Optimal Pricing under Diffusion-Choice Models

Hongmin Li
W.P. Carey School of Business
Arizona State University
Tempe, Arizona 85287, USA
hongmin.li@asu.edu
(480) 965-2232
Fax (480) 965-8629

September 2, 2019

Abstract

We develop a solution approach to the centralized pricing problem of a firm managing multiple substitutable products. Demand of these products undergoes a diffusion process and customers choose among the products, with the choice probability of each product given by the logit model. We examine the firm's optimal pricing problem when product demand can be described by such "diffusion-choice" models. In particular, we focus on two models with proven merits, proposed by Weerahandi and Dalal (1992) and Jun and Park (1999) respectively, and study a generalized version of the two models. To our knowledge, ours is the first to study multi-product pricing problem under the integrated diffusion-choice models which are of both theoretical appeal and practical advantage. We establish uniqueness of the optimal solution, propose an efficient solution approach, in addition to characterizing the optimal prices and their time trend. We show that the price trend can be attributed to diffusion dynamics, as well as intertemporal changes in product costs and customer price sensitivity, all of which are integrated into a unified framework in this paper. Our model applies to both simultaneous and sequential product introductions and adapts to stochastic demand.

Keywords: *Pricing, Revenue Management, Diffusion, Multinomial Logit Model, Diffusion-Choice Model, Product Life Cycle*

Optimal Pricing under Diffusion-Choice Models

Hongmin Li

W.P. Carey School of Business, Arizona State University
Tempe, Arizona 85287, USA

hongmin.li@asu.edu

We develop a solution approach to the centralized pricing problem of a firm managing multiple substitutable products. Demand of these products undergoes a diffusion process and customers choose among the products, with the choice probability of each product given by the logit model. We examine the firm's optimal pricing problem when product demand can be described by such "diffusion-choice" models. In particular, we focus on two models with proven merits, proposed by Weerahandi and Dalal (1992) and Jun and Park (1999) respectively, and study a generalized version of the two models. To our knowledge, ours is the first to study multi-product pricing problem under the integrated diffusion-choice models which are of both theoretical appeal and practical advantage. We establish uniqueness of the optimal solution, propose an efficient solution approach, in addition to characterizing the optimal prices and their time trend. We show that the price trend can be attributed to diffusion dynamics, as well as intertemporal changes in product costs and customer price sensitivity, all of which are integrated into a unified framework in this paper. Our model applies to both simultaneous and sequential product introductions and adapts to stochastic demand.

1 Introduction

Product adoption often resembles a diffusion process (Bass, 1969) and the resulting sales exhibit a time trend that can be described by a diffusion equation and characterized by a bell-shaped curve. When the diffusion of a product takes place in parallel with substitute products offered by the same firm, demand interactions among the products create interesting and challenging complications for decision makers. For example, sales of a new book exhibit diffusion characteristic, and often hardcover, paperback, and electronic versions of the same book are sold concurrently. It is also common for manufacturers of technology products to introduce multiple variations of a new product to the market, each variation as a separate model, diffusing into the market concurrently. For example, Apple introduces three different versions of iPhone7 (32GB, 128GB, and 256GB memory), as well as three different versions of iPhone7plus concurrently. Product prices and the resulting price differences among the products play an important role in shaping the diffusion of the products and customers' product choice. In this paper, we study the firm's optimal pricing strategy within this context. Specifically, we consider how the firm should set prices of products over the products' life cycle when demand exhibits diffusion characteristics and customers choose among multiple available product options.

In practice, firms adopt varied pricing strategies. Prices of microprocessors go through periodic downward readjustment as the products enter different phases of their life cycles (Li and Huh, 2012). By the end of a two-year life cycle, the retail prices of smart phones are redued by 34% on average (Entwistle, 2012). Prices of Apple's iPhone slided by 25% between 22 and 28 months of age (Entwistle, 2012) and the retail price of Apple's iPhone 5s declined by 7% one year after launch (Entwistle, 2015). Samsung uses an interesting strategy in which the average retail price peaks one month after release with the introduction price at a 10% discount of the peak price and the ending price at 50% discount (Maling, 2012). Price increase is observed also in other product categories. A study by Melser and Syed (2007) indicates that the introduction prices for new beer varieties are relatively cheap and prices are raised once a following is established. Amazon routinely adjusts prices for books with strong life cycle characteristics such as those in the "Mystery, Thriller and Suspense" category, where both price decreases and price increases are observed in the first few weeks of release (Amazon Website 2017). These anecdotal albeit prevalent phenomena suggest diverse pricing strategies for products through their life cycles.

The difficulty and complexity of pricing decisions in a diffusion often drive firms to myopic strategies. According to PriceBeam, a pricing solution provider, companies sometimes conduct extensive research on setting the proper launch price but fails to re-evaluate the pricing continuously over the product life cycle (PLC) (PriceBeam, 2017): *"This is problematic in two ways: Firstly, you will either be losing out on revenue or leaving money on the table as you fail to take into account the changes in the customer's willingness-to-pay, and secondly, it will most likely deteriorate the returns on other products in your portfolio. Proper life cycle pricing across both the early, mid- and late-stage of the PLC is absolutely crucial to reap the rewards from the R&D investment."* While practitioners are actively searching for a solution that continuously optimizes product prices and accounts for product interactions throughout the diffusion, pricing in multi-product diffusion is under-studied in the academic literature and existing models are complicated and intractable. We present in this paper a viable model and a novel and efficient solution approach for this complex problem.

Researchers have applied diffusion models and choice models separately in many industries. The diffusion model has been successful in modeling product adoption and the resulting aggregate demand pattern, and has been extended to include marketing-mix variables such as advertising and pricing. However, it does not extend easily to admit product and consumer characteristics data for sales prediction. On the contrary, choice models (in particular, logit choice models) have been very successful in linking consumer and product attributes to sales prediction, however, they do not address the diffusion dynamics that are inherent to the demand of a product over its life cycle. Therefore, the idea of integrating consumer choice with diffusion is appealing. As Weerahandi and Dalal (1992) point out, the combination of choice models with diffusion makes it possible to incorporate customer demographics, price, and product attributes and to utilize cross-sectional data on customers' purchasing decision to predict sales penetration.

Despite its theoretical appeal, integrating the choice model and the diffusion model is not straightforward. In this paper, we restrict our attention to integrated "diffusion-choice" models that have successfully fitted empirical data and provided good forecast and understanding of the underlying market. Specifically, we discuss in detail two important models – Weerahandi and Dalal (1992) ("Weerahandi-Dalal" model hereafter), and Jun and Park (1999) ("Jun-Park" model hereafter). The former is a single-product diffusion-choice model which treats each adoption as a choice decision but does not consider interactions among products. The latter addresses diffusion of competing products but omits the imitation (or word-of-mouth) effect which is often considered a key component in product adoption. Both models are shown to work well on real-world data set such as fax communication services (Weerahandi and Dalal, 1992), DRAM (Jun and Park, 1999; Kim et al., 2005) and large-screen TVs (Lee et al., 2006). In this paper, we examine a firm's pricing decision when product demand can be described by such diffusion-choice models. We present a general model structure that encompasses both the Weerahandi-Dalal and the Jun-Park models and solve and characterize the optimal pricing solution.

To our knowledge, ours is the first paper addressing the pricing problem in diffusion-choice models. From a theoretical perspective, we contribute to both the pricing and revenue management literature and the diffusion theory literature. While price extensions of the diffusion model are many, none examines pricing under diffusion-choice models despite their proven merits. Indeed, solving the optimal pricing problem under either the Weerahandi-Dalal or Jun-Park model makes significant technical and practical contributions while this paper effectively solves both. A key advantage of our pricing model is its tractability for multi-product diffusion and, unlike existing pricing models for diffusion, the complexity and computation burden grow very little with the number of products. In the pricing and revenue management literature, multinomial logit (MNL) and related choice models are well studied but none from within the context of product diffusion. The problem becomes significantly harder when examined with diffusion, as customer choices in the current period affect future sales. As a result of this dependency, the state-of-art solution approach does not apply. We present an analysis that circumvents this complexity and show that the optimal solution is unique and efficiently obtainable. Moreover, we characterize properties of the optimal solution, yielding insights into this complex pricing problem.

The rest of the paper is organized as follows. We review the related literature in Section 2 and describe the diffusion-choice model in Section 3. We present the detailed analysis in Section 4 and conduct numerical studies in Section 5. In Section 6, we present extensions that allow sequential product introduction, stochastic demand and product-specific imitation effect, respectively. We conclude in Section 7.

2 Prior Research

The literature on product diffusion is substantial, following the seminal work of Bass (1969) who identifies two key forces driving product adoption - innovative and imitative behaviors of the consumers. Innovators are early adopters whose decision is not influenced by previous adopters whereas imitators learn from others and feel increasing pressure for adoption as the number of adoptions increases. The imitation effect is also referred to as the word-of-mouth effect. The magnitude of this effect varies by industry, which is shown in the studies by both Mansfield (1961) and Bass (1969). Norton and Bass (1987) extend the diffusion model in Bass (1969) to successive generations of products. A series of Bass model extensions incorporate the price variable by modifying the total market potential or the adoption rate with a price term; see Robinson and Lakhani (1975), Bass (1980), Dolan and Jeuland (1981), Kalish (1983), Kamakura and Balasubramanian (1988), Jain and Rao (1990), and Krishnan et al. (1999). Kalish (1983) and Horsky (1990) show that the optimal price trend depends on the word-of-mouth effect and that penetration pricing (skim pricing) is optimal with strong (weak) word-of-mouth effect. Krishnan et al. (1999) characterize the optimal price trajectory under a variation of the generalized Bass model (Bass et al., 1994) and show that the transition point of price trend depends on the discount factor and the price sensitivity parameter instead of the innovation and imitation parameters; with zero discounting, the optimal price shows a monotonically increasing trend nearly for the entire planning horizon. In this paper, we establish a unimodal price path for time-invariant product quality, price sensitivity and cost, generalizing the result in Kalish (1983) and Horsky (1990) to multi-product diffusion. In addition, our model allows more general price trend to be justified by the effect of time-variant product attributes, price sensitivity, and cost.

More recently, Li and Huh (2012) examine pricing of a product going through diffusion assuming a linear demand-price relationship superimposed on the diffusion curve in which the effect of price does not carry through to future demand. They show that the price trend is driven by the interaction between the diffusion curve and the intertemporal change in customer price sensitivity. Our model is more flexible, capturing not only intertemporal changes in price sensitivity, but also its effect on diffusion and future sales. Shen et al. (2014) consider a new product diffusion model with supply constraints and show that the optimal price trajectory is complicated with inventory and capacity considerations: it can be bimodal under the lost sales assumption and may have up to three modes under a partial backlogging assumption. Most pricing models for product diffusion, including the aforementioned, are for a single product and those that do address more than one product are typically in a stylized setting with only two products (e.g., Bayus 1992); the primary reason is that these models become cumbersome and intractable for price optimization as the number of products increases beyond two. In contrast, we consider pricing in a multi-product diffusion with adoption decisions given by the MNL choice model and we provide a scalable algorithm to solve the problem for n products over T periods, which was unattainable with existing models. Consumer utility is a function of both product quality and price and each consumer chooses the product that maximizes his/her own utility. As a result, we are able to capture the effect not only of diffusion dynamics, but also of product attributes and the demand interactions among products. For example, we show that products with higher net quality have larger swings in sales, i.e., larger gap between peak and non-peak sales, a characteristic that prior models were not able to identify.

Several researchers integrate discrete choice models with product diffusion. Oren and Schwartz (1988) consider diffusion of a new product when each customer makes adoption decision by solving a utility maximization between the new and current product. They assume that a customer's utility depends on the uncertainty of product performance which decreases over time and show that this uncertainty and the heterogeneity in customer's risk aversion lead to a logistic growth curve. Weerahandi and Dalal (1992) propose a single-product diffusion-choice model and demonstrate with an application of the fax communication services that the model is useful for studying segmented market and the effect of marketing strategies and product features on sales. Jun and Park (1999) predict sales in the DRAM market with a multi-product diffusion-choice model. It has been later adopted in a series of applications for competing product or service offerings including DRAM (Kim et al., 2005), telephony and broadband service subscriptions (Chen and Watanabe, 2006), large-screen TVs (Lee et al., 2006), and home networking technologies (Lee et al., 2008). We study a firm's pricing decisions under such diffusion-choice models and explore how the diffusion dynamics play into the pricing decisions and vice versa.

