice. *Marketing Science* 12(4):378–394.
- Hariss R, Perakis G, Sinchaisri WP, Zheng Y (2020) Markdown pricing with quality perception. *Available at SSRN 3634129*.
- Hedgcock WM, Rao RS, Chen H (2016) Choosing to choose: The effects of decoys and prior choice on deferral. *Management Science* 62(10):2952–2976.
- 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 (2017) Are markets with loss-averse consumers more sensitive to losses? *Management Science* 64(3):1384–1395.
- Jagabathula S, Subramanian L, Venkataraman A (2020) A conditional gradient approach for nonparametric estimation of mixing distributions. *Management Science* 66(8):3635–3656.
- Jaggi M (2013) Revisiting frank-wolfe: Projection-free sparse convex optimization. *Proceedings of the 30th International Conference on Machine Learning*, 427–435.
- Kahneman D, Tversky A (1979) Prospect theory: An analysis of decision under risk. *Econometrica* 47(2):263–292.
- Kallio M, Halme M, et al. (2009) Redefining loss averse and gain seeking consumer price behavior: based on demand response. *Working paper, Helsinki School of Economics, Helsinki, Finland*.
- Keskin NB, Zeevi A (2014) Dynamic pricing with an unknown demand model: Asymptotically optimal semi-myopic policies. *Operations Research* 62(5):1142–1167.
- Kiefer J, Wolfowitz J (1956) Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters. *The Annals of Mathematical Statistics* 887–906.
- Klapper D, Ebling C, Temme J (2005) Another look at loss aversion in brand choice data: Can we characterize the loss averse consumer? *International Journal of Research in Marketing* 22(3):239–254.
- 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.
- Lindsay BG (1983) The geometry of mixture likelihoods: a general theory. *The Annals of Statistics* 86–94.
- Lindsay BG (1995) Mixture models: theory, geometry and applications. *NSF-CBMS regional conference series in probability and statistics*, i–163 (JSTOR).
- 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.
- Mazumdar T, Raj SP, Sinha I (2005) Reference price research: Review and propositions. *Journal of Marketing* 69(4):84–102.
- McFadden D (1973) Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*, 105–142 (Academic Press: New York).
- McFadden D, Train K (2000) Mixed MNL models for discrete response. *Journal of Applied Econometrics* 15(5):447–470.
- Meng XL, Rubin DB (1993) Maximum likelihood estimation via the ECM algorithm: A general framework. *Biometrika* 80(2):267–278.
- 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.
- Shen M, Tang CS, Wu D, Yuan R, Zhou W (2020) JD. com: Transaction-level data for the 2020 msom data driven research challenge. *Manufacturing & Service Operations Management* .
- Su X (2007) Intertemporal pricing with strategic customer behavior. *Management Science* 53(5):726–741.
- Train KE (2008) EM algorithms for nonparametric estimation of mixing distributions. *Journal of Choice Modelling* 1(1):40–69.
- Train KE (2009) *Discrete choice methods with simulation* (Cambridge university press).
- Udell M, Boyd S (2013) Maximizing a sum of sigmoids. *Optimization and Engineering* 1–25.
- 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 Y, Chen B, Simchi-Levi D (2021) Multimodal dynamic pricing. *Management Science* 67(10):6136–6152.
- Wang Z (2016) Intertemporal price discrimination via reference price effects. *Operations Research* 64(2):290–296.
- 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.
- Yu Q, Allon G, Bassamboo A (2021) The reference effect of delay announcements: A field experiment. *Management Science* 67(12):7417–7437.
- Zhang DJ, Dai H, Dong L, Qi F, Zhang N, Liu X, Liu Z, Yang J (2020) The long-term and spillover effects of price promotions on retailing platforms: Evidence from a large randomized experiment on alibaba. *Management Science* 66(6):2589–2609.
- 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.

E-companion for “Intertemporal Pricing via Nonparametric Estimation: Integrating Reference Effects and Consumer Heterogeneity”

EC.1. Proofs for Section 3

In this section, we present proofs of Proposition 1, and Proposition 2 in Section 3.

EC.1.1. Proof of Proposition 1

Proof of Lemma 1. When $R(x, p) < -R(-x, p)$ for any $x > 0$, we have

$$\frac{R(x, p) - R(0, p)}{x} + \frac{R(-x, p) - R(0, p)}{x} < 0, \text{ for any } x > 0,$$

where the inequality holds because $R(0, p) = 0$. Taking the limit $x \rightarrow 0$, it follows that $\rho > 1$. We can similarly get results for the other two cases. $\square$

Proof of Proposition 1. (i) The conclusion is straightforward by considering the signs of the second derivative of the logit demand function $R(x, p)$, with

$$R(x, p) = \frac{\exp(a - bp + \mathcal{C}(x))}{1 + \exp(a - bp + \mathcal{C}(x))} - \frac{\exp(a - bp)}{1 + \exp(a - bp)}.$$

Define $h(x) = \frac{\exp(a - bp + \mathcal{C}(x))}{1 + \exp(a - bp + \mathcal{C}(x))}$, then taking derivative with respect to x , we have

$$\frac{\partial R(x, p)}{\partial x} = \mathcal{C}'(x) \frac{\exp(a - bp + \mathcal{C}(x))}{(1 + \exp(a - bp + \mathcal{C}(x)))^2} = \mathcal{C}'(x) h(x) (1 - h(x))$$

and

$$\frac{\partial^2 R(x, p)}{\partial x^2} = (\mathcal{C}'(x))^2 h(x) (1 - h(x)) (1 - 2h(x)).$$

$R(x, p)$ is convex in x if and only if $(1 - 2h(x)) \geq 0$, and concave if and only if $(1 - 2h(x)) \leq 0$. It is therefore clear to verify these conditions hold.

(ii) We introduce shorthand notation $w_1 = \exp(a + c_+ x_1)$ and $w_2 = \exp(a + c_+ x_2)$, and $z = \exp(-bp)$, then we obtain the difference between $R(x_1, p)$ and $R(x_2, p)$ as

$$R(x_1, p) - R(x_2, p) = \frac{w_1 z}{1 + w_1 z} - \frac{w_2 z}{1 + w_2 z} = \frac{(w_1 - w_2) z}{(1 + w_1 z)(1 + w_2 z)} = \frac{w_1 - w_2}{1/z + w_1 w_2 z + w_1 + w_2}.$$

Without loss of generality, we assume $x_1 \geq x_2$. Since $c_+ \geq 0$, we have $w_1 - w_2 \geq 0$. Clearly, z decreases as p increases, while $R(x_1, p) - R(x_2, p)$ decreases as $1/z + w_1 w_2 z$ increases. Considering $1/z + w_1 w_2 z$ as a function of z , it is monotonically decreasing in z when $z \leq 1/\sqrt{w_1 w_2}$ and monotonically increasing in z when $z > 1/\sqrt{w_1 w_2}$. Here, $z \leq 1/\sqrt{w_1 w_2}$ is equivalent to the condition that $-2bp \leq -2a - c_+(x_1 + x_2)$. $\square$

EC.1.2. Proof of Proposition 2

Proof of Proposition 2. The proof is based on the following properties of $R(x, p)$. When $a - bp = 0$, $c_+ = c_-$, $R(x, p)$ as a function of x while fixing p , is a function with “S”-shape and symmetric with respect to the origin $x = 0$. Moreover, the absolute value of derivative of $R(x, p)$ is maximized at the origin $x = 0$ and decreases as $|x|$ increases. When $a - bp \neq 0$, we can equivalently think of the “S”-shape moved left or right. Therefore, statement (i), (ii) and (iii) follows from the basic properties of the logit function. To be more rigorous, we show a detailed proof of (i).

(i) Note that for $x > 0$, we have

$$R(x, p) + R(-x, p) = \frac{\exp(a - bp + c_+x)}{1 + \exp(a - bp + c_+x)} - \frac{\exp(a - bp)}{1 + \exp(a - bp)} + \frac{\exp(a - bp - c_-x)}{1 + \exp(a - bp - c_-x)} - \frac{\exp(a - bp)}{1 + \exp(a - bp)}.$$

When $a - bp < 0$ and $c_+ - c_- \geq 0$, we are able to bound it by

$$R(x, p) + R(-x, p) \geq \frac{\exp(a - bp + c_+x)}{1 + \exp(a - bp + c_+x)} - \frac{\exp(a - bp)}{1 + \exp(a - bp)} + \frac{\exp(a - bp - c_+x)}{1 + \exp(a - bp - c_+x)} - \frac{\exp(a - bp)}{1 + \exp(a - bp)} \quad (\text{EC.1})$$

Consider function $h(y) = \frac{\exp(a - bp + y)}{1 + \exp(a - bp + y)}$. Taking derivative with respect to y , we have $h'(y) = h(y)(1 - h(y))$. Then we can write $R(x, p) + R(-x, p)$ as

$$R(x, p) + R(-x, p) = \int_0^{c_+x} h'(y) dy - \int_{-c_+x}^0 h'(y) dy = \int_0^{c_+x} [h'(y) - h'(-y)] dy.$$

It suffices to prove that $h'(y) \geq h'(-y)$ for any $y > 0$. We can bound the difference by

$$h'(y) - h'(-y) = \frac{1}{(1 + \exp(-a + bp - y))(1 + \exp(a - bp + y))} - \frac{1}{(1 + \exp(-a + bp + y))(1 + \exp(a - bp - y))}.$$

Under a shorthand notation that $A = \exp(a - bp)$ and $B = \exp(y)$, then $A < 1$ for $a - bp < 0$ and $B > 1$ for $y > 0$, and we have

$$h'(y) - h'(-y) = \frac{1}{(1 + A^{-1}B^{-1})(1 + AB)} - \frac{1}{(1 + A^{-1}B)(1 + AB^{-1})} = \frac{AB}{(1 + AB)^2} - \frac{AB}{(A + B)^2}.$$

By comparing the denominators of the above two terms, it is clear that

$$(1 + AB)^2 - (A + B)^2 = (1 + AB + A + B)(A - 1)(B - 1) < 0.$$

Therefore we have shown that $h'(y) - h'(-y) > 0$ and the conclusion follows. The part 2) of (i) follows directly since the inequality in EC.1 is strict when $c_+ - c_- > 0$. Similarly, the conclusion of (ii)(iii) hold.

