3-2018

Multigeneration product diffusion in the presence of strategic consumers

Zhiling GUO
Singapore Management University, ZHILINGGUO@smu.edu.sg

Jianqing CHEN
University of Texas at Dallas

Citation

GUO, Zhiling and CHEN, Jianqing. Multigeneration product diffusion in the presence of strategic consumers. (2018). *Information Systems Research*. 29, (1), 206-224.
Available at: https://ink.library.smu.edu.sg/sis_research/3747

Multi-Generation Product Diffusion in the Presence of Strategic Consumers

Zhiling Guo
Singapore Management University
zhilingguo@smu.edu.sg

Jianqing Chen
The University of Texas at Dallas
chenjq@utdallas.edu

Abstract

Frequent new product releases pose significant challenges for firms as they manage successive generations of product diffusion. We develop an analytical model to study the effect of different purchase options by strategic consumers on a firm's profit and the firm's strategies for the timing and pricing of its successive generations of product diffusion. We show that consumers' strategic behavior, although adversely affecting the sales of the first-generation product, positively influences the sales of the second-generation product through an initial "seeding" effect. The influence of strategic consumers on profit and sales depends largely on the discount-to-price ratio of the first generation relative to the performance improvement in the second generation. When the relative discount is small, the "seeding" effect on the second-generation product dominates. When the relative discount is large, the "cannibalization" effect on the first-generation product dominates. We further demonstrate that the optimal entry timings recommended in the literature (i.e., "now," "maturity," or "never") can occur under different market conditions. In general, higher performance improvement and lower salvage value would support a higher optimal price, a larger discount, and a later introduction time. In addition, the firm can benefit from patient consumers when the performance improvement is relatively small, and it can induce the complete substitution of the later generation for the earlier generation when the performance improvement is relatively large. Overall, our model provides a theoretical foundation for understanding the effect of consumer strategic behavior on product diffusion, and our results offer important insights about firms' multi-generation product diffusion strategies.

1 Introduction

Rapid technological development in the industry has significantly sped up new product development, so that the price of a given model declines over time and several generations of the same product tend to coexist in the consumer marketplace. Anticipating the introduction of a new generation of technology in the near future, potential adopters of the earlier technology might choose to wait, cannibalizing the sales of the old technology—a decision by customers termed "inter-temporal substitution" (Norton and Bass, 1987). For example, Apple sold 14 million iPads in the fourth quarter of 2012, which was significantly less than analysts expected. Apple attributed the lower sales to the fact that customers were holding back and waiting for the newer models (Newton, 2012). A recent Mizuho Securities survey found that consumers have been taking a wait-and-see approach to Apple Watch before jumping in, partly because of the prospect of added innovations in next-gen watches (Seitz, 2015). In September 2015, Apple announced that opening weekend iPhone 6S and 6S Plus sales were "more than 13 million units." The sales looked particularly strong, and the long queue outside Apple stores clearly showed that many consumers had been waiting too long for the new product. These examples show that firms must fully anticipate forward-looking consumers' reactions to a future newer generation and take into account the effect of the consumers' strategic behavior on product sales of both generations.

The introduction of a newer generation product generally results in diminishing adoption of the first-generation product. On the one hand, we see cannibalization of sales from the newer generation (i.e., the substitution effect). On the other hand, when the new-generation product is available, some existing adopters upgrade from the first to the second generation (i.e., the switching effect). For example, Apple provides trade-in service for its old-generation iPhone models when the new-generation model is introduced. In addition, the price cut of the old-generation product has a market-expansion effect. Some consumers who have not been able to afford the old-generation product now can purchase the product at a lower price. In the social-technological system, consumers' adoption of multi-generation products follows an interactive diffusion process, where market segmentation is collectively determined by substitution and switching to the newer generation, continuous diffusion of the earlier generation, and both generations' penetration into new markets. How consumers make trade-offs between the performance improvement and the price discount of the two generations requires careful consideration.

Classical multi-generation product diffusion models, such as the Norton and Bass (1987) model, posit that the new-generation product follows an S-shaped growth curve similar to that of the earlier generation product. These models suggest an initial market development period in which the sales rate at the beginning of the new product introduction is low, and the market penetration rate gradually increases over time (see Figure 1). However, recent real market data from the high-tech industry rarely agree with such a growth trend. For example, Apple received more than two million pre-orders in just 24 hours after the firm started its sale of the iPhone 5 (Apple, 2012). The iPhone 5 took just 60 minutes to sell out of its launch-day stock, showing incredible demand from consumers at the time of the new product release. The dramatic jump in sales at the new product introduction can also be seen in the iPhone quarterly sales figure (see Figure 1). The jumpstart on the sales cannot be explained by existing multi-generation diffusion models, which typically assume a slow rate of market penetration in the initial stage of the product life cycle. The tendency to ignore individual consumers' adoption behavior is a key limitation in the aggregate-level, classical Bass-type of diffusion models (Chatterjee and Eliashberg, 1990).