One stream of related work extends the MNL model to include network effect in which the customers' utility for consuming a product depends on the total consumption level of the product. Miyao and Shapiro (1981) examine situations in which the network effect can be positive, negative, or zero and establish existence of a steady state equilibrium in which the choice probability of each product equals the actual consumption proportion of each product. Wang and Wang (2016) consider assortment planning under this endogenized MNL model. Du et al. (2016) study the pricing problem when this network effect is a linear function of the total consumption and show an interesting feature of the optimal prices in which one product has a low price whereas all other products are priced higher at a common price. While the steady-state description of the network effect encompasses the idea that choices are affected by total consumption, it does not fully capture the dynamics of product diffusion, which usually do not reach a steady state but exhibit a time-varying demand pattern. In contrast, we consider diffusion-choice models that explicitly model the diffusion dynamics and we characterize the optimal price behaviors throughout the diffusion cycle.

Our paper also builds upon advances in price optimization under the family of logit choice models. Hanson and Martin (1996) pioneer this research and show that the profit function under the MNL model is non-concave in prices. Dong et al. (2009) and Song and Xue (2007) prove concavity with respect to the choice probability vector under the MNL model. Li and Huh (2011) extend the concavity result to the nested Logit model for symmetric within-nest price sensitivities and Gallego and Wang (2014) identify a more general condition for a unique solution. Davis et al. (2014) and Rayfield et al. (2015) propose methods for optimizing prices under the nested logit model with price constraints. Huh and Li (2015) and Li et al. (2015) examine multi-level nested logit models. More recently, Wang and Sahin (2017) consider the impact of switching cost on the assortment and pricing decisions under MNL. Li and Webster (2017) present solution approaches for pricing under the paired combinatorial logit (PCL) and the cross-nested logit models. Zhang et al. (2018) study pricing under the broader class of Generalized Extreme Value (GEV) models. In this paper, we extend this literature and demonstrate how diffusion dynamics influence the profit function and the optimal prices and sales of products that form a MNL choice set.

3 Model

Figure 1 illustrates consumer choices in a multi-product diffusion, generalizing the single-product adoption in Weerahandi and Dalal (1992, Figure 1). We assume that the market starts with a certain potential pool of customers. The occurrence of a sale for a particular product i results from two sequential events: first, a purchase occasion arises for a potential customer; second, the customer chooses product i among the set of available options (including the no-purchase option). As in Weerahandi and Dalal (1992), we assume that each customer purchases a single unit. The rate at which purchase occasions arise is driven by two forces: (i) the innovation effect which may depend on the firm's efforts in advertising and promotion but not on the current adoption level, and (ii) the imitation effect which depends on how many customers have already made a purchase. Once a customer decides to make a purchase, he/she drops out of the potential customer population and enters the adopter population which influences future sales through the imitation effect; if a customer decides not to purchase anything, he/she stays in the potential customer population and may become a future adopter.

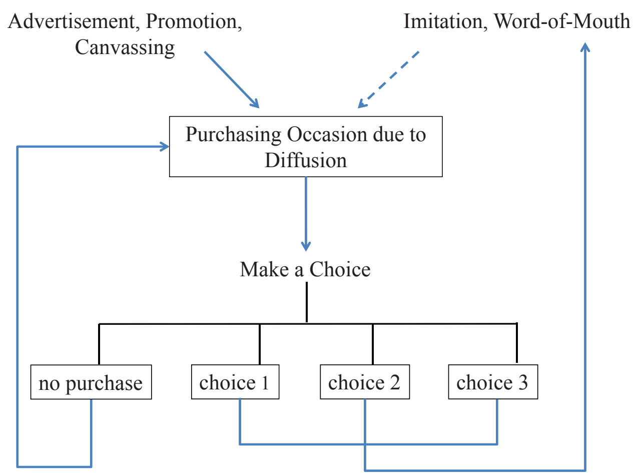


Figure 1: Choice Decision in Product Diffusion.

The conceptual model and assumptions are the same as the Weerahandi-Dalal model, with the binary logit choices extended to multinomial logit choices. Likewise, the model and assumptions are also the same as the Jun-Park model with the addition of word-of-mouth (i.e., the dashed arrow connection in Figure 1). Both models have been tested and applied empirically in the marketing and economics literature with proven validity and merits, but neither has been studied in a normative prescriptive decision context. In this paper, we adopt their conceptual model and assumptions to study the pricing decision in a multi-product diffusion. Given the empirical practicality and scalability of the model, it offers greater advantage than other multi-product diffusion models that are either too stylized for practical decision support (such as Bayus (1992)) or too complex to be tractable and scalable (such as Norton and Bass (1987)).

Consider a firm selling n products concurrently and maximizing profit within a given planning horizon of T time periods. Define M as the total market potential. Let Z_{it} be the sales of product i in period t and Y_t be the cumulative sales (including all products) by the end of period t. Thus

$$Y_t = \sum_{s=1}^{t} \sum_{i=1}^{n} Z_{is} \; . \tag{1}$$

Let the utility of acquiring product i in period t at price p_{it} be given by $a_{it} - b_t p_{it} + \epsilon_{it}$, where a_{it} stands for the quality of the product at period t, b_t is the time-dependent price sensitivity, and ϵ_{it} is a random noise term that is of Gumbel distribution. Here, quality is a measure of attractiveness of the product based on its non-price attributes and features. In addition, we normalize the utility of no purchase in each period to zero.

Following the Weerahandi-Dalal model, we consider both the innovation and imitation dynamics in product adoption. In particular, the number of customers facing a purchase decision in period t depends on the size of the remaining market potential, $(M - Y_{t-1})$, and on a fractional rate $(\alpha + \beta Y_{t-1})$ which we interpret as the fraction of customers in the remaining market potential who will face a purchase occasion. The linear dependency on the size of adopter population is a basic assumption in Bass (1969) and nearly all its extensions. As in Bass (1969), we assume that the time unit is chosen such that $(\alpha + \beta Y_{t-1}) \in [0, 1]$ for $Y_{t-1} \in [0, M]$, $t = 1, 2, \ldots, T$, with $\alpha > 0$ signifying the intensity of the innovation effect and $\beta > 0$ the intensity of the imitation effect (α denotes the portion of customers whose interest is not affected by the current number of adopters, and βY_{t-1} denotes the portion of customers whose interest is influenced by those who have already made a purchase, i.e., the two terms represent the innovators and the imitators respectively). Therefore, the demand for product i in period t is given by

$$Z_{it} = (M - Y_{t-1})(\alpha + \beta Y_{t-1}) q_{it} \; , \quad i = 1, \ldots, n \tag{2}$$

$$\text{where} \qquad q_{it} = \frac{\exp(a_{it} - b_t p_{it})}{1 + \sum_{j=1}^{n} \exp(a_{jt} - b_t p_{jt})} \tag{3}$$

is the purchase probability given by the MNL choice model. We denote the no-purchase probability in period t with

$$q_{0t} \quad = \quad \frac{1}{1 + \sum_{j=1}^{n} \exp(a_{jt} - b_t p_{jt})} \; .$$

Following Weerahandi and Dalal (1992), we let α and β be constants while incorporating the effect of product quality and price into the logit choice probability given in equation (3). In this paper, we consider the expected demand and optimize the firm's profit based on the expected demand given in (2) and we assume ample supply, both of which are simplifications commonly adopted in the diffusion pricing literature (e.g., Kalish 1983; Horsky 1990; Bayus 1992). We examine a dynamic pricing problem based on stochastic adoption in an extension.

The firm's price optimization problem is given by

$$\max_{\substack{p_{it},\\ i=1,\ldots,n,\\ t=1,\ldots,T}} \sum_{t=1}^{T}\sum_{i=1}^{n}(p_{it}-c_{it})Z_{it}$$

where c_{it} is the cost of product i in period t. This problem formulation generalizes both the Jun-Park model and the Weerahandi-Dalal model. Specifically, the special case of $\alpha = 1, \beta = 0$ leads to the Jun-Park model and the special case of $n = 1$ leads to the Weerahandi-Dalal model.

From equations (1) and (2), we can derive

$$M - Y_t = (M - Y_{t-1})\Big[1 - (\alpha + \beta Y_{t-1})\sum_{i=1}^{n} q_{it}\Big] , \tag{4}$$

$$\alpha + \beta Y_t = (\alpha + \beta Y_{t-1})\Big[1 + \beta(M - Y_{t-1})\sum_{i=1}^{n} q_{it}\Big] . \tag{5}$$

Define $H_t := M - Y_{t-1}$ and $F_t := \alpha + \beta Y_{t-1}$ for $t = 1, \ldots, T$. Then

$$H_t = H_{t-1}\Big(1 - F_{t-1}\sum_{i=1}^{n} q_{i,t-1}\Big) , \tag{6}$$

$$F_t = F_{t-1}\Big(1 + \beta H_{t-1}\sum_{i=1}^{n} q_{i,t-1}\Big) , \tag{7}$$

and sales of product i in period t can be rewritten as

$$Z_{it} = F_t H_t q_{it} . \tag{8}$$

We observe in equations (3)-(8) that more attractive products obtain higher q_{it} values and contribute more to the word-of-mouth effect; for the same reason, they also benefit more from the word-of-mouth effect than the less attractive products.

Note that $H_t = M - Y_{t-1}$ represents the remaining market potential, and $F_t = \alpha + \beta Y_{t-1}$ represents the fraction of customers in the remaining market potential that will face a purchase occasion; so we refer to F_t as the *diffusion intensity*. We can then interpret $F_t H_t$ as the number of customers facing a purchase decision at time t. The recursions in equations (6) and (7) indicate how the market potential and diffusion intensity evolve over time. The market potential monotonically decreases each period, which is reflected in the multiplicative term $(1 - F_{t-1} \sum_{i=1}^{n} q_{it-1})$, where $F_{t-1} \sum_{i=1}^{n} q_{it-1}$ signifies the proportion of customers who have made a purchase in period $t-1$. Since $(1 - F_{t-1} \sum_{i=1}^{n} q_{it-1}) \in [0, 1]$, we refer to it as the *shrinkage factor* of the market potential from period $t-1$ to t. Similarly, the diffusion intensity F_t monotonically increases each period due to the imitation effect, as reflected by the term $(1 + \beta H_{t-1} \sum_{i=1}^{n} q_{it-1})$; this term is greater than 1 and we refer to it as the *amplification factor* of the diffusion intensity.

The original single-product Bass diffusion model is fully parameterized with α and β, and sales reach peak volume when the multiplicative product of the remaining market potential H_t and the diffusion intensity F_t reaches its maximum. In the diffusion-choice model, however, sales of each product are also affected by quality a_{it} and prices p_{it} of *all* products through the MNL model, and such dependencies are then carried through to future periods through H_t and F_t values. These dynamics may shift the sales of each product differently. Therefore, the price of each product at any given time has complex rippling effect on the diffusion of this and other products.

4 Optimal Pricing Solution

Let π_{it} be the profit of product i in period t and define the profit-to-go, J_t, as

$$J_t = \max_{\substack{p_{is},\\ i=1,\cdots,n;\\ s=t,\ldots,T}} \sum_{s=t}^{T} \sum_{i=1}^{n} \pi_{is} \ . \tag{9}$$

From equation (8),

$$\begin{aligned} \pi_{it} &= (p_{it} - c_{it}) Z_{it} = F_t H_t \big(p_{it} - c_{it}\big) q_{it} \quad \text{and} \\ J_t &= \max_{\substack{p_{is},\\ i=1,\cdots,n;\\ s=t,\ldots,T}} \sum_{s=t}^{T} \sum_{i=1}^{n} F_s H_s \big(p_{is} - c_{is}\big) q_{is} \ . \end{aligned}$$

Rewrite the price p_{it} as a function of the purchase probability vector $\mathbf{q}_t = (q_{1t}, q_{2t}, ..., q_{nt})$

$$p_{it}(\mathbf{q}_t) = \frac{1}{b_t} \left(a_{it} - \log q_{it} + \log q_{0t}\right) \tag{10}$$

where $q_{0t} = 1 - \sum_{i=1}^{n} q_{it}$. Subsequently,

$$J_t = \max_{\substack{q_{it} \in [0,1],\\ i=1,\cdots,n}} \left[F_t H_t \sum_{i=1}^{n} \big(p_{it}(\mathbf{q}_t) - c_{it}\big) q_{it} + J_{t+1} \right] \ , \tag{11}$$

where $J_{T+1} = 0$. The first term on the right side of equation (11) is the current-period profit and the second term is the profit-to-go from the next period onward. We use the purchase probability vector $\mathbf{q}_t$ as the decision variable for the ease of both notation and derivation.