(iv) It is critical to see that the sign of $a - bp$ controls the behavioral type at large, because

$$\lim_{x \rightarrow +\infty} R(x, p) = 1 - \frac{\exp(a - bp)}{1 + \exp(a - bp)}, \quad \lim_{x \rightarrow +\infty} -R(-x, p) = \frac{\exp(a - bp)}{1 + \exp(a - bp)}.$$

We consider the scenario in (iv)(a) exclusively, since the scenario in (iv)(b) is symmetric to (iv)(a). We use shorthand notation $A = \exp(a - bp)$, $u(x) = \exp(c_+x)$, and $v(x) = \exp(-c_-x)$ for $x \geq 0$, then according to assumptions in (iv)(a), $A < 1$, $u(x) \geq 1$, $v(x) \leq 1$, $u(x)v(x) \leq 1$ and $\lim_{x \rightarrow +\infty} u(x)v(x) = 0$. By definition,

$$R(x, p) = \frac{Au(x)}{1 + Au(x)} - \frac{A}{1 + A} \quad \text{and} \quad R(-x, p) = \frac{Av(x)}{1 + Av(x)} - \frac{A}{1 + A}.$$

For any fixed p , let $S(x)$ be defined as $S(x) = R(x, p) + R(-x, p)$, then

$$S(0) = 0 \quad \text{and} \quad \lim_{x \rightarrow +\infty} S(x) = 1 - 2 \frac{A}{1+A} > 0.$$

To identify the type of consumer behavior, we want to know the relationship of $S(x)$ and 0. We take a derivative of $S(x)$ with respect to x as below:

$$\frac{dS}{dx} = \frac{\partial S}{\partial u} \frac{du}{dx} + \frac{\partial S}{\partial v} \frac{dv}{dx} = \frac{c_-}{(\sqrt{Au(x)} + \frac{1}{\sqrt{Au(x)}})^2} \left[\frac{c_+}{c_-} - \left(\frac{A\sqrt{u(x)} + \frac{1}{\sqrt{u(x)}}}{A\sqrt{v(x)} + \frac{1}{\sqrt{v(x)}}} \right)^2 \right]. \quad (\text{EC.2})$$

We now show that $\left(\frac{A\sqrt{u(x)} + \frac{1}{\sqrt{u(x)}}}{A\sqrt{v(x)} + \frac{1}{\sqrt{v(x)}}} \right)^2$ is non-increasing in x . Define $g(x) = \frac{A\sqrt{u(x)} + \frac{1}{\sqrt{u(x)}}}{A\sqrt{v(x)} + \frac{1}{\sqrt{v(x)}}}$, then taking derivative with respect to x , we have

$$\frac{dg(x)}{dx} = \frac{-(c_+ + c_-)(\frac{1}{\sqrt{u(x)v(x)}} - A^2\sqrt{u(x)v(x)}) - (c_- - c_+)(A\frac{\sqrt{u(x)}}{\sqrt{v(x)}} - A\frac{\sqrt{v(x)}}{\sqrt{u(x)}})}{2(A\sqrt{v(x)} + \frac{1}{\sqrt{v(x)}})^2}. \quad (\text{EC.3})$$

From the condition that $A < 1$, $u(x) \geq 1, v(x) \leq 1, u(x)v(x) \leq 1$, we have inequalities

$$\frac{1}{\sqrt{u(x)v(x)}} - A^2\sqrt{u(x)v(x)} \geq 0 \quad \text{and} \quad A\frac{\sqrt{u(x)}}{\sqrt{v(x)}} - A\frac{\sqrt{v(x)}}{\sqrt{u(x)}} \geq 0.$$

Combining with the condition $c_- > c_+ \geq 0$, we conclude that $\frac{dg(x)}{dx} \leq 0$, and the inequality is strict when $x > 0$. Since

$$\left(\frac{A\sqrt{u(x)} + \frac{1}{\sqrt{u(x)}}}{A\sqrt{v(x)} + \frac{1}{\sqrt{v(x)}}} \right)^2 \Big|_{x=0} = 1 \quad \text{and} \quad \lim_{x \rightarrow +\infty} \left(\frac{A\sqrt{u(x)} + \frac{1}{\sqrt{u(x)}}}{A\sqrt{v(x)} + \frac{1}{\sqrt{v(x)}}} \right)^2 = 0,$$

we have shown that $g^2(x) = \left(\frac{A\sqrt{u(x)} + \frac{1}{\sqrt{u(x)}}}{A\sqrt{v(x)} + \frac{1}{\sqrt{v(x)}}} \right)^2$ is decreasing from 1 to 0 as x goes from 0_+ to $+\infty$.

Correspondingly, according to (EC.2), combined with the condition that $\frac{c_+}{c_-} \in [0, 1)$, the derivative $\frac{dS(x)}{dx}$ goes from negative to positive. Let x' be the point satisfying $\frac{dS(x)}{dx} = 0$.

To sum up, $S(x)$ is strictly decreasing for $x \in (0, x')$ and then strictly increasing for $x \in (x', \infty)$. Combining with $S(0) = 0, \lim_{x \rightarrow +\infty} S(x) = 1 - 2 \frac{A}{1+A} > 0$, so there is a unique cross point $x^* \in (x', \infty)$ of the curve $y = S(x)$ and the horizontal line $y = 0$ in the two-dimensional space. Thus, we have completed our proof of (iv)(a), and (iv)(b) is a symmetric result of (iv)(a). $\square$

EC.2. Results on Myopic Pricing Policy

In this section, we present our results on the myopic pricing policy, which include the proof of Proposition 3 on the performance guarantee of myopic pricing policy and a simulation study.

EC.2.1. Performance Guarantee of Myopic Pricing Policy

A popular heuristic for the pricing optimization problem is the myopic pricing policy, denoted by p_m . Instead of solving the Bellman equation (9), the myopic pricing policy is defined by attaining the maximum for the single-period revenue during each period t , i.e.,

$$p_m(r) = \arg \max_{p \in \mathcal{P}} \Pi(r, p). \quad (\text{EC.4})$$

The pricing optimization problem (EC.4) is easy to solve since it is a univariate optimization problem. In the case where $\mathcal{P}$ is assumed to be finite, this can be solved by enumeration with time complexity $O(|\mathcal{P}|)$. The myopic policy is thus favorable since it is computationally effortless.

The long-term revenue under the myopic pricing policy is

$$V_m(r) = \sum_{t=0}^{\infty} \beta^t \Pi(r_t, p_m(r_t)), \quad (\text{EC.5})$$

where $r_t = \alpha r_{t-1} + (1 - \alpha)p_m(r_{t-1})$ for all $t \geq 1$, and $r_0 = r$.

In Proposition 3, we quantitatively bound the long-term revenue gap between the optimal pricing policy and the myopic pricing policy.

Proposition 3 (Restated) *Let V^* be the long-term revenue of the optimal pricing policy, defined as in (8) and let V_m be the long-term revenue of the myopic pricing policy, defined as in (EC.5). It holds for any initial reference price r that*

$$0 \leq V^*(r) - V_m(r) \leq \frac{\beta(1 - \alpha)}{(1 - \alpha\beta)(1 - \beta)} \eta p_H, \quad (\text{EC.6})$$

where $\eta := \min \left\{ 1, \sup_{(a, b, c_+, c_-) \in \Theta} \frac{\max(c_+, c_-)}{b + c_-} \right\}$.

REMARK EC.1. We make the following interpretations on the upper bound in Proposition 3. When $\alpha = 1$ or $\beta = 0$, the upper bound becomes 0, which indicates that the myopic pricing policy is optimal. This is true because the reference price is never updated when the memory parameter α is set to be 1, and the long-term revenue becomes the single-period revenue with discount parameter $\beta = 0$. The value of η depends on the responsiveness parameters among all the consumers. Recall that c_+ and c_- are the responsiveness parameters for reference prices and b is the responsiveness parameter for prices. When c_+ and c_- get larger, the reference effects become stronger. Since the myopic pricing policy neglects reference effects, it is expected that the revenue gap becomes larger.

EC.2.2. Proof of Proposition 3

In order to prove Proposition 3, we first present and prove Lemma EC.1, which provides key ingredients to our proof of Proposition 3. We use Lemma EC.1 to control the variation of maximal single-period revenue under perturbed reference prices.

LEMMA EC.1. *For two single-period revenue given reference price r and $\tilde{r}$ respectively, the maximum single-period revenue gap satisfies the following inequality:*

$$|\max_{q \in \mathcal{P}} \Pi(\tilde{r}, q) - \max_{p \in \mathcal{P}} \Pi(r, p)| \leq \eta |\tilde{r} - r|,$$

where $\eta = \min \left\{ 1, \sup_{(a, b, c_+, c_-) \in \Theta} \frac{\max(c_+, c_-)}{b + c_-} \right\}$.