In this paper, we develop an analytical model that reasonably explains the jump as a market seeding effect of consumers' strategic waiting. Because forward-looking, tech-savvy consumers are capable of anticipating the new product introduction, the way they trade off among various purchase options when deciding which product to buy and when to buy it affects the demand dynamics in multi-generation product diffusion. Having diffusion models that take into account consumer heterogeneity and consumers' strategic decision making,

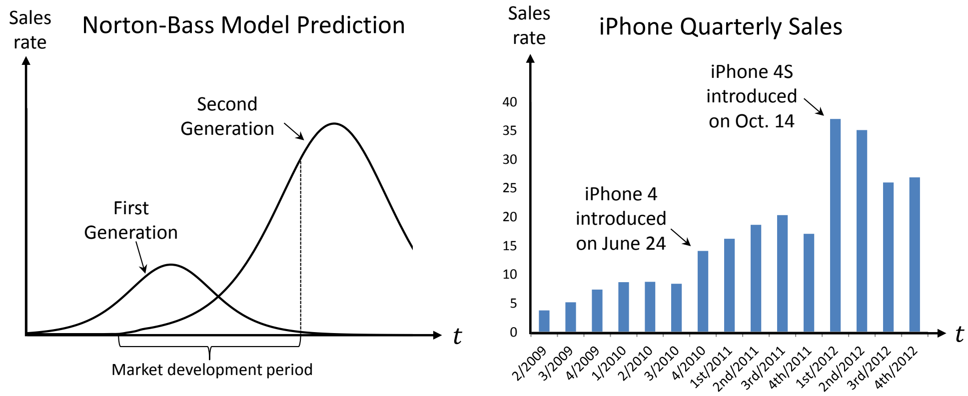


Figure 1: Comparison between the Norton-Bass Model and the iPhone Sales Data

and that are based on how market segmentation can influence the diffusion process, is highly desirable. We aim to fill this gap in the literature.

Building on the seminal Bass product diffusion framework (Bass, 1969), we propose a model that takes into account heterogeneous consumers' "buy now or later" strategic purchase decisions, which collectively influence product diffusion dynamics across successive generations of products. Using a direct microeconomic approach, our model provides a decision-theoretic foundation that explains individual adoption decisions and corresponding market segmentation. In contrast to most existing literature, which treats consumer demand and the substitution between the two generations as exogenous (Norton and Bass, 1987; Wilson and Norton, 1989; Mahajan and Muller, 1996), we allow for demand dependencies by endogenizing the substitution and switching effects based on strategic consumers' preference for different purchasing options. By fully anticipating the response of strategic consumers, we examine how the firm's pricing policy and introduction timing policy affect consumers' strategic behavior, which in turn affects the sales trajectory and the firm's profitability.

We find that consumers' strategic behavior, although adversely affecting the sales of the first-generation product, positively influences the sales of the second-generation product through an initial "seeding" effect. The influence of strategic consumers on profit and sales depends largely on the discount-to-price ratio of the first generation relative to the performance improvement in the second generation. When the relative discount is small, the "seeding" effect on the second-generation product dominates. When the relative discount is large, the "cannibalization" effect on the first-generation product dominates. We further demonstrate that the various optimal entry timings recommended in the literature ("now," "maturity," or "never") can occur under different market conditions. We also observe more extreme introduction timing (either at the beginning or at the end of the first-generation product life cycle) when the firm ignores consumers' strategic behavior and when product development and production costs are considered (see Online Appendix). In general, higher performance improvement and lower salvage value would support a higher optimal price, a larger discount, and a later introduction time. In addition, the firm can benefit from patient consumers when the performance improvement is relatively small, and it can induce the complete substitution of the later generation for the earlier generation when the performance improvement is relatively large. Overall, our results offer important insights regarding the firm's optimal timing and pricing policies in the presence of strategic consumers.

The organization of the paper is as follows. In the next section we review the relevant literature. Section 3 describes our base model and strategic consumers' purchase options. Section 4 derives the market segmentation and presents the resulting product diffusion dynamics of the two generations. Section 5 provides some important managerial insights based on numerical optimization. Section 6 concludes the paper. All proofs are presented in the Appendix. The Online Appendix further provides several model extensions to examine the effects of a more general pricing rule, an uncertain release time for the second-generation product, and the incorporation of product development and production costs on the consumers' behavior and the firm's strategies.

2 Literature Review

In this section, we review the technological diffusion models and discuss some recent extensions. We focus first on theoretical model development and then on empirical studies.

The earliest and most influential diffusion model is proposed by Bass (1969). Consistent with the studies of the adoption and diffusion of innovations in the social science literature (Roger, 1983), the Bass model assumes that the adoption process is affected by two sources of influence in the social system: the external source of information (e.g., mass advertising) and the internal source (e.g., word-of-mouth communication). The model describes how adoption probabilities and rates change over time as new products penetrate a fixed population, based on a hazard rate function (i.e., the conditional probability that an adoption will occur at time t given that an adoption has not yet occurred). Sales growth predicted by the Bass model follows a logistic curve (i.e., the S-shaped growth pattern), which gained substantial empirical support from a variety of durable goods.