Define $G_{T+1} := 0$ and

$$G_t(H_t, F_t, \mathbf{q}_t) := \sum_{i=1}^{n} \big(p_{it}(\mathbf{q}_t) - c_{it}\big) q_{it} + \left(1 - F_t \sum_{i=1}^{n} q_{it}\right) \left(1 + \beta H_t \sum_{i=1}^{n} q_{it}\right) G^*_{t+1}(H_{t+1}, F_{t+1}) \quad \text{for } t \le T \tag{12}$$

where

$$G^*_t(H_t, F_t) = \max_{\mathbf{q}_t} G_t(H_t, F_t, \mathbf{q}_t) \ . \tag{13}$$

Note that both H_{t+1} and F_{t+1} are functions of $\mathbf{q}_t$ (see equations (6) and (7)) and thus depend on the decision variables in period t. The product $(1 - F_t \sum_{i=1}^n q_{it})\,(1 + \beta H_t \sum_{i=1}^n q_{it})$ is equal to $\dfrac{F_{t+1}H_{t+1}}{F_t H_t}$, namely, the ratio of the number of customers facing a purchase occasion in period $t+1$ to that in period t. Therefore, from (11) and the definition of $\hat{G}^*_t$, the value-to-go J_t as defined in equation (11) is equivalent to

$$J_t(H_t, F_t) = F_t H_t G^*_t(H_t, F_t) \ . \tag{14}$$

Hence, to solve the optimal prices, it suffices to solve (13).

Before we proceed, it is useful to interpret G_t. Recall that $F_t H_t$ represents the number of customers facing a purchase decision at time t. From (14), G_t measures the expected profit "amortized" over the current number of customers facing purchase decisions, which includes potential profit from this customer and those that he/she will influence through word-of-mouth. The first term in equation (12) signifies the expected period-t profit from a customer who faces a purchase decision in this period; the second term is the expected future profit amortized over the the number of customers who face a purchase decision now. The price decision in period t, therefore, affects the immediate profit in this period and affects future profit through its impact on the remaining market potential H_{t+1} and on the diffusion intensity F_{t+1}. Since the relationship is recursive, decision on q_t affects not only G^*_{t+1}, but also G^*_{t+2} through F_{t+2} and H_{t+2}, G^*_{t+3} through F_{t+3} and H_{t+3}, and so on. This leads to a complex dependency that is polynomial and does not appear to provide a clear path for characterization. In this paper, we present a novel solution approach that circumvents this complexity. Using this approach, we are able to show not only that the profit optimization has a unique price solution, but also that the solution can be efficiently obtained and the optimal price trend can be analytically characterized. In the following, we present in detail the necessary steps leading to this characterization.

4.1 Reduction of the Choice Probability Vector $\mathbf{q}_t$ to a Single Variable θ_t

Given the relationships in (12), it should be clear that the choice of optimal prices in period t depends on previous price decisions, $\{p_s\}_{s=1,\cdots,t-1}$, only through F_t and H_t. Further, the definitions of F_t and H_t imply

$$F_t \;=\; \alpha + \beta(M - H_t)\ ,$$

hence we view F_t as a linear function of H_t and rewrite (12) and (13) as

$$\begin{aligned} G_t(H_t, \mathbf{q}_t) \;&:=\; \sum_{i=1}^{n} \big(p_{it}(\mathbf{q}_t) - c_{it}\big) q_{it} \\ &\quad + \left(1 - F_t(H_t) \sum_{i=1}^{n} q_{it}\right) \left(1 + \beta H_t \sum_{i=1}^{n} q_{it}\right) G^*_{t+1}(H_{t+1}(H_t, \mathbf{q}_t)) \end{aligned}$$

$$\begin{aligned} \text{where } p_{it}(\mathbf{q}_t) \;&=\; \frac{1}{b_t}\left(a_{it} - \log \frac{q_{it}}{1 - \sum_{j=1}^{n} q_{jt}}\right)\ , \\ F_t(H_t) \;&=\; \alpha + \beta(M - H_t)\ , \\ H_{t+1}(H_t, \mathbf{q}_t) \;&=\; H_t\Big(1 - F_t(H_t) \sum_{i=1}^{n} q_{it}\Big) \ \text{ and} \\ G^*_t(H_t) \;&=\; \max_{q_{it}\in[0,1], i=1,\ldots,n} G_t(H_t, \mathbf{q}_t)\ . \end{aligned}$$

For notation brevity, we suppress the function arguments of $F_t(H_t)$, $H_{t+1}(H_t, q_t)$, $G_t(H_t, q_t)$ and $G^*_t(H_t)$ where there is no ambiguity, and write them as F_t, H_{t+1}, G_t and G^*_t for the remainder of the paper.

Take the first-order partial derivative of G_t with respect to q_{it} and set it to zero to obtain a necessary condition for optimality:

$$p_{it}(\boldsymbol{q}_t) - c_{it} - \frac{1}{b_t} = \sum_{i'=1}^{n} \frac{q_{i't}}{b_t(1 - \sum_{j=1}^{n} q_{jt})} + (F_{t+1} - \beta H_{t+1})G^*_{t+1} + F_{t+1}H_{t+1}\frac{dG^*_{t+1}}{dH_{t+1}}\ . \tag{15}$$

where we use the relationships in (6) and (7) to simplify the expression (see the online appendix for the proof of (15)).

Note that the right hand side of the above is independent of the index i, which implies that, at optimality, there exists θ_t such that $p_{it} - c_{it} - \frac{1}{b_t} = \theta_t$ for all i. That is, the optimal markup is symmetric across products, which generalizes the "equal-adjusted-markup" property in Gallego and Wang (2014) to the multi-period diffusion setting. This observation can simplify the problem. In particular, we express the price and choice probability as functions of θ_t

$$
\begin{aligned}
p_{it}(\theta_t) &= \theta_t + c_{it} + \frac{1}{b_t} \ , && (16)\\
q_{it}(\theta_t) &= \frac{\exp(a_{it} - 1 - b_t\theta_t - b_t c_{it})}{1 + \sum_{j=1}^{n} \exp(a_{jt} - 1 - b_t\theta_t - b_t c_{jt})} \ , && (17)
\end{aligned}
$$

and consider a revised optimization problem over $\theta_t, t = 1, \ldots, T$. We rewrite $G_t(H_t, \mathbf{q}_t)$ as

$$
\begin{aligned}
\tilde{G}_t(H_t, \theta_t) &= \sum_{i=1}^{n} \big(p_{it}(\theta_t) - c_{it}\big) q_{it}(\theta_t) \\
&\quad + \Big(1 - F_t(H_t) \sum_{i=1}^{n} q_{it}(\theta_t)\Big)\Big(1 + \beta H_t \sum_{i=1}^{n} q_{it}(\theta_t)\Big) \tilde{G}^*_{t+1}(H_{t+1}) && (18)\\
\text{where} \quad H_{t+1} &= H_t\Big(1 - F_t(H_t) \sum_{i=1}^{n} q_{it}(\theta_t)\Big) \ , && (19)\\
\tilde{G}^*_t(H_t) &= \max_{\theta_t} \tilde{G}_t(H_t, \theta_t) \ . && (20)
\end{aligned}
$$

Hence, we have reduced an n-dimension optimization in each period t to a 1-dimension problem. Nonetheless, the longitudinal complexity is still prohibitive. In the next step, we establish a connection between the optimal θ_t values of two adjacent time periods, which forms a crucial building block of our solution approach.

4.2 Relationship between θ_t and θ_{t-1} at Optimality

With the transformation of the choice probability vector to a single variable θ_t, the first-order necessary condition of optimality becomes

$$
\theta_t = \sum_{i=1}^{n} \frac{q_{it}(\theta_t)/q_{0t}(\theta_t)}{b_t} + (F_{t+1} - \beta H_{t+1})\tilde{G}^*_{t+1} + F_{t+1} H_{t+1} \frac{d\tilde{G}^*_{t+1}}{dH_{t+1}} \qquad (21)
$$

where $q_{0t}(\theta_t) = 1 - \sum_{i=1}^{n} q_{it}(\theta_t)$. Note $q_{it}/q_{0t} = e^{a_{it} - 1 - b_t c_{it} - b_t\theta_t}$ and define

$$
\begin{aligned}
r_t(\theta_t) &:= \theta_t - \sum_{i=1}^{n} \frac{e^{a_{it} - 1 - b_t c_{it} - b_t\theta_t}}{b_t} \ , \\
\Lambda_t &:= F_t H_t \frac{d\tilde{G}^*_t}{dH_t} \ .
\end{aligned}
$$

Then the first-order condition is equivalent to

$$
r_t(\theta_t) = (F_{t+1} - \beta H_{t+1})\tilde{G}^*_{t+1} + \Lambda_{t+1} \ , \quad t \le T \ . \qquad (22)
$$

This transformed first-order condition for the period t decision of θ_t reveals a clear balance between the marginal impact on current (i.e., left side of (22)) and future (i.e., right side of (22)) periods.

In a myopic single-period decision setting, only $r_t(\theta_t)$ is relevant and the optimal decision solves $r_t(\theta_t) = 0$. With future periods in consideration, it is necessary to quantify the values of $(F_{t+1} - \beta H_{t+1})\tilde{G}^*_{t+1} + \Lambda_{t+1}$. A direct quantification is impractical due to the high computation load, the difficulty of computing the derivative $\frac{d\tilde{G}^*_{t+1}}{dH_{t+1}}$, and the fact that it is unclear whether the optimal solution of θ_t is even uniquely identifiable. Instead, by connecting the optimal values of θ_t with that of its adjacent period, we can establish uniqueness of the optimal solution and also dramatically reduce the complexity of the problem.

To accomplish this, we make use of two recursive relationships. Substituting (16) into (18) yields

$$\tilde{G}^*_t(H_t) = \theta_t \sum_{i=1}^{n} q_{it} + \sum_{i=1}^{n} \frac{q_{it}}{b_t} + \frac{F_{t+1}}{F_t}\frac{H_{t+1}}{H_t}\tilde{G}^*_{t+1}(H_{t+1}) \ . \tag{23}$$

In addition, we can show that

$$\Lambda_t = \left(\frac{F_{t+1}}{F_t} + \frac{H_{t+1}}{H_t}\right)(F_{t+1} - F_t)\tilde{G}^*_{t+1} + \left(\frac{F_{t+1}}{F_t} + \frac{H_{t+1}}{H_t} - 1\right)\Lambda_{t+1} \ . \tag{24}$$

(Derivation of (24) appears in the online appendix.)

From the recursive relationships in (23) and (24), as well as the first order condition in (22), we establish a connection between the optimal solutions of two adjacent time periods.

Lemma 1. *Let θ^*_t and θ^*_{t-1} be the optimal solution at time t and $t-1$ respectively for $t = 2, \ldots, T$. Then*

$$r_{t-1}(\theta^*_{t-1}) - r_t(\theta^*_t) = (F_t - \beta H_t) \sum_{i=1}^{n} \frac{e^{a_{it}-1-b_t c_{it}-b_t\theta^*_t}}{b_t} \ .$$

The relationship in Lemma 1 is much simpler and easier to quantify than the first-order condition (22). It is not only crucial for establishing uniqueness of the optimal solution and efficiently solving for it, but also useful for characterizing the optimal price trend, as we show in the next two subsections.