Proof of Lemma EC.1. Without loss of generality, we assume that $\tilde{r} > r$. Given any $\epsilon > 0$, we claim that it suffices to prove the inequality in Lemma EC.1 for $\tilde{r} - r < \epsilon$. Let $r_1, \dots, r_N$ be a price sequence between r and $\tilde{r}$ such that $r = r_1 < r_2 < \dots < r_N = \tilde{r}$ and $r_{i+1} - r_i \leq \epsilon$ for any $1 \leq i \leq N$. Suppose Lemma EC.1 holds whenever $\tilde{r} - r < \epsilon$, then

$$\left| \max_q \Pi(r_{i+1}, q) - \max_p \Pi(r_i, p) \right| \leq \eta |r_{i+1} - r_i|.$$

Therefore, we are able to decompose the revenue gap $\max_p \Pi(\tilde{r}, p) - \max_p \Pi(r, p)$ to

$$\begin{aligned} \max_p \Pi(\tilde{r}, p) - \max_p \Pi(r, p) &= \sum_{i=1}^{N-1} \left(\max_p \Pi(r_{i+1}, p) - \max_p \Pi(r_i, p) \right) \\ &\leq \sum_{i=1}^{N-1} \left| \max_p \Pi(r_{i+1}, p) - \max_p \Pi(r_i, p) \right| \leq \eta \sum_{i=1}^{N-1} |r_{i+1} - r_i| = \eta |\tilde{r} - r|. \end{aligned}$$

Now let $\epsilon = \frac{1}{2} \inf_{(a, b, c_+, c_-) \in \Theta} \frac{1}{b + \max(c_+, c_-)}$. It suffices to prove the revenue gap in the statement for any $\tilde{r}$ and r such that $\tilde{r} - r < \epsilon$. In the following proof, we assume $\tilde{r} - r < \epsilon$. We denote the maximizers of single-period revenue function $\Pi(\tilde{r}, p)$ and $\Pi(r, p)$ by p^* and $\tilde{p}^*$ respectively, i.e.,

$$p^* \in \arg \max_p \Pi(r, p) \quad \text{and} \quad \tilde{p}^* \in \arg \max_q \Pi(\tilde{r}, q).$$

We first prove that the price is sub-optimal if $p < \inf_{(a, b, c_+, c_-) \in \Theta} \frac{1}{b + \max(c_+, c_-)}$. We use $\frac{\partial \Pi(r, p)}{\partial p}$ to denote sub-gradient in case the revenue function is not smooth. Note that,

$$\begin{aligned} \frac{\partial \Pi(r, p)}{\partial p} &= \int_{\theta \in \Theta} \left(-(b + C'(\tilde{r} - p))p \frac{\exp(a - bp + C(\tilde{r} - p))}{[1 + \exp(a - bp + C(\tilde{r} - p))]^2} + \frac{\exp(a - bp + C(\tilde{r} - p))}{1 + \exp(a - bp + C(\tilde{r} - p))} \right) dG(\theta) \\ &= \int_{\theta \in \Theta} \frac{\exp(a - bp + C(\tilde{r} - p))}{1 + \exp(a - bp + C(\tilde{r} - p))} \left(1 - \frac{(b + C'(\tilde{r} - p))p}{1 + \exp(a - bp + C(\tilde{r} - p))} \right) dG(\theta) \\ &\geq \int_{\theta \in \Theta} \frac{\exp(a - bp + C(\tilde{r} - p))}{1 + \exp(a - bp + C(\tilde{r} - p))} \left(1 - \frac{1}{1 + \exp(a - bp + C(\tilde{r} - p))} \right) dG(\theta) \\ &> 0, \end{aligned}$$

which indicates that the price is not optimal. Now we construct a revenue lower bound for $\tilde{p}^* > \inf_{(a, b, c_+, c_-) \in \Theta} \frac{1}{b + \max(c_+, c_-)}$. Define

$$\hat{p} = \tilde{p}^* - \eta(\tilde{r} - r), \tag{EC.7}$$

where

$$\eta = \min \left(1, \sup_{(a,b,c_+,c_-) \in \Theta} \frac{\max(c_+, c_-)}{b + c_-} \right).$$

We will show that $\hat{p}$ can reach higher revenue than p^* when the reference price is $\tilde{r}$, i.e.,

$$(a - b\tilde{p}^* + \mathcal{C}(\tilde{r} - \tilde{p}^*)) - (a - b\hat{p} + \mathcal{C}(r - \hat{p})) \leq 0.$$

We prove the conclusion for two scenarios where $\sup_{(a,b,c_+,c_-) \in \Theta} \frac{\max(c_+, c_-)}{b + c_-} \geq 1$ and $\sup_{(a,b,c_+,c_-) \in \Theta} \frac{\max(c_+, c_-)}{b + c_-} < 1$, respectively.

1. Case $\eta = 1$: Equation $r - \hat{p} = \tilde{r} - \tilde{p}^*$ holds when $\eta = 1$, thus it holds that

$$(a - b\tilde{p}^* + \mathcal{C}(\tilde{r} - \tilde{p}^*)) - (a - b\hat{p} + \mathcal{C}(r - \hat{p})) = -b(\tilde{p}^* - \hat{p}) \leq 0.$$

2. Case $\eta = \sup_{(a,b,c_+,c_-) \in \Theta} \frac{\max(c_+, c_-)}{b + c_-} < 1$:

- (a) When $(r - \hat{p})(\tilde{r} - \tilde{p}^*) > 0$ or $c_+ = c_-$, we have

$$(a - b\tilde{p}^* + \mathcal{C}(\tilde{r} - \tilde{p}^*)) - (a - b\hat{p} + \mathcal{C}(r - \hat{p})) \leq \max \{ (-b\eta + c_+ - c_+\eta)(\tilde{r} - r), (-b\eta + c_- - c_-\eta)(\tilde{r} - r) \}.$$

- (b) When $c_+ \neq c_-$ and $(r - \hat{p})(\tilde{r} - \tilde{p}^*) < 0$. By our choice of $\hat{p}$ and the condition that $\eta < 1$, we have

$$r - \hat{p} = \tilde{r} - \tilde{p}^* - (1 - \eta)(\tilde{r} - r) \leq \tilde{r} - \tilde{p}^*.$$

It implies that $r - \hat{p} \leq 0 \leq \tilde{r} - \tilde{p}^*$. Note that $r \leq \hat{p} \leq \tilde{p}^* \leq \tilde{r}$. Consider the difference between two utility functions,

$$\begin{aligned} & (a - b\tilde{p}^* + \mathcal{C}(\tilde{r} - \tilde{p}^*)) - (a - b\hat{p} + \mathcal{C}(r - \hat{p})) = -b\eta(\tilde{r} - r) + c_+(\tilde{r} - \tilde{p}^*) - c_-(r - \hat{p}) \\ & = (c_+ - c_-\eta - b\eta)(\tilde{r} - r) - (c_+ - c_-)(\tilde{p}^* - r). \end{aligned}$$

We consider $c_+ \geq c_-$ and $c_- \leq c_+$ separately.

- i. If $c_+ \geq c_-$, then $(c_+ - c_-\eta - b\eta)(\tilde{r} - r) \leq 0$ and $-(c_+ - c_-)(\tilde{p}^* - r) \leq 0$, and we reach the conclusion.
- ii. Otherwise, if $c_+ \leq c_-$, then

$$\begin{aligned} & (c_+ - c_-\eta - b\eta)(\tilde{r} - r) - (c_+ - c_-)(\tilde{p}^* - r) \leq (c_+ - c_-\eta - b\eta + c_- - c_+)(\tilde{r} - r) \\ & = (c_-(1 - \eta) - b\eta)(\tilde{r} - r) = (c_- - \eta(c_- + b))(\tilde{r} - r) \leq 0, \end{aligned}$$

where the first inequality follows from $c_- - c_+ \geq 0$ and $\tilde{p}^* - r \leq \tilde{r} - r$.

To sum up, it follows that for $\eta = \min\{1, \sup_{(a,b,c_+,c_-) \in \text{supp}(G)} \frac{\max(c_+, c_-)}{b + c_-}\}$,

$$\int_{\theta \in \Theta} \frac{\exp(a - b\tilde{p}^* + \mathcal{C}(\tilde{r} - \tilde{p}^*))}{1 + \exp(a - b\tilde{p}^* + \mathcal{C}(\tilde{r} - \tilde{p}^*))} dG(\theta) \leq \int_{\theta \in \Theta} \frac{\exp(a - b\hat{p} + \mathcal{C}(r - \hat{p}))}{1 + \exp(a - b\hat{p} + \mathcal{C}(r - \hat{p}))} dG(\theta).$$

Therefore,

$$\begin{aligned} \left| \max_q \Pi(\tilde{r}, q) - \max_p \Pi(r, p) \right| &= \Pi(\tilde{r}, \tilde{p}^*) - \Pi(r, p^*) \leq \Pi(\tilde{r}, \tilde{p}^*) - \Pi(r, \hat{p}) \\ &\leq (\tilde{p}^* - \hat{p}) \cdot \int_{\theta \in \Theta} \frac{\exp(a - b\hat{p} + \mathcal{C}(r - \hat{p}))}{1 + \exp(a - b\hat{p} + \mathcal{C}(r - \hat{p}))} dG(\theta) \\ &\leq \eta |\tilde{r} - r|. \end{aligned}$$

This completes our proof. $\square$

Proof of Proposition 3. Recall that the optimal pricing policy and myopic pricing policy with respect to reference price r are denoted by $p^*(r)$ and $p_m(r)$ respectively. With the notation, for optimal pricing

$$V^*(r) = \sum_{t=0}^{\infty} \beta^t \Pi(r_t^*, p^*(r_t^*)), \quad (\text{EC.8})$$

where $r_t^* = \alpha r_{t-1}^* + (1 - \alpha)p^*(r_{t-1}^*)$ for all $t \geq 1$, and $r_0^* = r$. Based on the recursive definition of r_t^* , r_t^* can be expressed as

$$r_t^* = \alpha^t r + \alpha^{t-1}(1 - \alpha)p^*(r_0^*) + \cdots + (1 - \alpha)p^*(r_{t-1}^*) = \alpha^t r + (1 - \alpha) \sum_{\tau=0}^{t-1} \alpha^{t-1-\tau} p^*(r_\tau^*), \quad \forall t \geq 0. \quad (\text{EC.9})$$