Subsequent extensions have built into the model greater realism regarding consumer adoption behavior. For example, Kalish (1985) characterized the adoption of a new product as consisting of two steps: awareness and adoption. Awareness is generated by advertising and word of mouth and is the stage of being informed about the product. Conditional on awareness, adoption occurs if the perceived value of the product exceeds its price in a heterogeneous population. Much of the marketing literature has incorporated other marketing mix variables, such as price and advertising, into the model (Robinson and Lakhani, 1975; Dolan and Jeuland, 1981; Kalish, 1983; Feichtinger, 1982). For a comprehensive review of diffusion models, we refer readers to Mahajan et al. (1990) and Peres et al. (2010). In addition to a demand-side, Bass-type of diffusion model, a few studies in the field of operations management add the supply-side constraints (Jain et al. 1991;Kurawarwala and Matsuo, 1996;Ho et al., 2002). For example, Balakrishnan and Pathak (2014) considered the influence of production capacity on service quality in the presence of a supply shortage. In the context of IT services, Niculescu et al. (2012) assessed how prices, network effects, consumer heterogeneity, and associated awareness jointly govern the adoption paths. In contrast to our research, all these papers consider the diffusion of a single product.

As the product life cycle becomes shorter, the simultaneous coexistence of multiple product lines or of several generations of a single product category is a commonly observed phenomenon in the high-tech industry. Norton and Bass (1987) proposed a multi-generation diffusion model that focuses on the technological substitution and the diffusion patterns of multiple generations simultaneously. Jiang and Jain (2012) provided a generalization of the Norton and Bass (1987) model by separating the switching consumers from the leapfrogging consumers in the substitution process. Alternatively, Mahajan and Muller (1996) proposed a model that allows for partial leapfrogging and partial cannibalization between the two generations. It suggests that a firm should either introduce a new generation as soon as it is available or delay its introduction to a much later date, termed as a “now or maturity” strategy. Building on Kalish (1985), which considers the two-step (awareness and adoption) product diffusion in one product generation, Wilson and Norton (1989) extended the model to a two-generation product setting, resulting in a “now or never” timing strategy. With the focus on dynamic pricing, Padmanabhan and Bass (1993) suggested that the optimal pricing strategy depends on the degree of substitutability across the two generations. Furthermore, Krankel et al. (2006) and Ke et al. (2013) showed how operational decisions, such as order policy and inventory cost, can affect the optimal introduction timing decisions. Mehra et al. (2014) studied successive software upgrade strategies in the presence of technological obsolescence. They showed that the optimal upgrade intervals are monotonically increasing throughout the product’s life cycle because of demand and cost considerations. We complement these prior works by considering individual consumers’ strategic behavior. We demonstrate that both the “now or never” (Wilson and Norton, 1989) and the “now or maturity” (Mahajan and Muller, 1996) introduction timing rules can arise as the optimal strategy. We show that a wide range of timing choices could be optimal under different market conditions.

In addition to the continuous time models in the product diffusion literature, a few studies have used two-period discrete-time models to analyze the sequential and simultaneous introduction strategies of high- or low-end product line extensions. Moorthy and Png (1992) suggested that, if a firm can commit to the subsequent prices and product design, the introduction of a low-end product should be delayed to alleviate cannibalization of sales of the high-end product. However, Bhattacharya et al. (2003) showed that the introduction of a low-end product before its high-end variant might be optimal if technological improvement is taken into account.

In terms of strategic consumers’ decision making, Bala and Carr (2009) considered a two-period model in which consumers anticipate product prices and qualities while the firm decides upgrade pricing. They showed that both product improvement and user costs play a role in pricing software upgrades. Some recent studies in the operations management literature also have incorporated intertemporal consumer purchasing decisions in anticipation of a future price markdown (Aviv and Pazgal, 2008; Liu and van Ryzin, 2008; Su and Zhang, 2008). However, all of these works focus on inventory management rather than on multi-generation product sales.

One key assumption of the Bass model is that the potential adopter population is homogenous, which implies that, at any point in the process, all individuals who have not yet adopted a product have the same probability of adopting. To overcome this limitation, several studies consider the heterogeneous consumer's choice processes in deriving the product diffusion pattern. In the presence of competition between an existing mature product and an uncertain new product, Oren and Schwartz (1988) developed a model in which risk-averse consumers are Bayesian learners who use information generated by early adopters to update their prior knowledge about the new product's performance. Hiebert (1974) and Jensen (1982) studied the effects of risk attitude and of learning, given uncertain perceptions of an innovation, on the individual level adoption decision in a heterogeneous population. Roberts and Urban (1988) used a dynamic brand choice model to study individual consumers' acceptance of a new brand when they are uncertain about the brand value. Chatterjee and Eliashberg (1990) incorporated heterogeneity in the population with respect to initial uncertain perceptions about product quality, risk attitude, price sensitivity, and responsiveness to information about the innovation. Most of these models include Bayesian updating of uncertain perceptions. Heterogeneity in these models is captured by different initial perceptions. In contrast to this stream of literature, we assume heterogeneous consumer valuation of the product and consumers' strategic behavior in the presence of several inter-temporal purchase options.