4.3 Uniqueness and Solution Algorithm

It is clear from the problem formulation in (18)-(20) that the optimal θ_t is a function of the remaining market potential at period t, H_t. Let $\theta^*_t(H_t)$ solve the optimization in (18)-(20) for any given H_t. Rewrite the relationship in Lemma 1 as

$$r_{t-1}(\theta^*_{t-1}) = r_t(\theta^*_t) + (F_t - \beta H_t) \sum_{i=1}^{n} \frac{e^{a_{it}-1-b_t c_{it}-b_t\theta^*_t}}{b_t} \ . \tag{25}$$

We express the right side of (25) as a function of H_t and define

$$g_t(H_t) \quad := \quad r_t\big(\theta_t^*(H_t)\big) + \big(F_t(H_t) - \beta H_t\big) \sum_{i=1}^{n} \frac{e^{a_{it}-1-b_t c_{it}-b_t\theta_t^*(H_t)}}{b_t} \ .$$

From (19), $H_t = H_{t-1}\Big(1 - F_{t-1}(H_{t-1}) \sum_{i=1}^{n} q_{i,t-1}(\theta_{t-1})\Big)$ is a function of H_{t-1} and θ_{t-1}; we further define $\hat{g}_t(H_{t-1}, \theta_{t-1}) := g_t\big(H_t(H_{t-1}, \theta_{t-1})\big)$. Therefore, according to (25), the optimal solution $\theta_{t-1}^*(H_{t-1})$ must satisfy

$$r_{t-1}(\theta_{t-1}) = \hat{g}_t(H_{t-1}, \theta_{t-1}) \ , \quad t = 2, \ldots, T \ . \tag{26}$$

Given H_{t-1}, this is a single-variable equation for θ_{t-1}. From this, we establish uniqueness of the optimal solution and construct an efficient method to solve for it, as summarized in the following proposition.

Proposition 1. *(i) For any given H_t, there exists a unique solution $\theta_t^*(H_t)$ which maximizes $\tilde{G}_t(H_t, \theta_t)$. (ii) $\frac{d\theta_t^*(H_t)}{dH_t} \leq 0$. (iii) Suppose $\theta_{t+1}^*(H_{t+1})$ is the unique optimal solution in period $t+1$ for any given H_{t+1}. Then for any given H_t, $\theta_t^*(H_t)$ is the unique root to*

$$r_t(\theta_t) = \hat{g}_{t+1}(H_t, \theta_t) \tag{27}$$

where $\dfrac{dr_t(\theta_t)}{d\theta_t} > 0$ *and* $\dfrac{\partial \hat{g}_{t+1}(H_t, \theta_t)}{\partial \theta_t} < 0.$

The proof relies on inductively showing (i) and (ii) jointly, using the relationship in equation (26). Furthermore, equation (27) can be solved efficiently with a bisection search since the left side monotonically increases and the right side monotonically decreases in θ_t, leading to the following algorithm for computing $\mathbf{q}_t^*$ for all $t = 1, \ldots, T$.

Algorithm 1 (Solution Procedure)**.**

(i) First, solve the period T problem. Since $\tilde{G}_{T+1}^* = 0$, solve $r_T(\theta_T^*) = 0$ to obtain θ_T^* and then let $p_{iT}^* = c_{iT} + \frac{1}{b_t} + \theta_T^*$.

(ii) Next, for $t = T-1, T-2, \ldots, 1$, discretize the plausible range for H_t, e.g., [0,1], to N_H grid points. For each H_t value, solve the single-variable equation (27) via bisection search to obtain $\theta_t^*(H_t)$. Specifically, given H_t, the left hand side of equation (27), i.e., $r_t(\theta_t)$ is computed according to

$$r_t(\theta_t) = \theta_t - \sum_{i=1}^{n} \frac{e^{a_{it}-1-b_t c_{it}-b_t\theta_t}}{b_t} \ ,$$

and the right hand side of equation (27), i.e., $\hat{g}_{t+1}(H_t, \theta_t)$, is computed according to the relationships:

$$\begin{aligned}
\hat{g}_{t+1}(H_t, \theta_t) \quad &= \quad g_{t+1}\big(H_{t+1}(H_t, \theta_t)\big) \ , \\
H_{t+1} \quad &= \quad H_t\Big(1 - F_t(H_t) \sum_{i=1}^{n} q_{i,t}(\theta_t)\Big) \ , \\
g_{t+1}(H_{t+1}) \quad &= \quad r_{t+1}\big(\theta_{t+1}^*(H_{t+1})\big) + \big(F_{t+1}(H_{t+1}) - \beta H_{t+1}\big) \sum_{i=1}^{n} \frac{e^{a_{it+1}-1-b_{it+1}c_{it+1}-b_{it+1}\theta_{t+1}^*(H_{t+1})}}{b_{it+1}} \ .
\end{aligned}$$

(iii) Finally, identify the path of optimal prices and sales from the initial value $H_1 = M - Y_0$ for $t = 1, 2, \ldots, T-1$ according to

$$\begin{aligned}
\theta_t^* &= \theta_t^*(H_t), \\
p_{it}^* &= \theta_t^* + c_{it} + \frac{1}{b_t}\ , \\
q_{it}^* &= \frac{\exp(a_{it} - 1 - b_t\theta_t^* - b_t c_{it})}{1 + \sum_{j=1}^{n}\exp(a_{jt} - 1 - b_t\theta_t^* - b_t c_{jt})}\ , \\
Z_{it}^* &= [\alpha + \beta(M - H_t)]\, H_t q_{it}^*\ , \\
H_{t+1} &= H_t\left\{1 - [\alpha + \beta(M - H_t)]\sum_{i=1}^{n} q_{it}^*\right\},
\end{aligned}$$

where we obtain $\theta_t^*(H_t)$ by interpolating θ_t^* values on the grid computed in step (ii).

In the Weerahandi-Dalal model, the above procedure applies with $n = 1$. In the Jun-Park model, it turns out that with $\alpha = 1$ and $\beta = 0$, the problem becomes independent of the remaining market potential H_t. In particular, (12) reduces to

$$\hat{G}_t = \sum_{i=1}^{n}\big(p_{it}(\mathbf{q}_t) - c_{it}\big)q_{it} + \left(1 - \sum_{i=1}^{n} q_{it}\right)\hat{G}_{t+1}^* \tag{28}$$

which is concave in $\mathbf{q}_t$ and independent of H_t, leading to an even simpler solution.

Proposition 2. *In the Jun-Park model for which $\alpha = 1$ and $\beta = 0$,*

(i) $\hat{G}_t$ is concave in $\mathbf{q}_t = (q_{1t}, q_{2t}, ..., q_{nt})$.

(ii) let ρ_t^* be the unique solution to

$$\rho_t = \sum_{i=1}^{n}\frac{1}{b_t}\exp(a_{it} - 1 - b_t c_{it} - b_t\rho_t) + G_{t+1}^*\ . \tag{29}$$

Then $\hat{G}_t^* = \rho_t^*$.

(iii) the optimal pricing solution is given by

$$p_{it}^* = c_{it} + \frac{1}{b_t} + \hat{G}_t^*\ , \tag{30}$$

$$q_{it}^* = \frac{\exp(a_{it} - 1 - b_t c_{it} - b_t\hat{G}_t^*)}{1 + \sum_{j=1}^{n}\exp(a_{jt} - 1 - b_t c_{jt} - b_t\hat{G}_t^*)}\ , \tag{31}$$

where $\hat{G}_t^*$ is solved backward in time according to (29) and $\hat{G}_{T+1}^* = 0$.

When diffusion is driven predominantly by the innovation effect, the solution given in Proposition 2 suffices, which is simple to obtain due to independency on H_t.

4.4 Properties of Optimal Solution

Lemma 1 implies $r_{t-1}(\theta^*_{t-1}) \leq r_t(\theta^*_t)$ if and only if $F_t \leq \beta H_t$. Recall that H_t measures the remaining market potential which decreases with time and F_t measures the diffusion intensity which increases with time. Hence the term $(F_t - \beta H_t)$ monotonically increases in t; if the planning horizon covers a majority of the product's life cycle, this term should shift from negative to positive and crosses zero only once as t increases. Let $\bar{t}$ be the time period at which $(F_t - \beta H_t)$ switches from negative to positive. If the planning horizon only covers a portion of the product life cycle such that $(F_t - \beta H_t)$ is negative (positive) throughout the planning horizon, then $\bar{t} = T$ ($\bar{t} = 1$).

As a result, $r_t(\theta^*_t)$, and consequently θ^*_t change in a predictable manner with respect to time.

Corollary 1. *There exists $\bar{t} \in \{1, \ldots, T\}$ such that $r_{t-1}(\theta^*_{t-1}) \leq r_t(\theta^*_t)$ if $t < \bar{t}$ and $r_{t-1}(\theta^*_{t-1}) \geq r_t(\theta^*_t)$ if $t \geq \bar{t}$ where $t = 1, \ldots, T$.*

In the special case for which the cost and price sensitivity are time invariant, we can characterize the time trend of the optimal prices.

Corollary 2. *Suppose product qualities, costs and price sensitivities are time-invariant, i.e., $a_{it} = a_i$, $b_t = b$ and $c_{it} = c_i$ for $i = 1, \ldots, n$; $t = 1, \ldots, T$. Then there exists $\bar{t} \in \{1, \ldots, T\}$ such that $\theta^*_{t-1} \leq \theta^*_t$ if $t \leq \bar{t}$ and $\theta_{t-1} \geq \theta_t$ if $t \geq \bar{t}$. In addition, $p^*_{i,t-1} \leq p^*_{i,t}$ if $t \leq \bar{t}$ and $p^*_{i,t-1} \geq p^*_{i,t}$ if $t \geq \bar{t}$.*

Thus with time-invariant quality, cost and price sensitivity, the optimal prices of all products exhibit an up-and-then-down time trend.

In the special case of the Jun-Park model, it can be shown that the optimal price trend is a monotonic descending one.

Corollary 3. *Suppose $\alpha = 1$ and $\beta = 0$. Then (i) $\hat{G}^*_t$ decreases in t. (ii) Suppose that a_{jt}, b_t, and c_{jt} stay constant over time. Then the optimal price of each product, given by p^*_{it}, $i = 1, 2, \ldots, n$, decreases in t.*

The difference of price trend between the Jun-Park special case and the general model reflects the impact of the imitation effect. In the absence of imitation, the term $\hat{G}^*_t$ decreases in time, implying that the average profit from a customer remaining in the potential customer population decreases with time, which leads to decreasing prices. Intuitively, as time progresses, the time window to capture the remaining customers shortens; to counteract this effect, the firm has to lower prices to make the products more attractive. With imitation, the firm has an incentive to set the price low initially so as to speed up the word-of-mouth effect; this is most effective early in the diffusion. As the diffusion progresses, the firm faces a smaller pool of remaining customers and a shrinking time window to capture them and thus the pressure to reduce price increases over time. Samsung, for example, is reported to have priced its smart phones with such an up-and-then-down trend (Maling, 2012). This highlights the importance of aligning a firm's product introduction and pricing strategy with product diffusion characteristics: Optimal pricing strategy should be stage-dependent in a product life cycle; products that rely more on word-of-mouth for adoption are better candidates for discounted introduction prices than others.

A similar effect has been documented in the existing literature for a single-product diffusion – with strong word-of-mouth effect, it is optimal to use penetration pricing (i.e., increase price initially and decrease price later); with weak word-of-mouth, it is optimal to price skim (i.e., monotonically decrease price) (Kalish 1983, Horsky 1990). We have generalized this to the setting of multiple substitutable products, and the relationships in Lemma 1 and equation (16) quantify period-to-period price changes for every product in the choice set, which were not available in prior models. As we will show later in the paper, these relationships continue to hold for sequential product introductions, making the method a robust tool for analyzing price trend in a much broader context. In addition, our model allows more general price trend to be justified by the effect of diffusion as well as time-variant product qualities, price sensitivity, and cost; in this case, the price trend becomes further compounded with the time trend of these parameters. For example, since the optimal prices are given by

$$p_{it}^* = c_{it} + \frac{1}{b_t} + \theta_t^* \ , \tag{32}$$

it is evident that increasing price sensitivity and declining cost will superimpose additional descending trend on the optimal prices, and with our solution method, these impacts are easily quantified. Therefore, firms can use estimate of cost reduction (or price sensitivity change) to quantitatively project the optimal pricing strategy and diffusion path of each product over the product life cycle. Such capability facilitates other strategic decisions such as capacity and inventory planning. The common term θ_t^* depends on properties of all products and identifies the common price trend, emphasizing the need for joint determination of pricing of all products in a choice set.

From (32), we obtain the familiar "equal mark-up" property which is well-known for pricing under the MNL demand. Under this property, the product with higher cost is priced higher.

Corollary 4. *For any two products $i \neq j$, $p_{it}^* - c_{it} = p_{jt}^* - c_{jt}$, i.e., $p_{it}^* \geq p_{jt}^*$ if and only if $c_{it} \geq c_{ij}$.*

The choice probabilities and sales, however, are not equal across products, but instead differ across products based on product quality. The next corollary shows that products with higher "net quality" (i.e, higher $a_i - bc_i$ values) will, *ceteris paribus*, have higher swing in sales, i.e., steeper declining or increasing trend, than low net-quality products.