Similarly, for myopic pricing,

$$V_m(r) = \sum_{t=0}^{\infty} \beta^t \Pi(r_{m,t}, p_m(r_{m,t})) \quad (\text{EC.10})$$

where $r_{m,t} = \alpha r_{m,t-1} + (1 - \alpha)p_m(r_{m,t-1})$ for all $t \geq 1$, and $r_{m,0} = r$. And $r_{m,t}$ can be expressed as

$$r_{m,t} = \alpha^t r + \alpha^{t-1}(1 - \alpha)p_m(r_{m,0}) + \cdots + (1 - \alpha)p_m(r_{m,t-1}) = \alpha^t r + (1 - \alpha) \sum_{\tau=0}^{t-1} \alpha^{t-1-\tau} p_m(r_{m,\tau}), \quad \forall t \geq 0. \quad (\text{EC.11})$$

Combining (EC.9) and (EC.11), it follows that

$$|r_t^* - r_{m,t}| \leq (1 - \alpha) \sum_{\tau=0}^{t-1} \alpha^{t-1-\tau} |p^*(r_\tau^*) - p_m(r_{m,\tau})| \leq (1 - \alpha^t) p_H.$$

We now utilize Lemma EC.1 to finish our proof. Following (EC.8) and (EC.10), for any $r \in \mathcal{P}$,

$$\begin{aligned} V^*(r) - V_m(r) &= \sum_{t=0}^{\infty} \beta^t \Pi(r_t^*, p^*(r_t^*)) - \sum_{t=0}^{\infty} \beta^t \Pi(r_{m,t}, p_m(r_{m,t})) \\ &\leq \sum_{t=0}^{\infty} \beta^t \Pi(r_t^*, p_m(r_t^*)) - \sum_{t=0}^{\infty} \beta^t \Pi(r_{m,t}, p_m(r_{m,t})) \quad (\text{by definition of } p_m) \\ &= \sum_{t=1}^{\infty} \beta^t \Pi(r_t^*, p_m(r_t^*)) - \sum_{t=1}^{\infty} \beta^t \Pi(r_{m,t}, p_m(r_{m,t})) \quad (\text{because } r_0^* = r_{m,0} = r) \\ &= \sum_{t=1}^{\infty} \beta^t [\max_q \Pi(r_t^*, q) - \max_p \Pi(r_{m,t}, p)] \\ &\leq \sum_{t=1}^{\infty} \beta^t \eta |r_t^* - r_{m,t}| \quad (\text{by Lemma EC.1}) \\ &\leq \sum_{t=1}^{\infty} \beta^t (1 - \alpha^t) \eta p_H. \end{aligned}$$

Reorganizing the last term, it yields that

$$V^*(r) - V_m(r) \leq \sum_{t=1}^{\infty} \beta^t (1 - \alpha^t) \eta p_H = \frac{\eta \beta (1 - \alpha)}{(1 - \alpha \beta)(1 - \beta)} p_H,$$

which is the desired bound. $\square$

EC.2.3. Simulation Study of Myopic Pricing Policy

We present a simulation study of the myopic pricing policy and observe that its performance may deteriorate as reference effects grow. As defined in the maximization problem (EC.4), the myopic pricing policy maximizes the single period revenue but neglects the influences of the current price on future revenue. In Proposition 3, we provide an upper bound on the long-term revenue gap between the myopic pricing and the optimal pricing policy. Proposition 3 postulates that the revenue gap could be larger when the reference effects have a stronger influence on the revenue, as captured by the parameter η in (EC.6). To further illustrate the performance of the myopic pricing policy, we consider a case where the consumer responsiveness parameter c_+ varies but parameters $a = b = c_- = 2$ are fixed. Under the same choices of α, β, ϵ as before, we compute the average revenue gap ratio defined as $\frac{1}{|\mathcal{R}_\epsilon|} \sum_{r \in \mathcal{R}_\epsilon} \frac{V^*(r) - V_m(r)}{V^*(r)}$ for different c_+ and summarize the results in Table EC.1.

Table EC.1 Revenue Gap of Myopic Pricing Policy

Responsiveness Parameter c_+	1	1.5	2	2.5	3	3.5	4	4.5	5
Average Revenue Gap Ratio	0.290	0.378	0.423	0.447	0.460	0.467	0.471	0.472	0.472

Table EC.1 shows that that myopic pricing policy may lead to approximately 40% revenue loss compared to the optimal pricing policy. The myopic pricing policy may perform relatively better when the consumers are less gain-seeking such as when $c_+ = 1$. Overall, we can observe an increasing trend from Table EC.1 — the average revenue gap ratio increases as the magnitude of reference effects increases.

EC.3. Proofs for Section 4

In this section, we present proofs of Proposition 4, Proposition 5 and Proposition 6 in Section 4. The proof of Proposition 3 is in Section EC.2.

EC.3.1. Proof of Proposition 4

Proof of Proposition 4. The monotonicity of $\Pi(r, p)$ in r is clear from the non-negativity of b in the definition of $\Pi(r, p)$. Given r , we assume pricing policy $\{p_t^*\}_{t=1}^\infty$ achieves optimal long-term revenue $V^*(r)$. Let the associated reference price sequence be defined as $r_{t+1}^* = \alpha r_t^* + (1 - \alpha)p_t^*, t = 0, 1, 2, \dots$, where $r_0 = r$. For any reference price $\tilde{r}$, we define $\tilde{r}_t$ as $\tilde{r}_{t+1} = \alpha \tilde{r}_t + (1 - \alpha)p_t^*, t = 0, 1, 2, \dots$, where $\tilde{r}_0 = \tilde{r}$. Notice that $\tilde{r}_t$ and r_t^* are defined similarly, and the only difference lies in the initial reference price $\tilde{r} \geq r$, and therefore $\tilde{r}_t \geq r_t^*$ for all $t = 0, 1, 2, \dots$. Then the long-term revenue given initial reference price $\tilde{r}$ and pricing policy $\{p_t^*\}_{t=1}^\infty$ is $\sum_{t=0}^\infty \beta^t \Pi(\tilde{r}_t, p_t^*)$. By the monotonicity of the function $\Pi(r, p)$ in r , we conclude that $V^*(\tilde{r}) \geq \sum_{t=0}^\infty \beta^t \Pi(\tilde{r}_t, p_t^*) \geq \sum_{t=0}^\infty \beta^t \Pi(r_t^*, p_t^*) = V^*(r)$. Thus we reach our conclusion. $\square$

EC.3.2. Proof of Proposition 5

Proof of Proposition 5. We prove the conclusion by contradiction. Suppose there exists an optimal steady price p^* . Define the corresponding optimal revenue as

$$\pi^* = p^* \int_{\theta \in \Theta} \frac{\exp(a - bp^*)}{1 + \exp(a - bp^*)} dG(\theta).$$

Then the cumulative discounted revenue is $\pi^*/(1 - \beta)$. We prove the result by constructing a better pricing policy. Let q be a price lower than p^* and define δ as the ratio of q and p^* , then δ is a number in $(0, 1)$. According to our assumption, since for all $(a, b, c_+, c_-) \in \Theta$, it satisfies the condition that $\exp(a - bp^*) \leq \exp(a) < (1 - \beta)/\beta$, we can choose q such that

$$\sup_{(a, b, c_+, c_-) \in \Theta} \frac{\exp(a - bp^*)}{(1 - \beta)[1 + \exp(a - bp^*)]} < \delta < 1.$$

Suppose the given initial reference price is r_0 . We choose N large enough so that

$$N > \left\lceil \log_{1/\alpha} \left(\frac{p^* - r_0}{(1 - \delta)p^*} \right) \mathbf{1}\{p^* - r_0 > 0\} \right\rceil + 1, \quad (\text{EC.12})$$

where notation $\lfloor x \rfloor$ represents the largest integer that is less than or equal to number x . We consider the long-term revenue of the policy that price q is offered after price p^* being offered for N times consecutively. Notice that the reference price for time $N + 1$ is $\alpha^N r_0 + (1 - \alpha^N)p^*$, and the revenue at period $N + 1$ is

$$\Pi(p^*, q) = q \int_{\theta \in \Theta} \frac{\exp(a - bq + c_+(\alpha^N r_0 + (1 - \alpha^N)p^* - q))}{1 + \exp(a - bq + c_+(\alpha^N r_0 + (1 - \alpha^N)p^* - q))} dG(\theta).$$

We can rewrite the utility function for time $N + 1$ as

$$u_{N+1} = a - bp^* + b(1 - \delta)p^* + c_+((1 - \delta)p^* + \alpha^N(r_0 - p^*)). \quad (\text{EC.13})$$

According to our choice of N in (EC.12), the quantity $(1 - \delta)p^* + \alpha^N(r_0 - p^*)$ is positive. Therefore, the utility defined in (EC.13) goes to infinity when c_+ approaches infinity. Notice that the definition of π^* does not depend on c_+ since π^* is a constant policy. Thus,

$$\delta \int_{\theta \in \Theta} \frac{\exp(a - bq + c_+(\alpha^N r_0 + (1 - \alpha^N)p^* - q))}{1 + \exp(a - bq + c_+(\alpha^N r_0 + (1 - \alpha^N)p^* - q))} dG(\theta)$$

approaches to δ as c_+ goes to infinity, which could achieve a larger value compared to $\pi^*/[p^*(1 - \beta)]$. Therefore, for sufficiently large c_+ , we have $\Pi(p^*, q) > \pi^*/(1 - \beta)$. The new pricing policy produces a higher revenue since the revenue generated at time $N + 1$ is higher than the long-term revenue from stationary price p^* since time $N + 1$, which conflicts with the optimality of p^* . Specifically, if $r_0 = p^*$, then we have shown that the next optimal price is not p^* , which means that any price p^* cannot be a steady state. This concludes the proof. $\square$