In addition to the theoretical analysis already described, a substantial amount of empirical research also has applied the multi-generation diffusion model to forecast growth in the high-tech industry. For example, Danaher et al. (2001) studied both first-time sales and ongoing renewals of a subscription for two generations of analog cellular phone technology in a European country. They found that intergenerational interdependencies cannot be ignored when extending single-generation estimates of price response over time to successive-generation markets. Chu and Pan (2008) estimated the growth potential of the mobile Internet market in Taiwan. Islam and Meade (1997) fit the multi-generation model using mobile phone technology data from eleven countries. Jun and Park (1999) and Kim et al. (2005) adopted a consumer choice model to study the diffusion and substitution processes of successive generations of the IBM mainframe system and the worldwide dynamic random access memory (DRAM) market. Based on market survey data, Kim et al. (2001) proposed an individual-level adoption model to incorporate both initial and repeat consumer purchases in the multi-generation personal computer market.

Despite the effort of incorporating micro-level decision making, Bass-type model variations cannot sufficiently explain some actual market data. Song and Chintagunta (2003) provided a structural model to explain the long "time to take-off" phenomenon observed by Golder and Tellis (1997) and the saddle effect in the sales pattern observed by Goldenberg et al. (2002) in a single-generation product setting. They showed that forward-looking consumers might strategically hold out on their purchase of a durable product, anticipating a future price markdown. We complement this existing literature by considering strategic consumers' behavior in a multi-generation product diffusion framework. The micro-modeling approach provides a behavioral basis for explaining adoption at the individual level, which leads to various patterns of diffusion at the aggregate market level. Our new model lays a promising foundation for future empirical analysis.

3 Model Description

In this section, we focus on the key idea of our approach, which extends the classic, aggregate Bass diffusion model (Bass, 1969) and its multi-generation version (Norton and Bass, 1987; Mahajan and Muller, 1996) to a micro-level diffusion process that takes into account various types of behaviors of strategic consumers.

We consider a durable technological innovation, such as a high-tech product like an iPhone. Let m be the population size and $x(t)$ be the number of adopters by time t. Bass (1969) has suggested that the conditional likelihood of adoption increases linearly in the number of existing adopters—that is, $\alpha + \beta\frac{x(t)}{m}$, where α and β are parameters called the coefficient of innovation and the coefficient of imitation, respectively. The underlying rationale is that α captures external influences, such as the mass advertising effect, and β captures the internal influences, such as the word-of-mouth effect, which depends linearly on the market penetration in the product diffusion process. The adoption rate can be expressed as a continuous time differential equation, $\dot{x}(t) = \frac{dx(t)}{dt} = (\alpha + \beta\frac{x(t)}{m})(m - x(t))$, which is the probability of adoption multiplied by the number of potential adopters who have not yet adopted the product. Throughout this paper, we use the dot notation to denote the first derivative.

Taking into account the price effect, the deterministic market potential m can be expressed as a function of price. Kalish (1983) and Feichtinger (1982) have assumed $\dot{x}(t) = (\alpha + \beta\frac{x(t)}{m})(m(p) - x(t))$, where $m(p)$, the market size, is limited by the number of consumers who are willing to purchase the product at price p, and $m(p) - x(t)$ is interpreted as the remaining market potential. We extend this line of literature by considering strategic consumers. In the presence of strategic consumers, the conditional probability of adoption is less than $\alpha + \beta\frac{x(t)}{m}$ because some consumers prefer not to buy immediately after evaluating their purchase options. Some consumers cannot afford the product, and others choose to strategically delay their purchase until the second-generation product release. Therefore, only a segment of the penetrated market is converted to final adoption based on the available options. We express the fraction of conversion as $f(\cdot)$ and call it the conversion rate. The explicit form of $f(\cdot)$ is derived in Section 4.2. Accordingly, our model modifies the conditional probability of purchase by multiplying the awareness probability by the conversion rate. The rate of adoption is expressed as $\dot{x}(t) = f(\cdot)(\alpha + \beta\frac{x(t)}{m})(m(p) - x(t))$. Note that the rate parameter $f(\cdot)$ was exogenously given in Kalish (1985) (parameter k in his model), whereas in our model, $f(\cdot)$ is endogenously derived, based on the market segmentation of aware consumers.

In the context of two generations of technological innovation, we assume the first-generation product is available at $t = 0$. The introduction time of the second-generation product, $\tau > 0$, is common knowledge. The firm pre-announces its two-generation markdown-pricing policy (p, δ) over $[0, T]$ as follows. The first-generation product is priced at p. The second generation sells at the same price as the first-generation product when it is released. At the same time, the first generation's selling price is discounted to $p - \delta$. This pricing strategy is frequently observed in the high-tech product market, as with Apple's iPhone product family. For example, in late 2013, the newest 16G iPhone5S was sold at \$199—the same price as the iPhone5 before the iPhone5S was introduced. The 16G iPhone5C, which used the earlier iPhone5 generation technology, was sold at a discounted price of \$99 at that time. Use of price commitments in connection with strategic consumers has been widely discussed in the literature. The benefit of a price commitment is that it encourages consumers not to gamble on the future price movement, so they can make a purchase decision right away. Stokey (1981) has suggested that a durable goods monopolist cannot charge a price above its cost if the monopolist is unable to commit to future prices, in part because strategic consumers delay their purchase in anticipation of the future price cut. Besanko and Winston (1990) showed that firms can be better off when they commit to a declining price path.