Corollary 5. *Suppose $a_{it} = a_i$, $c_{it} = c_i$ and $b_t = b$ for all i, t. If $a_i - bc_i > a_j - bc_j$ for $i \neq j$ then at the optimal price solution, $|q^*_{i,t-1} - q^*_{i,t}| > |q^*_{j,t-1} - q^*_{j,t}|$ and $|Z^*_{i,t-1} - Z^*_{i,t}| > |Z^*_{j,t-1} - Z^*_{j,t}|$ for $t = 2, \ldots, T$.*

Therefore, the gap between peak and non-peak sales is larger for high-end products than for low-end products. This has important implications for inventory and capacity decisions as products with larger demand swing are harder to manage and require more careful planning.

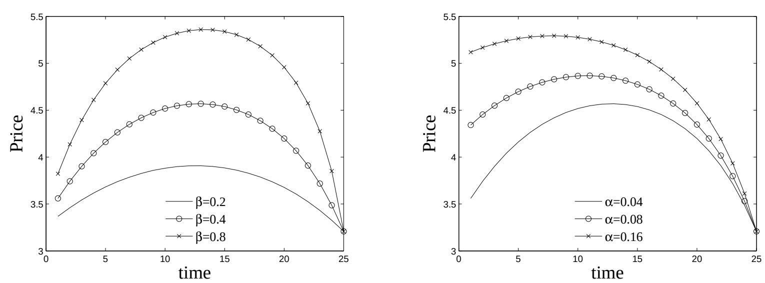
Lastly, we comment on the diffusion under the optimal prices. Note that sales under the optimal path are given by $Z_{it} = \frac{F_t H_t \exp(a_{it} - 1 - b_t\theta^*_t - b_t c_{it})}{1 + \sum_{j=1}^{n} \exp(a_{jt} - 1 - b_t\theta^*_t - b_t c_{jt})}$. Hence, the time trend of $F_t H_t$ indicates how the overall diffusion evolves over time while the exact time when sales peak for each product depends on the values of $a_{it} - b_t c_{it}$. For example, all else equal, a product with faster value depreciation will reach sales peak sooner than others while a product with faster cost reduction may start to see sales decline later than other products (as the firm is able to optimally reduce price of this product at a faster speed than for other products). Therefore, using our model, one can project not only the optimal price path, but also how each individual product diffuses into the market under the optimal prices, given the projections of a_{it}, b_t and c_{it} over time.

5 Numerical Illustration

The solution method we have developed in Section 4 applies to any intertemporal pattern of parameters a_{it}, b_t and c_{it}. However, conclusive analytical properties of the price trend can only be said when restricting to particular parameter patterns. Next, we numerically examine the impact of diffusion dynamics, the effect of product quality, price sensitivity and cost as well as the compounding effect of their intertemporal trend. We illustrate this with a single-product case (Sections 5.1 and 5.2) as figures plotting the time trends of multiple products under varying parameter values are illegible. To illustrate the effect of product interactions, we present the multiple product scenarios (Section 5.3) and we consider both the breadth (i.e., the number of products) and the depth (i.e., quality differences) of product variety.

5.1 Effects of Diffusion Parameters, Product Quality and Price Sensitivity

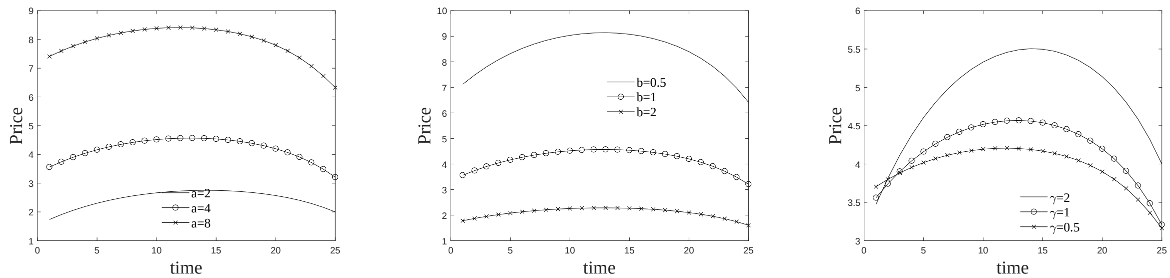
In Bass (1969), as well as in subsequent applications, products of different categories exhibit distinctive diffusion characteristics which is reflected in the relative magnitude of innovation versus imitation effect. In the examples shown in Figure 2, we compute the optimal price path by varying the values of α and β while keeping other parameter values fixed with $M = 1, Y_0 = 0, T = 25$; the quality, cost and price sensitivity are held constant at $a_t = 4, b_t = 1, c_t = 0$ for $t = 1, \ldots, T$ in order to focus on the diffusion effect. As the imitation effect β increases, the optimal price displays a steeper up-and-then-down trend, which is in agreement with the discussion following Corollary 2. However, as the innovation effect α becomes more dominant (Figure 2(b)), the initial price is to be set higher, which offsets the imitation-driven initial upward price trend and ultimately leads to a predominantly descending price trend.



(a) Effect of β (α is fixed at 0.04) (b) Effect of α (β is fixed at 0.4)

Figure 2: Optimal Prices versus Diffusion Parameters.

The effect of product quality a and price sensitivity b are illustrated in Figures 3(a) and 3(b) respectively. As expected, higher quality or lower price sensitivity leads to higher prices. In the literature, a common scale factor of a and b parameters is interpreted as a measure for the cognitive limitations of customers, i.e., choice probability is given by $q_i = \frac{e^{u_i/\gamma}}{\sum_{j=0}^{n} e^{u_j/\gamma}}$ with large γ values indicating less rational choice behaviors (Su, 2008). Figure 3(c) shows how such customer rationality affects the optimal prices by scaling down values of a and b by a common parameter γ. As consumer rationality decreases (i.e., γ increases), the optimal prices are generally higher, which is intuitive since customers are less able to differentiate between good and poor choices (i.e., less sensitive to price or quality differences). In addition, we observe larger intertemporal price adjustments for higher γ values, which is less intuitive but also reflects the firm's adaptation to reduced rationality - with irrational customers, it takes a larger initial discount to induce word-of-mouth as well as a larger markdown later in the diffusion to counteract the depletion of potential customers.

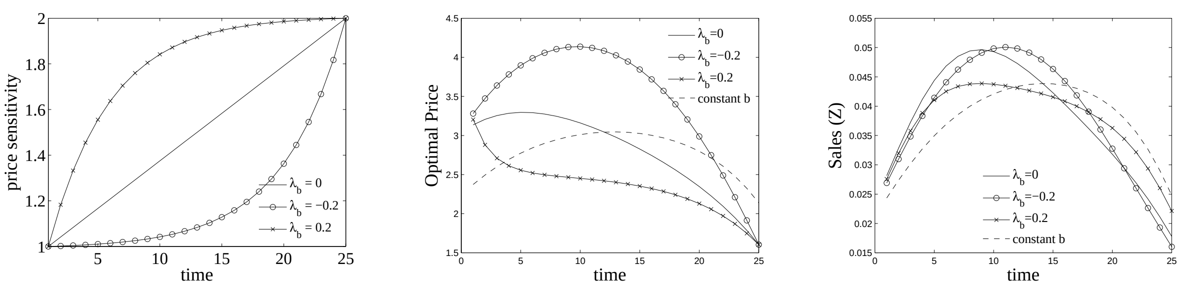


(a) $b = 1$ (b) $a = 4$ (c) $a = 4, b = 1$

Figure 3: Optimal Prices versus Parameters a and b (other parameters are the same as in Figure 2.)

5.2 Effect of Time-Varying Price Sensitivity and Cost

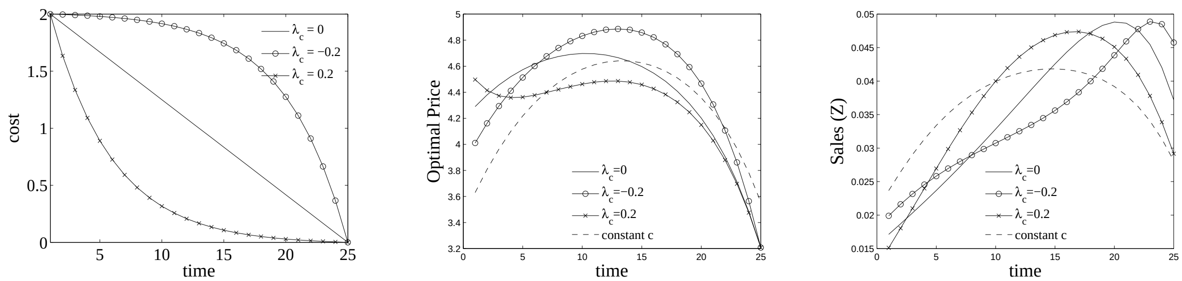
While the examples in Figure 2 have time-invariant cost and price sensitivity (so as to isolate the effect of diffusion on prices), in most realistic settings these vary with time and typically cost declines over time and price sensitivity increases over time. Figure 4 illustrates the optimal price path and the corresponding sales for each price sensitivity curve specified in Figure 4(a) where the curve is concave (convex) when $\lambda_b > 0$ ($\lambda_b < 0$) and linear when $\lambda_b = 0$. For reference, we also include the optimal price path and sales for constant price sensitivity with $b(t) = (b_1 + b_T)/2$, shown as the dashed curve. Evidently, the optimal price displays a stronger declining trend with increasing price sensitivity than with constant b. With sharply increasing price sensitivity ($\lambda_b = 0.2$), we obtain a monotonically descending price trend, masking any imitation-driven upward price trend (Figure 4(b)). In addition, with increasing price sensitivity, peak sales occur earlier in the planning horizon because the firm has a stronger incentive to sell early when the market is less price-sensitive (Figure 4(c)).



(a) Increasing Price Sensitivity (b) Effect on Optimal Prices (c) Effect on Sales at Optimal Prices

Figure 4: Effect of Increasing Price Sensitivity (price sensitivity curves in Figure 4(a) are given by the equation $b_t = b_1 + (b_T - b_1) \cdot \frac{1-e^{-\lambda_b(t-1)}}{1-e^{-\lambda_b(T-1)}}$ with $b_1 = 1, b_T = 2$.)

We observe in Figure 5 that declining cost (shown in Figure 5(a)) results in a stronger downward price trend (Figure 5(b)) and delayed peak sales (Figure 5(c)) relative to constant cost (which is fixed at $c_t = (c_1 + c_T)/2$ and shown as the dashed curve). That is, anticipating cost decline and potentially higher margin later in time, the firm finds it more profitable to delay peak sales; like in the case of increasing price sensitivity, this dampens the imitation-driven upward price trend. Interestingly, although both price-sensitivity increase and cost decline reduce the upward price trend



(a) Declining Cost (b) Effect on Optimal Prices (c) Effect on Sales at Optimal Prices

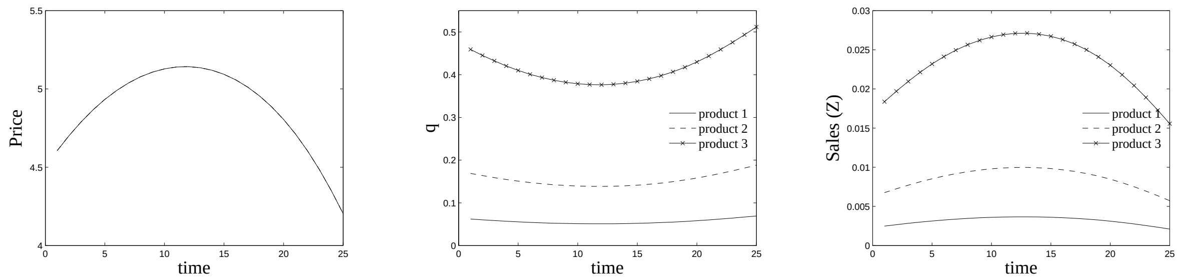
Figure 5: Effect of Declining Cost (cost curves in Figure 5(a) are given by equation $c_{it} = c_1 - (c_1 - c_T) \cdot \frac{1-e^{-\lambda_c(t-1)}}{1-e^{-\lambda_c(T-1)}}$ with $c_1 = 2, c_T = 0$.)

during the early stage of diffusion, the former accelerates peak sales whereas the latter postpones peak sales.