EC.3.3. Proof of Proposition 6

Proof of Proposition 6. Let p^* denote the optimal pricing policy. We define a pricing policy $\hat{p}_\epsilon$ for the discretized pricing optimization problem as follows. For any $r \in \mathcal{R}_\epsilon$, $\hat{p}(r)$ is defined as

$$\begin{aligned} \hat{p}(r) &\in \arg \min_{p \in \mathcal{P}} |p - p^*(r)| \\ \text{s.t. } &\alpha r + (1 - \alpha)p \in \mathcal{R}_\epsilon. \end{aligned}$$

In case there are multiple minimizers to the optimization problem above, $\hat{p}(r)$ is chosen as any of the minimizers. By definition of $\mathcal{R}_\epsilon$ and $\hat{p}(r)$, the reference price gap $|(\alpha r + (1 - \alpha)\hat{p}(r)) - (\alpha r + (1 - \alpha)p^*(r))| \leq \epsilon$, and therefore $|\hat{p}(r) - p^*(r)| \leq \frac{\epsilon}{1 - \alpha}$. By optimality of V^* and p^* , the Bellman's equation holds, i.e., $V^*(r) = \Pi(r, p^*(r)) + \beta V^*(\alpha r + (1 - \alpha)p^*(r))$. By optimality of V_ϵ^* ,

$$\begin{aligned} V_\epsilon^*(r) &= \max_{\alpha r + (1 - \alpha)q \in \mathcal{R}_\epsilon} \{\Pi(r, q) + \beta V_\epsilon^*(\alpha r + (1 - \alpha)q)\} \\ &\geq \Pi(r, \hat{p}(r)) + \beta V_\epsilon^*(\alpha r + (1 - \alpha)\hat{p}(r)). \end{aligned}$$

Therefore,

$$\begin{aligned} V^*(r) - V_\epsilon^*(r) &\leq \Pi(r, p^*(r)) + \beta V^*(\alpha r + (1 - \alpha)p^*(r)) - \Pi(r, \hat{p}(r)) - \beta V_\epsilon^*(\alpha r + (1 - \alpha)\hat{p}(r)) \\ &= (\Pi(r, p^*(r)) - \Pi(r, \hat{p}(r))) + \beta (V^*(\alpha r + (1 - \alpha)p^*(r)) - V_\epsilon^*(\alpha r + (1 - \alpha)\hat{p}(r))) \\ &\quad + \beta (V^*(\alpha r + (1 - \alpha)\hat{p}(r)) - V_\epsilon^*(\alpha r + (1 - \alpha)\hat{p}(r))) \\ &\leq (\Pi(r, p^*(r)) - \Pi(r, \hat{p}(r))) + \beta (V^*(\alpha r + (1 - \alpha)p^*(r)) - V^*(\alpha r + (1 - \alpha)\hat{p}(r))) \\ &\quad + \beta \max_{\tilde{r} \in \mathcal{R}_\epsilon} \{V^*(\tilde{r}) - V_\epsilon^*(\tilde{r})\}, \end{aligned} \tag{EC.14}$$

where the last inequality follows from the fact that $\alpha r + (1 - \alpha)\hat{p}(r) \in \mathcal{R}_\epsilon$. For the simplicity of notation, we define

$$D_1 = \Pi(r, p^*(r)) - \Pi(r, \hat{p}(r)) \quad \text{and} \quad D_2 = V^*(\alpha r + (1 - \alpha)p^*(r)) - V^*(\alpha r + (1 - \alpha)\hat{p}(r)).$$

We first prove uniform bounds on D_1 and D_2 that do not depend on choice of r . Note that we are using slightly different proof technique here compared to proof of Proposition 3. The main reason is that we want to make sure the final bound depends linearly on ϵ without an intercept term. For D_1 , the proof idea is using Lipschitz property, and for D_2 , we consider value iteration and prove by induction. Note that

$$\begin{aligned} D_1 &= \Pi(r, p^*(r)) - \Pi(r, \hat{p}(r)) \\ &\leq \sup_{p \in [p^*(r), \hat{p}(r)]} \int_{\theta \in \Theta} \frac{\exp(a - bp + \mathcal{C}(r - p))}{(1 + \exp(a - bp + \mathcal{C}(r - p)))^2} \mathcal{C}'(r - p) dG(\theta) \cdot |p^*(r) - \hat{p}(r)| \\ &\leq \frac{1}{4} \int_{\theta \in \Theta} \max(b + c_+, b + c_-) \frac{\epsilon}{1 - \alpha} dG(\theta), \end{aligned}$$

where the first inequality follows from the Lipschitz property of differentiable functions, and the second inequality follows from the basic inequality that $\frac{y}{(1+y)^2} \leq \frac{1}{4}$ for any $y > 0$ and the fact that $|\hat{p}(r) - p^*(r)| \leq \frac{\epsilon}{1-\alpha}$. Namely, we show that γ_1 can be taken as $\int_{\theta \in \Theta} \max(b + c_+, b + c_-) \frac{1}{4(1-\alpha)} dG(\theta)$.

Consider the following value iteration, $V^{(k+1)}(r) = \max_p \{\Pi(r, p) + \beta V^{(k)}(\alpha r + (1-\alpha)p)\}$. Now we prove $V^{(k)}(\tilde{r}) - V^{(k)}(r) \leq \gamma_2 |\tilde{r} - r|$ by induction. The induction holds trivially for $V^{(0)}$, and furthermore

$$\begin{aligned} V^{(k+1)}(\tilde{r}) - V^{(k+1)}(r) &\leq \max_p \{\Pi(r, p) - \Pi(\tilde{r}, p)\} + \beta \max_p \{V^{(k)}(\alpha \tilde{r} + (1-\alpha)p) - V^{(k)}(\alpha r + (1-\alpha)p)\} \\ &\leq \frac{1}{4} |\tilde{r} - r| \int_{\theta \in \Theta} \max(c_+, c_-) dG(\theta) + \beta \gamma_2 \alpha |\tilde{r} - r|, \end{aligned}$$

where the first inequality follows from the triangle inequality of maximums, the first bound in the last line follows from the same proof idea as when we prove D_1 , and the second in the last line is by induction hypothesis. Therefore, we have

$$V^{(k+1)}(\tilde{r}) - V^{(k+1)}(r) \leq \left(\frac{1}{4} \int_{\theta \in \Theta} \max(c_+, c_-) dG(\theta) + \beta \gamma_2 \alpha \right) |\tilde{r} - r|.$$

We set $\gamma_2 = \frac{1}{4(1-\alpha\beta)} \int_{\theta \in \Theta} \max(c_+, c_-) dG(\theta)$ then the proof by induction is completed. Taking the limit of k , and it becomes $V^*(\tilde{r}) - V^*(r) \leq \gamma_2 |\tilde{r} - r|$. Plugging this into D_2 , it follows that $D_2 \leq \gamma_2 \epsilon$. Now that we have shown $D_1 \leq \gamma_1 \epsilon$ and $D_2 \leq \gamma_2 \epsilon$ where γ_1, γ_2 are some parameters independent of r , then for any $r \in \mathcal{R}_\epsilon$, (EC.14) yields that

$$V^*(r) - V_\epsilon^*(r) \leq (\gamma_1 + \beta \gamma_2) \epsilon + \beta \max_{\tilde{r} \in \mathcal{R}_\epsilon} \{V^*(\tilde{r}) - V_\epsilon^*(\tilde{r})\}. \quad (\text{EC.15})$$

Since the above inequality (EC.15) holds for any r while the right-hand side of (EC.15) does not depend on the initial reference price r , taking the maximum over r , the same bound holds uniformly, i.e.,

$$\max_{r \in \mathcal{R}_\epsilon} V^*(r) - V_\epsilon^*(r) \leq (\gamma_1 + \beta \gamma_2) \epsilon + \beta \max_{\tilde{r} \in \mathcal{R}_\epsilon} \{V^*(\tilde{r}) - V_\epsilon^*(\tilde{r})\}. \quad (\text{EC.16})$$

Since $\max_{r \in \mathcal{R}_\epsilon} \{V^*(r) - V_\epsilon^*(r)\} = \max_{\tilde{r} \in \mathcal{R}_\epsilon} \{V^*(\tilde{r}) - V_\epsilon^*(\tilde{r})\}$, after re-arranging terms and dividing by $(1-\beta)$ on both sides of (EC.16), it follows that $\max_{r \in \mathcal{R}_\epsilon} \{V^*(r) - V_\epsilon^*(r)\} \leq \frac{\gamma_1 + \beta \gamma_2}{1-\beta} \epsilon$. Plugging the expression of γ_1, γ_2 back to the bound, we have

$$\max_{r \in \mathcal{R}_\epsilon} V^*(r) - V_\epsilon^*(r) \leq \left[\frac{C_1}{4(1-\alpha)(1-\beta)} + \frac{\beta C_2}{4(1-\alpha)(1-\alpha\beta)} \right] \epsilon,$$

where $C_1 = \int_{\theta \in \Theta} \max(b + c_+, b + c_-) dG(\theta)$, $C_2 = \int_{\theta \in \Theta} \max(c_+, c_-) dG(\theta)$. Thus, we reach our conclusion. $\square$

EC.3.4. Intuition on Optimal Pricing Policy

We provide some intuitions on the optimality of the cyclic pricing policy when facing heterogeneous consumers. As shown in Figure 4(d), there is a higher revenue gain from consumer B when the price is set high around $1/5^{\text{th}}$ of the time. The high price increases the reference price for the following periods. As prices drop, the major revenue gain is from consumer A , which happens around $4/5^{\text{th}}$ of the time. To conclude, we observe that this cyclic pricing policy provides an interesting dynamic pattern to boost revenue from two consumer types at alternating periods. In contrast, a constant price has to be either away from p_A^* or away from p_B^* during all periods. Therefore, a constant pricing policy hinders the revenue significantly from at least one consumer type and is unlikely to be optimal.