The population is heterogeneous with respect to consumers' valuation of the product. We assume that each consumer derives an instantaneous utility u per unit time for using the first-generation product and a higher instantaneous utility ρu, where $\rho > 1$ represents the performance improvement, for the second-generation product because of enhanced functionality, better quality, and so on. Because we focus on durable products, all consumers consider the time period of the use of the product to be infinite, with a discount rate r. If a consumer buys the first-generation product and uses it forever, the total utility generated from using the product is $u \int_0^\infty e^{-rt} dt = \frac{u}{r}$. We define the first-generation product lifetime value as $v \equiv \frac{u}{r}$; similarly, the second-generation product lifetime value is ρv. For simplicity, we assume v follows a uniform distribution over $[0, 1]$.

The new generation of technology provides an opportunity for technological substitution of the first-generation product, which results in several types of strategic behavior. If a strategic consumer prefers to wait for the second-generation product, she might forgo the opportunity to buy the first-generation product, even though she would have adopted the first generation if the second generation were not available. We call this option the *Leapfrog* option, which captures the intergenerational substitution effect. If a strategic consumer is only interested in buying the first-generation product at a discounted price after the second-generation product is introduced, then she is a laggard adopter of the first-generation product. We call this option the *Laggard* option.

If a strategic consumer prefers to buy the first-generation product immediately, two cases emerge: Some consumers might decide to buy the first-generation product as a "once and for all" decision. They enjoy the use of the first-generation product and do not consider adopting the second-generation product even after it becomes available. In contrast, other first-generation adopters might prefer to switch to the second-generation product when it is launched. We call the former the *Adopt* option and the latter the *Upgrade* option. The upgrade option is a widely observed industry practice—for example, note the Apple's iPhone upgrade program in the United States. We call the consumers who decide not to buy either generation of product the *Non-Adopters*. As a result, strategic consumers who are aware of the firm's product at time t essentially have the following five options:

Adopt: Buy the first-generation product immediately at price p and continue to use the product for the product's lifetime; the expected total payoff is $v - p$.

Leapfrog: Wait and buy the second-generation product with expected payoff $e^{-r(\tau-t)}(\rho v - p)$.

Laggard: Wait until the release of the second-generation product and buy the first-generation product at a discounted price; the expected payoff is $e^{-r(\tau-t)}(v - p + \delta)$.

Upgrade: Buy the first-generation product immediately at price p and upgrade to the second generation at time τ; the expected payoff is $v - p + e^{-r(\tau-t)}[(\rho - 1)v - p + \mu]$, where μ is the salvage value of the first-generation product. We assume that the salvage value of a used first-generation product is lower than the selling price of the new first-generation product at time τ; that is, $p - \delta > \mu$, which implies no arbitrage.

Non-Adopt: Do not buy any of the products, with payoff 0.

Note that the Upgrade consumers are the existing adopters of the first-generation product who upgrade to the new generation, which captures the switching behavior among existing consumers. In contrast, the Leapfrog consumers are the potential adopters who skip the first-generation product to adopt the second generation, which captures the substitution behavior between the two generations. Although previous research recognizes the importance of distinguishing the two types of behavior (Norton and Bass, 1987; Wilson and Norton, 1989; Mahajan and Muller, 1996), none has provided a sound behavioral basis to explain the effects. In the following section, we endogenously determine the market segmentation based on consumers' evaluation of the different purchase options.

4 Analysis

In this section, we first analyze strategic consumers' purchase options and the resulting market segmentation. We then examine the interdependent diffusion processes and characterize the intergenerational diffusion dynamics of the two generations of products.

4.1 Market Segmentation

Regarding strategic consumers' wait-or-buy decision, the Adopt and Upgrade options correspond to an immediate buy decision, and the Leapfrog and Laggard options correspond to a wait decision. The Non-Adopt option is a no-buy decision. Strategic consumers consider these options and choose the one that gives them the highest expected payoff. Under certain circumstances, some options can be dominated by others, and not all strategic actions can be observed. The following Lemma pinpoints the circumstances.

Lemma 1. (Dominated Strategy) *(a) If $\rho - 1 \leq p - \mu$, no strategic consumers choose the Upgrade option.*

(b) If $\rho - 1 \leq \delta$, no strategic consumers choose the Leapfrog option.

(c) If $\delta \leq \frac{p(\rho-1)}{\rho}$, no strategic consumers choose the Laggard option.

Intuitively, only high-valuation consumers might have an incentive to upgrade because of the relatively high increase in the valuation $(\rho - 1)v$ available from the second-generation product. Notice that the effective price that a consumer pays for an upgrade is $(p - \mu)$. When $(\rho - 1)$ is less than the price $(p - \mu)$, even the highest valuation consumer has no incentive to upgrade because the benefit cannot compensate for the cost. Therefore, in this case, no strategic consumers choose the Upgrade option.

Note also that the tradeoff between the Leapfrog and Laggard options is the increase in valuation of the second-generation product and the price discount for the first-generation product. Lemma 1(b) indicates that if the value increase in the second-generation product is small compared to the price discount of the first-generation product, no consumers would be interested in purchasing the second-generation product. In contrast, Lemma 1(c) shows that if the price discount is too small, no consumers would be interested in the older generation product if the newer generation product is available.