5.3 Effect of Product Interactions

To isolate the effect of product interactions, we consider examples with time-invariant qualities, costs and price sensitivities. Let $M = 1, Y_0 = 0, T = 25$, $\alpha = 0.04$, $\beta = 0.2$, $n = 3$, $a_{1t} = 3, a_{2t} = 4, a_{3t} = 5$, $b_t = 1$, $c_{it} = 0$ for $t = 1, \ldots, T$. Apply Algorithm 1 to obtain the optimal θ_t, p_{it}, q_{it} and Z_{it} for $i = 1, 2, 3$ and $t = 1, \ldots, T$. Since cost and price sensitivity are symmetric across products and time-invariant, the optimal prices of the products are equal and given by $p^*_{it} = c_{it} + 1/b_t + \theta_t = 1 + \theta_t$. Figures 6(a)-6(c) illustrate the optimal price, choice probability and sales, respectively. Note that the phenomenon identified in Corollary 5 is apparent: Not only are sales and choice probability of the high-quality product (product 3 with $a_{3t} = 5$) much higher, but they also exhibit stronger time trend (i.e., larger swing) than the low-quality product (product 1 with $a_{1t} = 3$).

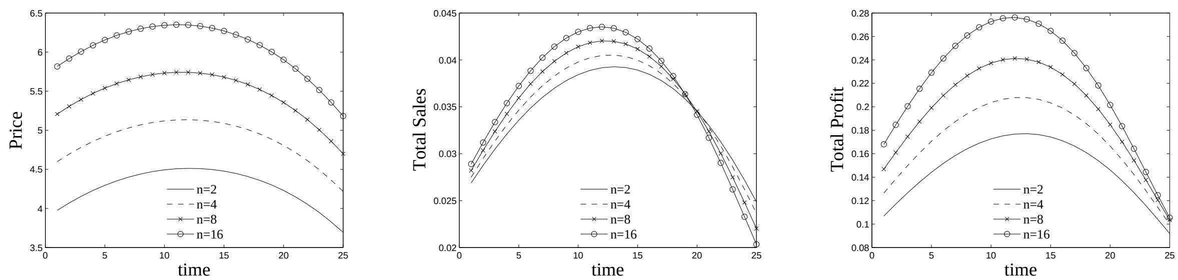
Next, we examine how the breadth of product variety, measured by the number of products offered in the choice set, affects the optimal solution. For the ease of comparison, we let $a_{it} = 4$ for $i = 1, \ldots, n$, $t = 1, \ldots, T$ and vary n while keeping all other parameters the same as in Figure 6.



(a) Optimal Price (b) Optimal Choice Probability (c) Optimal Sales

Figure 6: An Example Illustrating Corollary 5.

Results are shown in Figure 7. As n increases, both the optimal prices and the firm's profit increase.

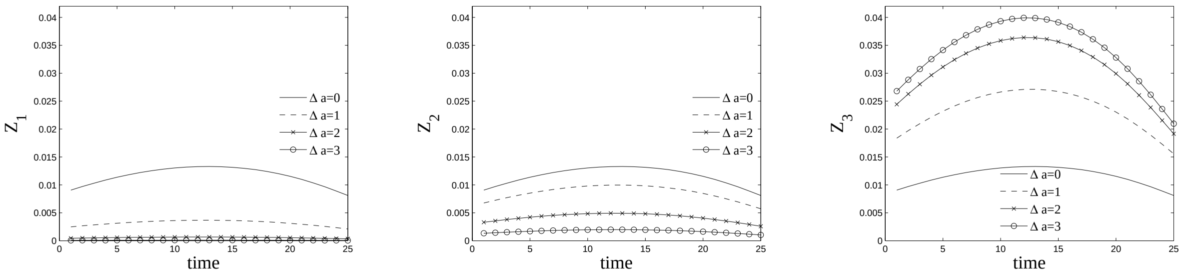


(a) Optimal Prices (b) Total Sales (c) Total Profit

Figure 7: Impact of Product Variety Breadth.

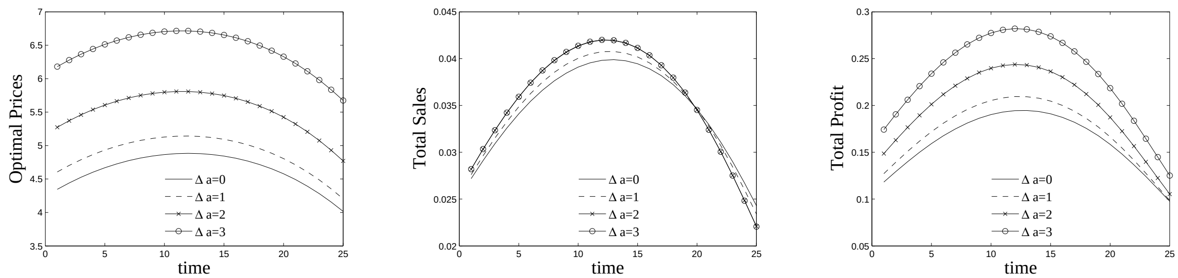
In addition, with more product options for the customers, the diffusion takes off more rapidly and reaches peak sooner. A similar pattern is observed for the optimal prices, which can be explained when considered in conjunction with diffusion: By adding breadth (i.e., more products), the firm is able to capture more customers early on, which speeds up the diffusion and consequently lessens the imitation-driven upward price trend but strengthens the depletion-driven downward price trend; this results in the optimal prices declining sooner. Such observations may help justify firms' strategy of introducing multiple variations of a new product, although the increased profit shown in Figure 7(c) need be balanced with the cost of managing more product variety which is omitted here.

Figure 9 illustrates the effect of depth of product variety, measured by the quality differences among products. We let $a_{1t} = 4 - \Delta a$, $a_{2t} = 4$, and $a_{3t} = 4 + \Delta a$ for $t = 1, \ldots, T$ and vary the value of Δa (other parameters are the same as in Figure 6). As the depth (i.e., Δa) increases, the optimal solution becomes increasingly dominated by the highest-quality product (i.e., product 3) as shown in Figure 8. As a result, not only is the firm able to charge higher prices (Figure 9(a)), but it also achieves a faster diffusion (Figure 9(b)) and a higher profit (Figure 9(c)). Therefore, having a more competitive leading product (versus three average products) accelerates the diffusion and benefits the firm.



(a) Sales of Product 1 (b) Sales of Product 2 (c) Sales of Product 3

Figure 8: Impact of Product Variety Depth on Sales.



(a) Optimal Prices (b) Total Sales (c) Total Profit

Figure 9: Impact of Product Variety Depth.

6 Extensions

6.1 Sequential Product Introductions

In Section 4, we have examined multiple substitute products going through diffusion concurrently. In practice, these products may be introduced nonconcurrently but with overlapping periods of sales. For example, Amazon sometimes introduces the paperback version of a book a few months after the hardcover version. Technology firms often introduces successive generations of products which are substitutes of one another and have overlapping sales. Indeed, the June-Park model itself considers such a scenario. Researchers have extended Bass (1969) and similar models to address diffusion of overlapping product generations (see Meade and Islam (2006) and Peres et al. (2010)). Most prominently, Norton and Bass (1987) assume that the adoption of the next generation product originates from both the untapped market potential and adopters of the previous generation upgrading to the new product. Their model, however, only addresses substitution between adjacent generations, not across multiple generations; it grows increasingly more complex with additional number of products and it is not clear how one can incorporate prices and product characteristics without making the model intractable. A related stream of work builds upon the population growth model (for example, Mahajan and Muller (1996) on the demand for IBM mainframe computers, Kim et al. (2000) on subscriptions of telecommunication services, and Li et al. (2013) on sales of microprocessors). These models perform well for fitting empirical data, but offer poor tractability for price optimization. Consequently, normative price optimizations based on multi-product diffusion models are few and are typically in a stylized setting with only two products (e.g., Bayus 1992) which is unsuitable for decision support. In this section, we show that the mathematical model presented in Section 4 can be extended to this scenario.

Let M_t be the total market potential at the beginning of period t, which may be adjusted based on market change. For example, when a new product is introduced, M_t may change due to additional market potential brought forth by the new product. This modeling choice is commonly adopted in the literature for successive product generations (see, for example, Norton and Bass (1987) and Jun and Park (1999)). Let κ_t be the set of products being offered at time t. As new products are introduced or old products are retired over time, the set κ_t and the total market potential M_t evolve accordingly. Assuming that M_t and κ_t are both exogenously given, we solve the pricing problem. The choice probability for product i, $i \in \kappa_t$ in period t is

$$q_{it} = \frac{\exp(a_{it} - b_t p_{it})}{1 + \sum_{j \in \kappa_t} \exp(a_{jt} - b_t p_{jt})}$$

and the demand for product i in period t is given by $Z_{it} = (M_t - Y_{t-1})(\alpha + \beta Y_{t-1})q_{it}$, where $Y_t = \sum_{s=1}^{t} \sum_{i \in \kappa_s} Z_{is}$. From this, we can derive

$$M_{t+1} - Y_t = \Delta M_{t+1} + (M_t - Y_{t-1})\Big[1 - (\alpha + \beta Y_{t-1}) \sum_{i \in \kappa_t} q_{it}\Big]$$

where $\Delta M_t := M_t - M_{t-1}$. We note the two main differences in the problem formulation for the sequential introduction model: (i) the set of product κ_t is allowed to change with time, (ii) there is a period-to-period change ΔM_t for the total market potential.

It can be shown that Lemma 1 and Proposition 1 continue to hold under this generalization (see the online appendix for proofs). Such results warrant a similar efficient algorithm for solving the model with sequential product introductions (provided in the online appendix). In addition, it can be verified that Corollaries 4 and 5 continue to hold. Suppose the remaining market potential H_t monotonically decreases with time, then Corollaries 1 and 2 hold in the sequential model. However, this may not always be the case if M_t varies with time; for example, M_t could increase due to the introduction of a new product, which may cause a temporary increase in H_t and consequently a price kink. Nonetheless, the relationship in Lemma 1 still helps identify the overall price trend with such potential kinks. Therefore, the model framework and our solution approach are flexible for studying sequential product introductions and the impact of the pricing decisions in this context. Our method accommodates multiple product introductions without creating prohibitive computation burdens.

6.2 Stochastic Demand

Our main model in Section 4 considers the expected demand, following the current literature on diffusion-choice models. Next, we extend this to stochastic demand, which is useful for dynamically pricing the products based on the realized adoption.

Let Y be the level of cumulative adoption, and $M - Y$ be the remaining market potential. The diffusion intensity is given by $\breve{\alpha}+\breve{\beta}Y$ where $\breve{\alpha}$ parameterize the innovation effect and $\breve{\beta}$ parameterizes the imitation effect. Let $(M - Y)(\breve{\alpha} + \breve{\beta}Y)$ be the probability that there is a customer arrival in each time period. We choose the time unit sufficiently small that there can be at most one arrival in each period. For a given customer arrival, the customer chooses among n products (and the no-purchase option) with the choice probability given in equation (3). If the customer purchases product i, then the firm obtains profit $p_{it} - c_{it}$ and the cumulative adoption increases by one unit. Define J_t as the profit-to-go from period t onward. The firm maximizes the total profit over the planning horizon of T time periods by solving the following dynamic programming problem:

$$
\begin{aligned}
J_t(Y, \mathbf{q}_t) &= (M - Y)(\breve{\alpha} + \breve{\beta}Y) \sum_{i=1}^{n} q_{it} \left(p_{it}(\mathbf{q}_t) - c_{it} + J^*_{t+1}(Y + 1) \right) \\
&\quad + \left[1 - (M - Y)(\breve{\alpha} + \breve{\beta}Y) + (M - Y)(\breve{\alpha} + \breve{\beta}Y) q_{0t} \right] J^*_{t+1}(Y) \qquad (33) \\
\text{where} &\quad J^*_t(Y) := \max_{\mathbf{q}_t} J_t(Y, \mathbf{q}_t) \ , \quad t = 1, \ldots, T, \quad J^*_{T+1}(Y) = 0 \ .
\end{aligned}
$$

Note that the adoption level stays at Y in period $t + 1$ if there is no arrival in period t, which occurs with probability $1 - (M - Y)(\breve{\alpha} + \breve{\beta}Y)$, or there is an arrival but the customer chooses not to purchase any of the n products, which occurs with probability $(M - Y)(\breve{\alpha} + \breve{\beta}Y)q_{0t}$. This formulation is consistent with stochastic MNL demand models adopted in the literature (e.g., Dong et al. (2009)); in our formulation, the arrival rate is endogenous and given by the diffusion equation instead of exogenous and constant.