EC.4. Technical Details of Nonparametric Estimation

This section provides technical details of nonparametric estimation to complement the introduction in Section 5.

EC.4.1. Subproblem: Characterization and Algorithm

Focusing on our specific problem formulation, we provide a thorough characterization of the subproblem (16) and its optimal solution in Proposition EC.1.

PROPOSITION EC.1. *For any optimal solution $\mathbf{g}^*$ to the subproblem (16),*

$$\mathbf{g}^*(n) = \begin{cases} d_n & \text{if } \mathbf{x}_n^\top \boldsymbol{\vartheta}^* > 0, \\ \mathcal{L}(\mathbf{x}_n, d_n | \delta_{\boldsymbol{\theta}^*}) & \text{if } \mathbf{x}_n^\top \boldsymbol{\vartheta}^* = 0, \text{ for all } 1 \leq n \leq N, \\ 1 - d_n & \text{if } \mathbf{x}_n^\top \boldsymbol{\vartheta}^* < 0, \end{cases} \quad (\text{EC.17})$$

where $(\boldsymbol{\vartheta}^*, \boldsymbol{\theta}^*)$ is the optimal solution to the following maximization problem,

$$\max_{\boldsymbol{\vartheta}, \boldsymbol{\theta}} \sum_{n=1}^N \frac{(2d_n - 1)}{\mathbf{f}_{(t)}(n)} \left(\mathbf{1}\{\mathbf{x}_n^\top \boldsymbol{\vartheta} > 0\} + \mathbf{1}\{\mathbf{x}_n^\top \boldsymbol{\vartheta} = 0\} \frac{\exp(\mathbf{x}_n^\top \boldsymbol{\theta})}{1 + \exp(\mathbf{x}_n^\top \boldsymbol{\theta})} \right). \quad (\text{EC.18})$$

As we recall from the likelihood definition in Section 5.1, each likelihood $\mathbf{f}^\theta$ vector can be captured by one set of parameters $\boldsymbol{\theta}$. But this is no longer true for the boundary points in $\text{cl}(\{\mathbf{f}^\theta | \boldsymbol{\theta} \in \Theta\})$. Equation (EC.17) of Proposition EC.1 states that the optimal solution $\mathbf{g}^*$ to the subproblem (16) can be characterized by two sets of parameters, $\boldsymbol{\vartheta}^*$ and $\boldsymbol{\theta}^*$. When $\mathbf{x}_n^\top \boldsymbol{\vartheta}^* = 0$ for all n , $\mathbf{g}^*$ corresponds to a standard atomic likelihood vector. When the sign of $\mathbf{x}_n^\top \boldsymbol{\vartheta}^*$ varies over n , $\mathbf{g}^*$ belongs to the boundary of the convex hull of all atomic likelihood vectors. The indicator $\mathbf{1}\{\mathbf{x}_n^\top \boldsymbol{\vartheta}^* > 0\}$ corresponds to a consumer type that makes purchase with probability 1 in the presence of $\mathbf{x}_n$, and $\mathbf{1}\{\mathbf{x}_n^\top \boldsymbol{\vartheta}^* < 0\}$ corresponds to no-purchase with probability 1. Therefore, the additional parameter $\boldsymbol{\vartheta}^*$ captures those “extreme” consumer behaviors that may exist in the market, which is not captured by typical finite mixture of logit models. Equation (EC.18) is obtained by plugging two-parameter structure as

explicitly stated in (EC.17) into the original subproblem (16). The complete proof of Proposition EC.1 is in Section EC.4.2.

We then introduce Algorithm 3 that is inspired by the structure of the optimal solution revealed in Proposition EC.1. Algorithm 3 is an alternating maximization procedure embedded into Algorithm 2 (line 4-8) in Section 5. In the θ step of Algorithm 3, the optimization problem over θ is a continuous optimization problem with the objective being a weighted sum of logistic functions. The logistic function can be viewed as a special sigmoid function and maximizing a sum of sigmoids is NP-hard in general (Udell and Boyd 2013). In the case of maximizing a sum of logistic functions with binary covariates and nonnegative weights, Akçakus and Mišić (2021) prove that such a maximization problem is NP-hard via reduction from the MAX 3SAT problem. Due to the hardness of solving this problem exactly, we tackle this continuous optimization problem with powerful off-the-shelf solvers, similar to Jagabathula et al. (2020). Specifically, we have found that the Powell method or BFGS method in the Python `scipy` package is effective in our numerical analysis. To further circumvent sub-optimal local maxima, we run the solver multiple times with random initialization and pick the θ that achieves the highest function value. In the ϑ step of Algorithm 3, the optimization objective over ϑ resembles the 0-1 loss for a binary classification problem with linear classifier and sample weights. Directly minimizing the 0-1 loss is non-convex and NP-hard, even when all weights are equal (Ben-David et al. 2003). To facilitate the computation, we maximize a convex surrogate loss following the popular SVM approach. The hyperplane computed by maximizing the convex surrogate loss does not directly maximize the 0-1 loss but we found that it performs decently well in improving the objective function values in our numerical experiments. Specifically, we use the standard SVM solver in the Python `sklearn` package in ϑ step.

Algorithm 3 Alternating Maximization Algorithm for Subproblem

- 1: Initialize: $\vartheta_{(0)}$ and $\theta_{(0)}$
 - 2: **for** $l = 1, \dots, L$ **do**
 - 3: θ step: $\theta_{(l)} = \arg \max_{\theta} \sum_{n: \mathbf{x}_n^\top \vartheta_{(l-1)} = 0} \frac{(2d_n - 1)}{\mathbf{f}_{(t)}(n)} \frac{\exp(\mathbf{x}_n^\top \theta)}{1 + \exp(\mathbf{x}_n^\top \theta)}$
 - 4: ϑ step: $\vartheta_{(l)} = \arg \max_{\vartheta} \sum_{n=1}^N \frac{(2d_n - 1)}{\mathbf{f}_{(t)}(n)} \left(\mathbf{1}\{\mathbf{x}_n^\top \vartheta > 0\} + \mathbf{1}\{\mathbf{x}_n^\top \vartheta = 0\} \frac{\exp(\mathbf{x}_n^\top \theta_{(l)})}{1 + \exp(\mathbf{x}_n^\top \theta_{(l)})} \right)$
 - 5: **end for**
-

We thus alternate the θ step and the ϑ step until the two steps no longer significantly improve the value of the subproblem objective function (EC.18). The final output is constructed based on the (ϑ_L, θ_L) of the subproblem procedure according to Proposition EC.1. Algorithm 4 in (Jagabathula et al. 2020, Online Appendix) can be viewed as a heuristic approach for our multi-time-period case.

When applying Algorithm 4 in Jagabathula et al. (2020) to our problem, after obtaining $\theta_{(1)}$, ϑ is intuitively computed by applying some upper and lower cut-off values of choice probabilities to approximate extreme consumer type. It is unclear whether the direct approximation gives near optimal solution to the subproblem. In contrast, the alternating maximization procedure allows the ϑ step to generate a ϑ that makes the maximization objective of the subproblem increase or at least stay the same.

EC.4.2. Proof of Proposition EC.1

Proof of Proposition EC.1. By definition of $\nabla\ell(\mathbf{f}_{(t)})$, the subproblem (16) at iteration t of Algorithm 2 is equivalent to

$$\max_{\mathbf{g} \in \text{cl}(\{\mathbf{f}^\theta | \theta \in \Theta\})} \sum_{n=1}^N \frac{\mathbf{g}(n)}{\mathbf{f}_{(t)}(n)}. \quad (\text{EC.19})$$

When $\mathbf{g}$ is a non-boundary point of $\text{cl}(\{\mathbf{f}^\theta | \theta \in \Theta\})$, i.e., $\mathbf{g} \in \{\mathbf{f}^\theta | \theta \in \Theta\}$, there exists θ such that $\mathbf{g} = \mathbf{f}^\theta$. The non-trivial case is when $\mathbf{g}$ is a boundary point of $\text{cl}(\{\mathbf{f}^\theta | \theta \in \Theta\})$, i.e., $\mathbf{g} \in \text{cl}(\{\mathbf{f}^\theta | \theta \in \Theta\}) \setminus \{\mathbf{f}^\theta | \theta \in \Theta\}$. By definition of $\text{cl}(\{\mathbf{f}^\theta | \theta \in \Theta\})$, there exists a sequence $\{\theta_k\}_{k=1}^\infty$ such that $\lim_{k \rightarrow \infty} \mathbf{f}^{\theta_k} = \mathbf{g}$, where $\mathbf{f}^{\theta_k}(n) = \mathcal{L}(\mathbf{x}_n, d_n | \theta_k)$.

Furthermore, the sequence $\{\theta_k\}_{k=1}^\infty$ is unbounded. Otherwise, if $\{\theta_k\}_{k=1}^\infty$ is bounded, $\{\theta_k\}_{k=1}^\infty$ contains a convergent sub-sequence that converges to some vector $\tilde{\theta}$. It follows that $\mathbf{g} = \mathbf{f}^{\tilde{\theta}} \in \{\mathbf{f}^\theta | \theta \in \Theta\}$, but this contradicts with the fact that $\mathbf{f}$ is a boundary point.