In reality, not all consumers who adopt the first-generation product $(v \geq p)$ choose the Upgrade option. In the following, we assume Upgrade occurs among some of the existing adopters; that is, we assume that $p < \frac{p-\mu}{\rho-1} < 1$.[1] Under this assumption, together with the no-arbitrage assumption $p - \delta > \mu$ imposed in the model, we have $\delta < \rho - 1$, which means, by Lemma 1, that the value increase in the second-generation product is not too small and that, as a result, the Leapfrog option is always attractive to some strategic consumers.

We define the indifference curves, $l_{ul}(t)$, $l_{al}(t)$, and $l_{ag}(t)$, where strategic consumers are indifferent between the Upgrade and Leapfrog options, between the Adopt and Leapfrog options, and between the Adopt and Laggard options, respectively. These curves are determined by equalizing payoffs with the respective options: $v - p = e^{-r(\tau-t)}(v - \mu)$, $v - p = e^{-r(\tau-t)}(\rho v - p)$, and $v - p = e^{-r(\tau-t)}(v - p + \delta)$; they can be derived as:

$$\begin{aligned} l_{ul}(t) &= \frac{p-\mu e^{-r(\tau-t)})}{1-e^{-r(\tau-t)}} \\ l_{al}(t) &= \frac{p-pe^{-r(\tau-t)}}{1-\rho e^{-r(\tau-t)}} \\ l_{ag}(t) &= \frac{p-(p-\delta)e^{-r(\tau-t)}}{1-e^{-r(\tau-t)}} \end{aligned} \tag{1}$$

The following proposition presents the strategic consumers' decisions. The market segmen-

[1] If these assumptions cannot be satisfied, then some purchase options might be dominated by others, as shown in Lemma 1. These scenarios can be easily analyzed because they are degenerated cases of the general model presented in Figure 2.

tation is illustrated in Figure 2.

Proposition 1. (Market Segmentation with Markdown Pricing Strategy) *There exist thresholds $t_1 = l_{ag}^{-1}(\frac{\delta}{\rho-1})$, $t_2 = l_{al}^{-1}(\frac{p-\mu}{\rho-1})$, and $t_3 = l_{ul}^{-1}(1)$, where $l_{ul}(t)$, $l_{al}(t)$, and $l_{ag}(t)$ are defined in Equation (1), and $t_1 < t_2 < t_3 < \tau$, such that*

(a) When $t \leq \max\{0, t_1\}$, consumers whose $v \in [p-\delta, l_{ag}(t))$ choose Laggard; those whose $v \in [l_{ag}(t), \frac{p-\mu}{\rho-1})$ choose Adopt; and those whose $v \in [\frac{p-\mu}{\rho-1}, 1]$ choose Upgrade;

(b) When $\max\{0, t_1\} \leq t < \tau$, consumers whose $v \in [\min\{p-\delta, \frac{\delta}{\rho-1}\}, \frac{\delta}{\rho-1}]$ choose Laggard; denote $\underline{v}^* \equiv \max\{\frac{\delta}{\rho-1}, \frac{p}{\rho}\}$:

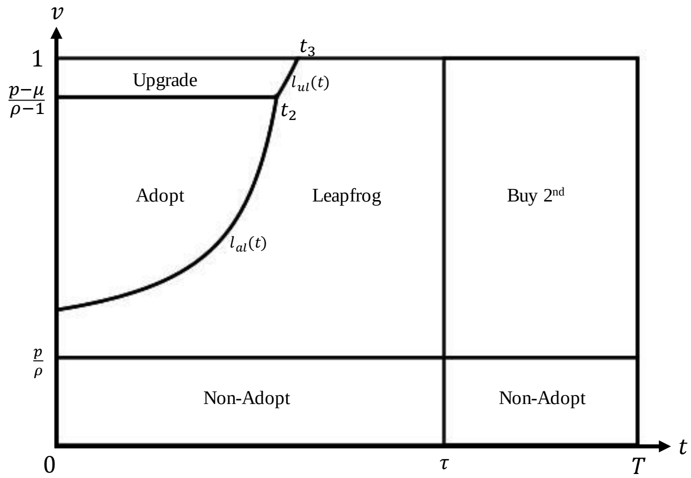
(b.1) When $t < t_2$, consumers whose $v \in [\underline{v}^*, l_{al}(t))$ choose Leapfrog; those whose $v \in [l_{al}(t), \frac{p-\mu}{\rho-1})$ choose Adopt; and those whose $v \in [\frac{p-\mu}{\rho-1}, 1]$ choose Upgrade;

(b.2) When $t_2 \leq t < t_3$, consumers whose $v \in [\underline{v}^*, l_{ul}(t))$ choose Leapfrog, and those whose $v \in [l_{ul}(t), 1]$ choose Upgrade;