Taking derivative of $J_t(Y, \mathbf{q}_t)$ with respect to q_{it}, we obtain the following first-order condition

(derivation provided in the online appendix)

$$p_{it} - c_{it} - \frac{1}{b_t} = \sum_{i'=1}^{n} \frac{q_{i't}/q_{0t}}{b_t} + J_{t+1}^*(Y) - J_{t+1}^*(Y+1) \ . \tag{34}$$

The right hand side is independent of the index i. Thus at optimality, equal markup holds and there exists θ_t such at

$$p_{it} - c_{it} - \frac{1}{b_t} = \theta_t \quad \forall i. \tag{35}$$

We can rewrite (34) as

$$\theta_t = \frac{1}{b_t} \sum_{i=1}^{n} e^{a_{it}-1-b_t c_{it}-b_t\theta_t} + J_{t+1}^*(Y) - J_{t+1}^*(Y+1) \ , \tag{36}$$

or equivalently,

$$r_t(\theta_t) = J_{t+1}^*(Y) - J_{t+1}^*(Y+1), \tag{37}$$

where

$$r_t(\theta_t) := \theta_t - \frac{1}{b_t} \sum_{i=1}^{n} e^{a_{it}-1-b_t c_{it}-b_t\theta_t} \ . \tag{38}$$

Let $\theta_t^*(Y)$ be the solution to the first-order condition in (37). We derive a similar connection between $r_{t-1}(\theta_t^*(Y))$ and $r_t(\theta_t^*(Y))$ as in Lemma 1.

Lemma 2.

$$\begin{aligned} r_{t-1}(\theta_{t-1}^*(Y)) - r_t(\theta_t^*(Y)) \quad &= \quad (M-Y)(\breve{\alpha} + \breve{\beta} Y) \sum_{i=1}^{n} \frac{e^{a_{it}-b_t c_{it}-1-b_t\theta_t^*(Y)}}{b_t} \\ &\quad - (M-Y-1)(\breve{\alpha} + \breve{\beta} Y + \breve{\beta}) \sum_{i=1}^{n} \frac{e^{a_{it}-b_t c_{it}-1-b_t\theta_t^*(Y+1)}}{b_t} \ . \end{aligned} \tag{39}$$

This relationship simplifies the dynamic programming problem and leads to a unique optimal price solution.

Proposition 3. *The solution to* (37) *is unique for any given* $Y = 0, \ldots, M-1$ *and* $t = 1, \ldots, T$.

Rewrite the relationship in Lemma 2 as

$$\begin{aligned} r_{t-1}(\theta_{t-1}^*(Y)) \quad &= \quad \theta_t^*(Y) + [(M-Y)(\breve{\alpha} + \breve{\beta} Y) - 1] \sum_{i=1}^{n} \frac{e^{a_{it}-b_t c_{it}-1-b_t\theta_t^*(Y)}}{b_t} \\ &\quad - (M-Y-1)(\breve{\alpha} + \breve{\beta} Y + \breve{\beta}) \sum_{i=1}^{n} \frac{e^{a_{it}-b_t c_{it}-1-b_t\theta_t^*(Y+1)}}{b_t} \ . \end{aligned} \tag{40}$$

Given $\theta_t^*(Y)$ for $Y = 0, \ldots, M-1$, it is straightforward to compute $r_{t-1}(\theta_{t-1}^*(Y))$; due to the monotonicity of $r_{t-1}(\cdot)$, one can then apply a bisection search to find $\theta_{t-1}^*(Y)$. Therefore, based on these analytical results, the dynamic programming problem is easily solved with backward computation. Furthermore, we can derive the following structural properties of the optimal solution.

Proposition 4. *(i) $\theta_t^*(Y+1) \geq \theta_t^*(Y)$ for $Y = 0, \ldots, M-1$ and $t = 1, \ldots, T$. (ii) $r_{t-1}(\theta_{t-1}^*(Y)) \geq r_t(\theta_t^*(Y))$ if $Y \geq \frac{M-1}{2} - \frac{\breve{\alpha}}{2\breve{\beta}}$, $t = 2, \ldots, T$.*

Corollary 6. *Suppose product qualities, costs and price sensitivities are time-invariant, i.e., $a_{it} = a_i$, $b_t = b$ and $c_{it} = c_i$ for $i = 1, \ldots, n$; $t = 1, \ldots, T$. Let $p_{it}^*(Y)$ be the optimal price of product i at period t when the current adoption level is Y. Then (i) $p_{it}^*(Y+1) \geq p_{it}^*(Y)$ for $Y = 0, \ldots, M-1$ and $t = 1, \ldots, T$ (ii) $p_{it-1}^*(Y) \geq p_{it}^*(Y)$ if $Y \geq \frac{M-1}{2} - \frac{\breve{\alpha}}{2\breve{\beta}}$, $t = 2, \ldots, T$.*

This indicates that, all else equal, the optimal price is higher at a higher adoption level. In addition, once adoption exceeds a certain level, the optimal prices decrease (weakly) over time.

The stochastic model examines the perspective of dynamic pricing and is solvable using a similar technique developed for the main model. However, it does not supersede the main model for two reasons. First, the main model is based on a family of models that have been proven to work empirically. Methods are well established to parameterize such models using sales data. The stochastic model is based on infinitesimally small time intervals and single-unit arrivals; it is more challenging to parameterize such a conceptual model with real data. In practice, single-unit demand arrival is not observed in most data set, one has to use the main model parameters α and β and rescale them to approximate parameters $\breve{\alpha}$ and $\breve{\beta}$ for the stochastic model. Second, the stochastic model focuses more on the tactical response strategy of dynamic pricing, i.e., how should prices be adjusted at a given time based on the stochastic realization of the adoption level, while the main model allows us to characterize the optimal adoption path along with the optimal price trend *a priori*, and thus offers practical guidance for strategic pricing over product life cycles.

6.3 Product-specific Word-of-mouth Effect

It is possible that the word-of-mouth effect may also be driven by the cumulative adoption of each individual product in addition to the aggregate adoption. For example, a consumer who bought an iPhone with a 256 GB memory may communicate the benefit of a large size memory to non-adopters and positively influence the sales of the 256GB model. Hence it may be desirable to consider product-specific imitation in diffusion-choice models. In this extension, we examine a general model in which this product-specific word-of-mouth effect is also considered. In particular, the demand for product i in period t is given by

$$Z_{it} = (M - Y_{t-1})(\alpha + \beta Y_{t-1} + \gamma Y_{it-1})q_{it} \tag{41}$$

where Y_{it-1} is the cumulative sales of product i by the end of period $t-1$, $\gamma > 0$ parameterizes the intensity of the product-specific imitation effect, and other terms are as defined in Section 3. As we show next, accounting for the product-specific imitation effect necessitates tracking product-specific diffusion intensity, leading to a much more complex problem. As a result, the reduction of choice probability vector $\boldsymbol{q}_t$ to a single variable θ_t adopted in Section 4.1 does not carry through. To see this, note that the remaining market potential and the diffusion intensity become

$$H_{t+1} = H_t \left(1 - \sum_{j=1}^{n} F_{jt} q_{jt}\right) \tag{42}$$

$$F_{it+1} = F_{it} + H_t \left(\beta \sum_{j=1}^{n} F_{jt} q_{jt} + \gamma F_{it} q_{it}\right) \tag{43}$$

and the optimization problem becomes

$$J_t = \max_{\substack{q_{it}\in[0,1],\\ i=1,\cdots,n}} \left[H_t \sum_{i=1}^{n} F_{it}\big(p_{it}(\mathbf{q}_t) - c_{it}\big)q_{it} + J_{t+1}\right] ,$$

where $J_{T+1} = 0$. Denote the vector of product-specific diffusion intensity with $\boldsymbol{F}_t = (F_{1t}, F_{2t}, \ldots, F_{nt})$. It can be shown that $H_t = M - \frac{\sum_{i=1}^{n} F_{it} - n\alpha}{n\beta+\gamma}$. Thus from (42) and (43), the vector $\boldsymbol{F}_{t+1}$ is a function of $\boldsymbol{F}_t$ and $\mathbf{q}_t$. Define

$$\breve{G}_t(\boldsymbol{F}_t, \mathbf{q}_t) := \sum_{i=1}^{n}\big(p_{it}(\mathbf{q}_t) - c_{it}\big)q_{it}F_{it} + \left(1 - \sum_{i=1}^{n} F_{it}q_{it}\right)\breve{G}^*_{t+1}(\boldsymbol{F}_{t+1}(\boldsymbol{F}_t, \mathbf{q}_t))$$

$$\text{where } \breve{G}^*_t(\boldsymbol{F}_t) = \max_{q_{it}\in[0,1], i=1,\ldots,n} \breve{G}_t(\boldsymbol{F}_t, \mathbf{q}_t) .$$

We note that the interpretation of $\breve{G}(\cdot)$ differs from that of $G(\cdot)$ in Section 4 due to the term F_{it}.

The first-order condition of optimality is given by (derivation provided in the online appendix):

$$p_{it}(\boldsymbol{q}_t) - c_{it} - \frac{1}{b_t} = \sum_{i'=1}^{n} \frac{q_{i't}}{b_t(1 - \sum_{j=1}^{n} q_{jt})} \cdot \frac{F_{i't}}{F_{it}} + \breve{G}^*_{t+1} - H_{t+1}\left(\gamma \frac{\partial \breve{G}^*_{t+1}}{\partial F_{it+1}} + \beta \sum_{i'} \frac{\partial \breve{G}^*_{t+1}}{\partial F_{i't+1}}\right) . \tag{44}$$

Contrasting this with equation (15), we observe that the right hand side is dependent on the index i through the term $\gamma \frac{\partial G^*_{t+1}}{\partial F_{it+1}}$. Hence the equal mark-up property no longer holds with product-specific imitation. As a result, the transformation given in equation (16)-(17) is no longer viable.

However, despite the added complexity of product-specific imitation, we can reduce the longitudinal complexity employing a similar technique as in Section 4. Define

$$R_{it}(\mathbf{p}_t) := p_{it} - c_{it} - \frac{1}{b_t} - \sum_{i'=1}^{n} \frac{e^{a_{i't} - b_t p_{i't}}}{b_t} \frac{F_{i't}}{F_{it}} \ .$$

We can rewrite the first-order condition as

$$R_{it} = \breve{G}^*_{t+1} - H_{t+1} \left(\gamma \frac{\partial G^*_{t+1}}{\partial F_{it+1}} + \beta \sum_{i'=1}^{n} \frac{\partial G^*_{t+1}}{\partial F_{i't+1}} \right) \ . \tag{45}$$

Proposition 5. *The optimal price solution* $\mathbf{p}^*_t, t = 1, \ldots, T$ *satisfies*

$$\begin{aligned} R_{it-1}(\mathbf{p}^*_{t-1}) - R_{it}(\mathbf{p}^*_t) \quad = \quad & \sum_{i'=1}^{n} \left(\frac{1}{b_t} + \sum_{j=1}^{n} \frac{e^{a_{jt} - b_t p^*_{jt}}}{b_t} \cdot \frac{F_{jt}}{F_{i't}} \right) q_{i't}(\mathbf{p}^*_t)(F_{i't} - \beta H_t) \\ & - \gamma H_t \left(\frac{1}{b_t} + \sum_{j=1}^{n} \frac{e^{a_{jt} - b_t p^*_{jt}}}{b_t} \cdot \frac{F_{jt}}{F_{it}} \right) q_{it}(\mathbf{p}^*_t) \quad \forall i = 1, \ldots, n \ . \end{aligned} \tag{46}$$

Proposition 5 connects the optimal price vectors of two adjacent time periods, consequently enabling reduction of the longitudinal complexity. The relationship in Proposition 5 decouples the price decision in each period from all future periods except the immediate next one. Proceeding backward in time, one can solve a system of n equations in each period to obtain the optimal solution path. Although still a problem of substantial size (particularly when the number of products is large), the complexity in the time dimension is significantly reduced.

The time trend of the optimal prices is not as clear-cut as in the case without product-specific imitation, but some general character persists such as indicated in the following corollary.