Without loss of generality, we can assume that $\lim_{k \rightarrow \infty} \|\theta_k\| = \infty$, because $\{\theta_k\}_{k=1}^\infty$ always contains a subsequence whose norm goes to infinity. Let S be a set of indexes defined as $S = \{n \in [N] | \mathbf{g}(n) \notin \{0, 1\}\}$. Let V be the linear subspace spanned by $\{\mathbf{x}_n | n \in S\}$. We consider the orthogonal decomposition with respect to V and $V^\perp$, where $V^\perp$ is the orthogonal complement of subspace V . For all k , θ_k can be decomposed to $\theta_k = \alpha_k + \beta_k$, where $\alpha_k \in V, \beta_k \in V^\perp$.

For $n \in S$, we have $\mathbf{g}(n) \notin \{0, 1\}$, thus $\{\mathbf{x}_n^\top \theta_k\}_{k=1}^\infty$ is bounded and $\lim_{k \rightarrow \infty} \mathbf{x}_n^\top \theta_k$ exists. On the other hand, $\mathbf{x}_n^\top \theta_k = \mathbf{x}_n^\top \alpha_k$, so $\{\mathbf{x}_n^\top \alpha_k\}_{k=1}^\infty$ is bounded. Note that $\{\mathbf{x}_n | n \in S\}$ is a set of basis of V and $\alpha_k \in V$, so $\{\alpha_k\}_{k=1}^\infty$ is also bounded. To see why $\{\alpha_k\}_{k=1}^\infty$ is bounded, without loss of generality, assume $\mathbf{x}_1, \dots, \mathbf{x}_r$ are linearly independent and spans space V . Then for each α_k , it can be decomposed to $\alpha_k = \sum_{i=1}^r a_{ki} \mathbf{x}_i$. Therefore $\mathbf{x}_j^\top \alpha_k = \sum_{i=1}^r a_{ki} \mathbf{x}_j^\top \mathbf{x}_i$. Let A be the matrix such that $A[i, j] = \mathbf{x}_i^\top \mathbf{x}_j$. Then we have the linear equation $A\alpha_k = \mathbf{b}$ where $\mathbf{b}_j = \mathbf{x}_j^\top \alpha_k$. Since $\mathbf{b}$ is uniformly bounded and A is invertible, $\alpha_k = A^{-1}\mathbf{b}$ is bounded, which implies that $\{\alpha_k^\top \alpha_k\}_{k=1}^\infty$ is bounded.

Thus, there exists a convergent subsequence of $\{\alpha_k\}_{k=1}^\infty$ and let the limit be α . Therefore $\mathbf{x}_n^\top \alpha = \lim_{k \rightarrow \infty} \mathbf{x}_n^\top \alpha_k = \lim_{k \rightarrow \infty} \mathbf{x}_n^\top \theta_k$. For $\{\beta_k\}_{k=1}^\infty \in V^\perp$, we consider a normalized sequence $\{\beta_k / \|\beta_k\|\}_{k=1}^\infty$. Here, the normalized term equals $\mathbf{0}$ if $\beta_k = \mathbf{0}$. The bounded sequence $\{\beta_k / \|\beta_k\|\}_{k=1}^\infty$ contains a convergent subsequence and let the limit be β . For the limit β , it holds that $\mathbf{x}_n^\top \beta = 0$ for all $n \in S$.

For $n \notin S$, $\mathbf{g}(n)$ is equal to either 0 or 1, depending on the signs of $\mathbf{x}_n^\top \boldsymbol{\theta}_k$ for sufficiently large k . The reason why $\mathbf{g}(n)$ reaches the bound value 0,1 is that $|\mathbf{x}_n^\top \boldsymbol{\theta}_k| \rightarrow \infty$ as $k \rightarrow \infty$. Since $\boldsymbol{\alpha}_k$ is bounded, the sign of $\mathbf{x}_n^\top \boldsymbol{\theta}_k$ is determined by $\mathbf{x}_n^\top \boldsymbol{\beta}_k$. Therefore, the limit $\mathbf{x}^\top \boldsymbol{\beta}$ carries the same sign as $\mathbf{x}_n^\top \boldsymbol{\theta}_k$. Therefore, we have shown that every boundary type can be written as

$$\mathbf{g}^*(n) = \begin{cases} d_n & \text{if } \mathbf{x}_n^\top \boldsymbol{\beta} > 0, \\ \frac{d_n \exp(\mathbf{x}_n^\top \boldsymbol{\alpha}) + (1 - d_n)}{1 + \exp(\mathbf{x}_n^\top \boldsymbol{\alpha})} & \text{if } \mathbf{x}_n^\top \boldsymbol{\beta} = 0, \\ 1 - d_n & \text{if } \mathbf{x}_n^\top \boldsymbol{\beta} < 0. \end{cases}$$

As for non-boundary type, the above expression holds trivially with $\boldsymbol{\beta} = \mathbf{0}$.

More specifically, for the optimal solution $\mathbf{g}^*$ to the subproblem, such $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ exist as well. Suppose $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ does not maximize (EC.18), let $\tilde{\mathbf{g}}$ be constructed based on $(\boldsymbol{\theta}_*, \boldsymbol{\vartheta}_*)$ according to equation (EC.17), then it holds that

$$\begin{aligned} & \sum_{n=1}^N \frac{\tilde{\mathbf{g}}(n)}{\mathbf{f}^{(t)}(n)} \\ &= \max_{\boldsymbol{\theta}, \boldsymbol{\vartheta}} \sum_{n=1}^N \frac{(2d_n - 1)}{\mathbf{f}_{(t)}(n)} \left(\mathbf{1}\{\mathbf{x}_n^\top \boldsymbol{\vartheta} > 0\} + \mathbf{1}\{\mathbf{x}_n^\top \boldsymbol{\vartheta} = 0\} \frac{\exp(\mathbf{x}_n^\top \boldsymbol{\theta})}{1 + \exp(\mathbf{x}_n^\top \boldsymbol{\theta})} \right) \\ &= \sum_{n=1}^N \frac{(2d_n - 1)}{\mathbf{f}_{(t)}(n)} \left[\mathbf{1}\{\mathbf{x}_n^\top \boldsymbol{\vartheta}^* > 0\} + \mathbf{1}\{\mathbf{x}_n^\top \boldsymbol{\vartheta}^* = 0\} \frac{\exp(\mathbf{x}_n^\top \boldsymbol{\theta}^*)}{1 + \exp(\mathbf{x}_n^\top \boldsymbol{\theta}^*)} \right] \\ &> \sum_{n=1}^N \frac{(2d_n - 1)}{\mathbf{f}_{(t)}(n)} \left[\mathbf{1}\{\mathbf{x}_n^\top \boldsymbol{\beta} > 0\} + \mathbf{1}\{\mathbf{x}_n^\top \boldsymbol{\beta} = 0\} \frac{\exp(\mathbf{x}_n^\top \boldsymbol{\alpha})}{1 + \exp(\mathbf{x}_n^\top \boldsymbol{\alpha})} \right] \\ &= \sum_{n=1}^N \frac{(2d_n - 1)}{\mathbf{f}^{(t)}(n)} \mathbf{g}^*(n), \end{aligned}$$

where the inequality holds because $(\boldsymbol{\theta}^*, \boldsymbol{\vartheta}^*)$ is the optimal solution to the maximization problem (EC.18). It is a contradiction to the fact that $\mathbf{g}^*$ achieves the maximum value. $\square$

EC.5. Details and Further Discussion of Numerical Experiments

EC.5.1. Aggregate Demand Model versus Individual Demand Model

So far, we have discussed the aggregate and individual-level demand models in terms of their power in expressing consumer reference effects in Section 3. In addition, the rise of online retailing is a stimulus for individual-level demand models. Compared to traditional retailing at brick-and-mortar stores, online retailers acquire richer information of individual consumers through the online platform. Apart from private identity information, each consumer's shopping activities, including browsing, clicking and purchasing, are tracked and recorded in data. The online retailer possesses price records that each consumer actually encounters and the exact purchase decisions of each consumer. Such information, as provided in the JD.com dataset, enables us to leverage our individual demand model and conduct

effective estimation and optimization. Finally, although we do not allow personalized pricing in this work, the estimated individual demand model provides a foundation for personalization, which can be conducted similarly to Chen et al. (2022).

From the perspective of demand forecasting, the rapidly changing environment in online retailing makes it more difficult to estimate aggregate demand models. With frequently changing prices, only a few data points can be aggregated at each price, which may lead to greater noise when estimating aggregate demand models. In addition, aggregate-level models are extremely restrictive in capturing consumer heterogeneity in their behaviors. For example, if the online retailer wants to do retargeting or provide personalized recommendations, only limited information can be extracted from aggregate demand models. In contrast, each data point is utilized to estimate the parameters at the individual-level demand model. In Section 6.2, we further compare the two models using the JD.com dataset.

EC.5.2. Implementation and Metrics

In Section 6.1, our nonparametric MLE approach is compared with three approaches: finite mixed logit, single logit, and linear model. The most common benchmark for fitting the finite mixed logit model is via the expectation-maximization (EM) algorithm. For the EM algorithm, we follow the method described in Train (2008) and the implementation of the `logisregmixEM` function from a popular R package `mixtools` (Benaglia et al. 2010). Algorithm 2 works well when the number of components increases, ranging from 7 to 18 in our numerical experiments. The single logit model and linear model are fitted via standard functions provided in the Python package `sklearn`.