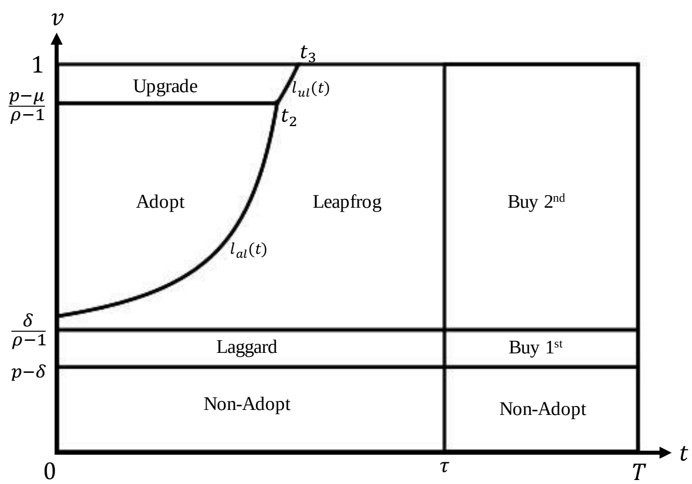
(b.3) When $t \geq t_3$, consumers whose $v \in [\underline{v}^*, 1]$ choose Leapfrog.

Three possible scenarios emerge that depend on the discount-to-price ratio $\frac{\delta}{p}$, as shown in Figure 2, in which the horizontal axis is the time dimension and the vertical axis represents the valuation dimension.

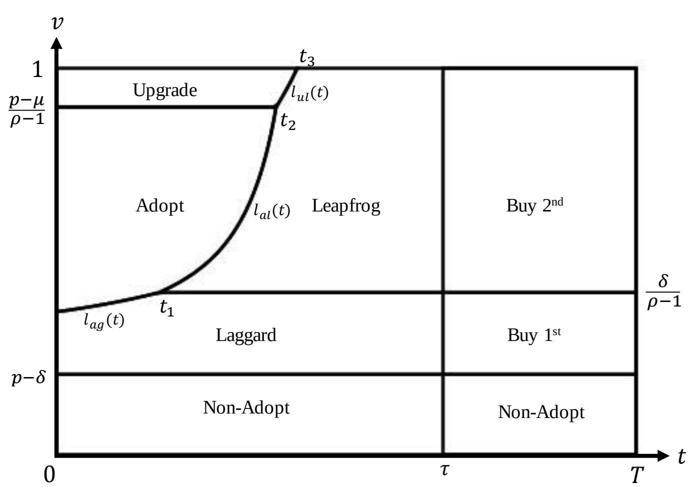
When the discount-to-price ratio is smaller than the relative performance improvement ratio (as in Figure 2a), the discounted first-generation product is not attractive at all, and, by Lemma 1, no strategic consumers choose the Laggard option. All sales after $t > \tau$ go to the second generation. We call this case "Complete Substitution" because no consumers will be interested in the first-generation product after the second-generation product is released. When the discount-to-price ratio is large enough, we have the general case (as in Figures 2b and 2c) where both generations of the product are sold and co-diffuse in the market. The lower the discount-to-price ratio, the larger is the proportion of strategic consumers who prefer the second-generation product. For a given discount-to-price ratio, as t increases, more consumers choose to wait for the second-generation product. In addition, among the early adopters, only high valuation consumers choose the Upgrade strategy. Proposition 1 has important implications for firms as they seek to adopt appropriate marketing strategies.



(a) When $\frac{\delta}{p} \leq \frac{\rho-1}{\rho}$ (Complete Substitution)



(b) When $\frac{\rho-1}{\rho} < \frac{\delta}{p} \leq \frac{(\rho-1)(1-e^{-r\tau})}{1-\rho e^{-r\tau}}$ (General Case)



(c) When $\frac{\delta}{p} > \frac{(\rho-1)(1-e^{-r\tau})}{1-\rho e^{-r\tau}}$ (General Case)

Figure 2: Market Segmentation with Markdown Pricing Strategy

It provides marketers with a better understanding of how the untapped consumers at different valuations evaluate their purchase options and how the composition of the segmented potential consumer pool changes over time. This understanding is important for forecasting the growth of sales.

After the release of the second-generation product, the proportion of buying consumers is $1-\frac{p}{\rho}$ or $1-p+\delta$ (see Figure 2). We see that both performance improvement and price discount can be effective means to attract purchasers. This insight is important for firms crafting their market expansion strategy. In the presence of strategic consumers, if the discount-to-price ratio is large enough, performance improvement changes the relative proportion of consumers who purchase the second-generation product among the buying consumers, but it does not contribute to the total market expansion. In contrast, if the discount-to-price ratio is relatively small, then performance improvement is crucial for expansion beyond the existing market.

Because the three indifference curves $l_{ul}(t)$, $l_{al}(t)$, and $l_{ag}(t)$ define the boundaries between consumers who adopt the first-generation product immediately and those who do not, the following result shows how the proportion of consumers who are interested in the immediate purchase of the first-generation product changes when key market parameters change.