Corollary 7. *There exist* $\underline{t}, \bar{t} \in \{1, \ldots, T\}$ *such that* $R_{it-1}(\mathbf{p}^*_{t-1}) \leq R_{it}(\mathbf{p}^*_t)$ *for* $t \leq \underline{t}$ *and* $R_{it-1}(\mathbf{p}^*_{t-1}) \geq R_{it}(\mathbf{p}^*_t)$ *for* $t \geq \bar{t}$.

7 Conclusion

We have examined pricing of multiple products in a diffusion process in which customers choose among the products and the choice probabilities are given by the MNL model. Technical challenge of the problem arises in two ways. First, interactions among multiple substitutable products complicate the pricing problem as the price of one product affects returns on this and all other products in the portfolio. Second, the diffusion dynamics, in particular, the imitation effect creates complex dependency of future sales on current prices, leading to a seemingly unsolvable problem. In this paper, we overcome both obstacles and present a scalable algorithm to solve the problem for n products over T periods, which was unattainable with existing models. We establish uniqueness of the optimal solution, characterize the time path of the optimal prices, and show how the diffusion dynamics and product-specific differences enter the optimal solution. Our model applies to both simultaneous and sequential product introductions and easily accommodates changes to the product portfolio as the firm introduces or retires products during the planning horizon. Moreover, we consider extensions including stochastic demand and product-specific imitation effect.

We identify a few crucial elements determining the time path of the optimal prices in this paper. First, the relative magnitude of the innovation versus imitation effect of diffusion can lead to differing price patterns. Second, quality differences across products affect the optimal solution. On one hand, the structure of the optimal solution indicates that the optimal price gap only has to do with cost and price sensitivity and is independent of quality; this is in spirit consistent with the "equal mark-up" finding for pricing under MNL. On the other hand, the optimal sales of products with higher net quality exhibit steeper increasing or declining trend than those with lower net quality. Managerially, this means that within the same choice set, the gap between peak and non-peak sales is larger for high-end products than for low-end products. This has important implications for inventory and capacity decisions as products with larger demand swing are harder to manage and require more careful planning. Third, our model is flexible and allows time-varying cost and price sensitivities, whose impacts compound the price trend but are easily quantified using our solution method. We show that increasing price sensitivity and declining cost both reduce the upward price trend, but the former accelerates peak sales whereas the latter delays peak sales. Lastly, both the breadth and depth of product variety accelerate the diffusion, increasing the optimal prices and the firm's profit.

We conclude the paper by discussing two practical considerations concerning our method. First, our model is a discrete-time model, which allows a firm to optimally adjust its product prices at the same frequency as the demand model. If a firm chooses to or is constrained to adjust price at a lesser frequency, we describe in the online appendix a heuristic implementation of our method and show that it can adapt to such practical considerations with sound performance. Second, there may be practical scenarios in which firms are constrained to non-increasing prices even though it might not be optimal. In this case, solving the period t problem requires not only the remaining market potential, but also the last-period prices. As a result, the current method becomes less tractable. However, insights afforded by our model shed light on how this may affect the price solution. In the absence of imitation, we have shown analytically that the optimal price trend is monotonically decreasing when other parameters are time invariant (Corollary 3). The imitation effect may cause an initial upward price trend because lower price early in the diffusion speeds up the word-of-mouth effect, benefiting the firm. If the firm is constrained to non-increasing prices, then it may not be able to set lower prices early in the diffusion cycle to take full advantage of the word-of-mouth effect; this will consequently slow down the diffusion and drive the firm to apply lower prices later in the diffusion, relative to the unconstrained price path.

Acknowledgement

The author is grateful to Professor Scott Webster for his advice on the paper. The author also thanks two anonymous referees, the Associate Editor and Professor Dan Edelman for their valuable and constructive comments and suggestions, which made this paper an outcome of teamwork by the author and the review team.

References

Amazon Website 2017. Amazon Hot New Releases. Published Online https://www.amazon.com/gp/new-releases/books/18/ref=zg_bsnr_nav_b_1_b.

Bass, F. 1969. A new product growth model for consumer durables. *Management Science* 15:215–227.

Bass, F. 1980. The relationship between diffusion rates, experience curves, and demand elasticities for consumer durable technological innovations. *The Journal of Business* 53 (3): S51–S67.

Bass, F., T. Krishnan, and D. Jain. 1994. Why the Bass model fits without decision variables. *Marketing Science* 13 (3): 203–223.

Bayus, B. L. 1992. The dynamic pricing of next generation consumer durables. *Marketing Science* 11 (3): 251–265.

Chen, C., and C. Watanabe. 2006. Diffusion, substitution and competition dynamism inside the ICT market: The case of Japan. *Technological Forecasting and Social Change* 73 (6): 731–759.

Davis, J., G. Gallego, and H. Topaloglu. 2014. Assortment optimization under variants of the nested logit model. *Operations Research* 62 (2): 250–273.

Dolan, R. J., and A. P. Jeuland. 1981. Experience curves and dynamic demand models: implications for optimal pricing strategies. *Journal of Marketing* 45 (1): 52–62.

Dong, L., P. Kouvelis, and Z. Tian. 2009. Dynamic pricing and inventory control of substitute products. *Manufacturing and Service Operations Management* 11 (2): 317–339.

Du, C., W. L. Cooper, and Z. Wang. 2016. Optimal pricing for a multinomial logit choice model with network effects. *Operations Research* 64 (2): 441–455.

Entwistle, S. 2012. HTC, Apple & Samsung Top Smartphone Retail Price Decline Comparison. *https://www.strategyanalytics.com/strategy-analytics/blogs/devices/pricetrax/pricetrax/2012/09/27/htc-apple-samsung-top-smartphone-retail-price-decline-comparison.*

Entwistle, S. 2015. Apple and HTC avoid major retail price erosion. *https://www.strategyanalytics.com/strategy-analytics/blogs/devices/pricetrax/pricetrax/2015/12/21/apple-and-htc-avoid-major-retail-price-erosion—unlike-other-flagship-smartphones-complimentaryreport-offer#.WfoTU4hrw2w.*

Gallego, G., and R. Wang. 2014. Multi-product price optimization and competition under the nested logit model with product-differentiated price sensitivities. *Operations Research* 62 (2): 450–461.

Hanson, W., and K. Martin. 1996. Optimizing multinomial logit profit functions. *Management Science* 42 (7): 992–1003.

Horsky, D. 1990. A diffusion model incorporating product benefits, price, income and information. *Marketing Science* 9 (4): 342–365.

Huh, W. T., and H. Li. 2015. Technical Note – Pricing under the nested attraction model with a multi-stage choice structure. *Operations Research* 63 (4): 840–850.

Jain, D. C., and R. C. Rao. 1990. Effect of price on the demand for durables:modeling, estimation, and findings. *Journal of Business Economics and Statistics* 8 (2): 163–170.

Jun, D. B., and Y. S. Park. 1999. A choice-based diffusion model for multiple generations of products. *Technological Forecasting and Social Change* 61:45–58.

Kalish, S. 1983. Monopolist pricing with dynamic demand and production cost. *Marketing Science* 2 (2): 135–159.

Kamakura, W., and S. Balasubramanian. 1988. Long-term view of the diffusion of durables: a study of the role of price and adoption influence process via tests of nested models. *International Journal of Research in Marketing* 5:1–13.

Kim, N., D. Chang, and A. Shocker. 2000. Modeling inter-category dynamics for a growing information technology industry. *Management Science* 46 (4): 496–512.

Kim, W.-J., J.-D. Lee, and T.-Y. Kim. 2005. Demand forecasting for multigenerational products combining discrete choice and dynamics of diffusion under technological trajectories. *Technological Forecasting and Social Change* 72 (7): 825–849.

Krishnan, T. V., F. M. Bass, and D. C. Jain. 1999. Optimal pricing strategy for new products. *Management Science* 45 (12): 1650–1663.

Lee, C.-Y., J.-D. Lee, and Y. Kim. 2008. Demand forecasting for new technology with a short history in a competitive environment: the case of the home networking market in South Korea. *Technological Forecasting and Social Change* 75 (1): 91–106.

Lee, J., Y. Cho, J.-D. Lee, and C.-Y. Lee. 2006. Forecasting future demand for large-screen television sets using conjoint analysis with diffusion model. *Technological Forecasting and Social Change* 73 (4): 362–376.

Li, G., P. Rusmevichientong, and H. Topaloglu. 2015. The d-level nested logit model: Assortment and price optimization problems. *Operations Research* 63 (2): 325–342.

Li, H., D. Armbruster, and K. G. Kempf. 2013. A population-growth model for multiple generations of technology products. *Manufacturing and Service Operations Management* 15 (3): 343–360.

Li, H., and W. T. Huh. 2011. Pricing multiple products with the multinomial logit and nested logit models: Concavity and implications. *Manufacturing and Service Operations Management* 13 (4): 549–563.

Li, H., and W. T. Huh. 2012. Optimal pricing for a short life-cycle product when customer price-sensitivity varies over time. *Naval Research Logistics* 59 (7): 552–576.

Li, H., and S. Webster. 2017. Optimal pricing of correlated product options under the paired combinatorial logit model. *Operations Research* 65 (5): 1215–1230.

Mahajan, V., and E. Muller. 1996. Timing, diffusion, and substitution of successive generations of technological innovations: the IBM mainframe case. *Technological Forecasting and Social Change* 51:109–132.

Maling, P. 2012. NEW: Smartphone Price Slide. *https://www.strategyanalytics.com*.

Mansfield, E. 1961. Technical Change and the Rate of Imitation. *Econometrica* 29 (4): 741–766.

Meade, N., and T. Islam. 2006. Modelling and forecasting the diffusion of innovation A 25-year review. *International Journal of Forecasting* 22:519–545.

Melser, D., and I. Syed. 2007. Life Cycle Pricing and the Measurement of Inflation.

Miyao, T., and P. Shapiro. 1981. Discrete choice and variable returns to scale. *International Economic*

Review:257–273.

Norton, J., and F. Bass. 1987. A diffusion theory model of adoption and substitution for successive generations of high- technology products. *Management Science* 33 (9): 1069–1086.

Oren, S. S., and R. G. Schwartz. 1988. Diffusion of new products in risk-sensitive markets. *Journal of Forecasting* 7 (4): 273–287.

Peres, R., E. Muller, and V. Mahajan. 2010. Innovation diffusion and new product growth models: A critical review and research directions. *International Journal of Research in Marketing* 29:91–106.

PriceBeam 2017. Life Cycle Pricing: Capturing Value Throughout the PLC. Published Online https://blog.pricebeam.com/life-cycle-pricing-capturing-value-throughout-the-plc.

Rayfield, W., P. Rusmevichientong, and H. Topaloglu. 2015. Approximation methods for pricing problems under the nested logit model with price bounds. *INFORMS Journal on Computing* 27 (2): 335–357.

Robinson, B., and C. Lakhani. 1975. Dynamic pricing models for new product planning. *Management Science* 21 (10): 1113–1122.

Shen, W., I. Duenyas, and R. Kapuscinski. 2014. Optimal pricing, production, and inventory for new product diffusion under supply constraints. *Manufacturing and Service Operations Management* 16 (1): 28–45.

Song, J.-S., and Z. Xue. 2007. Demand management and inventory control for substitutable products. *Working Paper, Duke University*.

Su, X. 2008. Bounded rationality in newsvendor models. *Manufacturing and Service Operations Management* 10 (4): 566–589.

Wang, R., and O. Sahin. 2017. The impact of consumer search cost on assortment planning and pricing. *Management Science*.

Wang, R., and Z. Wang. 2016. Consumer choice models with endogenous network effects. *Management Science* 63 (11): 3944–3960.

Weerahandi, S., and S. R. Dalal. 1992. A choice-based approach to the diffusion of a service: Forecasting fax penetration by market segments. *Marketing Science* 11 (1): 39–53.

Zhang, H., P. Rusmevichientong, and H. Topaloglu. 2018. Pricing under the generalized extreme value models with homogeneous price sensitivity parameters. *Operations Research* forthcoming.

Author Biography

Hongmin Li is an associate professor of Supply Chain Management at the W.P. Carey School of Business, Arizona State University. She received her Ph.D. from Massachusetts Institute of Technology. Her research interests include supply chain management, production and inventory control, pricing, and capacity planning, focusing on multi-generation manufactured products, and companies managing complex and evolving product lines.