Out of four estimation approaches under comparison, only the piece-wise linear model is at the aggregate level while the other three are based on individual level logit models. In each replication k , we aggregate the data points over a span of S_0 ($S_0 = 25$) so that the aggregate demand is the sum of demand from S_0 data points and the aggregate price is the average price across S_0 data points. The aggregate linear model is then fitted on the $M = \lfloor N/S_0 \rfloor$ data points. The aggregate version of metrics RMSE and MAE are defined similarly.

$$\text{RMSE} = \sqrt{\frac{1}{K \times N} \sum_{k=1}^K \sum_{m=1}^M (D_m - \hat{D}_m)^2}, \quad \text{MAE} = \frac{1}{K \times N} \sum_{k=1}^K \sum_{m=1}^M |D_m - \hat{D}_m|$$

where D_m is the total expected demand from the aggregated data points and $\hat{D}_m$ represents an estimation via one of these approaches.

EC.5.3. Fitting with Aggregate-level Demand Model

Using transaction data from JD.com, we apply aggregate-level demand models for demand estimation under reference effects. For the linear model (18), $D_t = A - Bp_t + C_+(r_t - p_t)_+ + C_-(r_t - p_t)_-$, where A, B, C_+, C_- are four parameters that need to be estimated from the data. Further specifications are needed for definitions of periods and prices. Firstly, it is natural to segment time by day since

it is a complete time cycle and demand per period is more stationary. Secondly, as reflected in the dataset, the actual transaction prices over the same day may vary. This is partly due to various coupons (limited availability), promotion (limited time), and group buying strategy on the JD.com platform. For linear models, it is reasonable to take an average over the observed prices on each day. The transaction data is aggregated by day and there are 31 days in March 2018. Table EC.2 shows the estimation result for three selected products.

Table EC.2 Estimation of Responsiveness Parameters in the Linear Demand Model

	A	B	C_+	C_-
SKU1	120.327	-0.721	0.782	-0.226
SKU2	165.242	-2.454	0.723	-3.187
SKU3	-17.878	0.701	11.446	-0.219

EC.5.4. Results of the JD.com Dataset

The nonparametric MLE estimation results of three SKUs are shown in Table EC.3, EC.4, and EC.5. All three results demonstrate strong consumer heterogeneity in their response to reference effects, although there exist some dominant market segments with significantly higher proportion (0.424, 0.535, and 0.310 in three cases respectively). We note that the signs of coefficients of the dominant market segments align with the intuition that demand increases as reference price increases and price decrease. The common appearances of 0's in the last element of the ϑ parameter are due to a special fact in the real dataset. For most transactions in the real dataset, the reference price is higher than the price due to a rather large initial reference price, as is commonly observed in practice that the retailers may show a rather high original price in the product page. Consequently, the last element $(r - p)_-$ of the feature vector $\mathbf{x}$ is 0 for most data points. After running a SVM based procedure in the ϑ step of Algorithm 3, the last element of ϑ often becomes 0 because those few units with nonzero $(r - p)_-$ often do not play a role in deciding the position of the separating hyperplane. On the other hand, the last element of θ is not necessarily 0 because in the θ step of Algorithm 3 the optimization procedure aims to maximize the sum of log likelihood functions of all data points.

Table EC.3 Complete Estimation Results of SKU1 'adfedb6893'

Proportion ω	Responsiveness parameters θ					Extreme parameter ϑ			
0.931	-527864.03	236067.989	-527864.03	999999.981		0	0	0	0
0.046	-10925.237	-245805.52	-103169.52	318714.318		0	0	0	0
0.023	-0.228	0.793	0.795	0.137		0	0	0	0

In Table 5 of Section 6, we compare the long-term revenue of all four methods under the demand response simulated by the nonparametric MLE estimation results. On the contrary, we may simulate

Table EC.4 Complete Estimation Results of SKU2 '3c79df1d80'

Proportion ω	Responsiveness parameters θ				Extreme parameter ϑ			
0.656	124716.238	123425.097	-275988.2	522839.046	0	0	0	0
0.281	-0.973	0.207	0.292	0.635	0	0	0	0
0.063	62183.915	-171452.9	-244198.6	120987.604	0	0	0	0

Table EC.5 Complete Estimation Results of SKU3 'b5cb479f7d'

Proportion ω	Responsiveness parameters θ				Extreme parameter ϑ			
0.647	697034.754	406886.594	267553.335	45915.053	0	0	0	0
0.13	-527864.03	236067.989	216953.347	999999.981	1.023	3.263	2.718	0
0.059	115716.362	430522.912	519592.215	995693.066	22.151	1.299	0.656	0
0.052	-78949.07	-3387.263	-5026.518	487742.478	-80.494	-2.244	-1.463	0
0.032	-74162.321	115676.876	98583.433	490049.124	-17.641	3.325	3.349	0
0.025	-11275.278	150931.9	158840.032	639360.55	0	0	0	0
0.022	-527864.03	236067.989	252204.707	999999.981	0	0	0	0
0.018	65769.991	-51847.703	-132755.98	750776.405	-34.938	-1.433	-1.49	0
0.014	-268989.76	242804.373	475836.851	-75457.33	-53.217	-0.393	1.88	0

the demand response simulated by the remaining three methods respectively and compare the long-term revenues. Table EC.6 shows the result when the demand is simulated by the aggregate linear model. We focus on this case because this linear model is most widely used in practice. As expected, policy by the aggregate linear model performs the best since the demand response is simulated by itself. Notably, among the three individual level models, nonparametric MLE still outperforms the other two methods. For both SKU1 and SKU2, the performance of nonparametric MLE are decently close to the optimal policy by the aggregate linear model. We provide some explanation on the different behavior in SKU3 in Table EC.6. For SKU3, the estimated linear model has a negative intercept (the parameter A in Definition (18)) as shown in Table EC.2. Because linear model itself does not naturally prevent negative demand, it is possible that the calculated linear demand is negative under certain pairs of price and reference price. On the contrary, the demands under all three individual level models are always non-negative by definition. Therefore, the pricing policy for individual and aggregate linear models are rather different. In the case of SKU3, we see in Table EC.2 a major increase in revenue gap between aggregate linear model and the individual level models compared to SKU1 and SKU2.

EC.5.5. Myopic Pricing Policy Applied to JD.com Data

Besides the optimal pricing policies, the myopic pricing policy could be of interest to practitioners. Previously in Section EC.2, we have obtained that the performance of myopic pricing policies in simulation. Using the estimation results of the JD.com dataset, we empirically study the performances of myopic pricing policies and find that the mean (median) long-term revenue for three SKUs are 145.117 (145.119), 797.257 (823.848), and 64.433 (66.714) respectively, which are decently close to

Table EC.6 Long-term Revenue in Simulations (with Aggregate Linear Model as Ground Truth)

Policy\Revenue	SKU1		SKU2		SKU3	
	Mean	Median	Mean	Median	Mean	Median
Nonparametric MLE	758082.020	758027.074	368106.529	369167.012	11217.728	10397.791
Mixed Logit via EM	758082.020	758027.074	337919.187	339719.390	498.146	0.000
Single Logit	683486.940	688191.602	336555.397	337411.023	0.000	0.000
Aggregate Linear	758399.425	758316.097	385382.842	385356.246	32827.100	31630.088

Notes: Zero revenue occurs when the generated pricing policy leads to negative demand based on the aggregate linear model, which is then truncated to zero when computing revenue.

the long-term revenue of nonparametric MLE in Table 5. This suggests that myopic pricing policy may serve as a good approximation in practice.

EC.5.6. Examining Price Endogeneity

We examine the issue of price endogeneity in our case study of the JD.com dataset. As suggested in Ferreira et al. (2016, Appendix C), we calculate the sample correlation coefficient between the error $\bar{d}_t - \hat{d}_t$ and the price p_t , and use it as a test of influences due to endogeneity. We calculate the correlation coefficients based on the processed data and estimated results introduced in Section 6.2. Among all tested SKUs, the magnitudes of correlation coefficients range from 0.016 to 0.30 with an average of only 0.16, which suggests that our estimation method on the JD.com dataset might not be troubled by the endogeneity issue in any significant way. On the other hand, if a reasonable instrumental variable for price-demand, potentially from additional data sources, were available, there are multiple methods to further correct endogeneity in the data, such as adopting the control function approach in Petrin and Train (2010).

References for E-Companion

- Akçakuş İ, Mišić V (2021) Exact logit-based product design. *Available at SSRN 3875986*.
- Ben-David S, Eiron N, Long PM (2003) On the difficulty of approximately maximizing agreements. *Journal of Computer and System Sciences* 66(3):496–514.
- Benaglia T, Chauveau D, Hunter DR, Young DS (2010) mixtools: an r package for analyzing mixture models. *Journal of statistical software* 32:1–29.
- Chen X, Owen Z, Pixton C, Simchi-Levi D (2022) A statistical learning approach to personalization in revenue management. *Management Science* 68(3):1923–1937.
- Ferreira KJ, Lee BHA, Simchi-Levi D (2016) Analytics for an online retailer: Demand forecasting and price optimization. *Manufacturing & Service Operations Management* 18(1):69–88.
- Jagabathula S, Subramanian L, Venkataraman A (2020) A conditional gradient approach for nonparametric estimation of mixing distributions. *Management Science* 66(8):3635–3656.

- Petrin A, Train K (2010) A control function approach to endogeneity in consumer choice models. *Journal of Marketing Research* 47(1):3–13.
- Train KE (2008) EM algorithms for nonparametric estimation of mixing distributions. *Journal of Choice Modelling* 1(1):40–69.
- Udell M, Boyd S (2013) Maximizing a sum of sigmoids. *Optimization and Engineering* 1–25.