Corollary 1. (Wait-or-Buy) *(a) The number of consumers who choose to adopt the first-generation product increases in the released time τ; that is, $\frac{\partial l_{ul}(t)}{\partial \tau} < 0$, $\frac{\partial l_{al}(t)}{\partial \tau} < 0$, and $\frac{\partial l_{ag}(t)}{\partial \tau} < 0$;*

(b) The number of consumers who choose to adopt the first-generation product decreases in the price p; that is, $\frac{\partial l_{ul}(t)}{\partial p} > 0$, $\frac{\partial l_{al}(t)}{\partial p} > 0$, and $\frac{\partial l_{ag}(t)}{\partial p} > 0$;

(c) The number of consumers who choose to adopt the first-generation product (weakly) decreases in the price discount δ; that is, $\frac{\partial l_{ul}(t)}{\partial \delta} = 0$, $\frac{\partial l_{al}(t)}{\partial \delta} = 0$, and $\frac{\partial l_{ag}(t)}{\partial \delta} > 0$;

(d) The number of consumers who choose to adopt the first-generation product (weakly) decreases in the performance improvement in the second-generation product; that is, $\frac{\partial l_{ul}(t)}{\partial \rho} = 0$, $\frac{\partial l_{al}(t)}{\partial \rho} > 0$, and $\frac{\partial l_{ag}(t)}{\partial \rho} = 0$;

(e) The number of consumers who choose to adopt the first-generation product (weakly) increases in the salvage value μ; that is, $\frac{\partial l_{ul}(t)}{\partial \mu} < 0$, $\frac{\partial l_{al}(t)}{\partial \mu} = 0$, and $\frac{\partial l_{ag}(t)}{\partial \mu} = 0$.

Other things being equal, Corollary 1(a) implies that, as the release time τ increases, the indifference curves shift to the right and the areas of Upgrade and Adopt in the graphs enlarge. Intuitively, consumers find that waiting would forgo too much utility when the release time of the second generation is far away.

In contrast, Corollary 1(b) suggests that, as the price increases, the indifference curves shift to the left. In other words, compared to a lower price, a higher price makes the net utility derived from using the product immediately smaller. The lower opportunity cost from delaying induces more consumers to defer their purchase to a later time.

Corollary 1(c) implies that the price discount of the first generation only negatively affects consumer choice between the Adopt and Laggard options. Intuitively, a higher discount makes more consumers defer their purchase of the first generation.

Corollary 1(d) indicates that the performance improvement in the second-generation product has a (weakly) negative effect on adoption because it induces strategic consumers to wait for the better product. The performance of the second-generation product does not affect consumer choice between the Upgrade and Leapfrog options and between the Adopt and Laggard options because the focal product being considered is the same.

Corollary 1(e) also shows that salvage value does not directly affect the first-generation product sales at the beginning, but a higher salvage value increases the proportion of consumers who upgrade. Therefore, it has a negative effect on the first-generation installed base after the second generation is introduced.

In sum, we expect to see more strategic consumers adopt the first-generation product immediately when the second generation introduction time is later, when either the price or the price discount is lower, when the performance improvement of the second-generation product is relatively small, and when the salvage value is higher.

4.2 Diffusion Dynamics of Two Generations of Product

In the context of two generations of technological innovation, we use subscript i, where $i = 1, 2$, to denote the ith generation. In the following, we explicitly derive the expression of the conversion rate $f_i(\cdot)$ and the market dynamics.

4.2.1 Diffusion when $t < \tau$

We denote the maximum market potential for the first-generation product as m_1. We interpret m_1 as the population that eventually would have adopted the first generation had the second generation not been introduced and had the first generation been priced at 0. Consumers who are aware of the product and who find the price below their reservation utility are the potential adopters. However, the actual adoption timing depends on the evaluation of available purchase options. The diffusion process is as follows.

At any time, a certain portion of market penetration is achieved, and some consumers become aware of the firm's product. These consumers make the purchase decision accordingly, and they exit the potential market after doing so. Those who decide to buy immediately become adopters of the product, and those who decide to delay their purchase and those who decide not to buy are also excluded from the remaining potential market in the current diffusion process. So the remaining potential market consists only of consumers who are not yet aware of the product. A portion of this remaining market is reached in the next time period, resulting in awareness of the product by some consumers and in their decision making. Market penetration continues to occur until the release of the second-generation product.

We assume any consumer is a strategic consumer with probability λ and non-strategic (myopic) with probability $1-\lambda$. A myopic consumer makes immediate purchase decisions as long as her valuation of the product is greater than the product price. In contrast, strategic consumers, who anticipate a future new product release and a price discount of the current product, time their purchase to maximize their expected payoffs based on Proposition 1.

As shown in Figure 2, different market segmentation yields different rates of diffusion at different times, leading to different market penetration patterns and diffusion dynamics. Because of space limitations, we focus here on the most complicated scenario, presented in Figure 2(c). Market dynamics under other scenarios can be derived in a similar way.

When $t \in (0, t_1]$, all Upgrade and Adopt strategic consumers buy the first-generation product. All myopic consumers whose product lifetime value is above the product price ($v \geq p$) buy the first-generation product without considering the upgrade opportunity in the future. Therefore, the Upgrade conversion rate is $\lambda(1-\frac{p-\mu}{\rho-1})$, the Adopt conversion rate is $\lambda(\frac{p-\mu}{\rho-1}-l_{ag}(t))+(1-\lambda)(1-p)$, the Leapfrog conversion rate is 0, the Laggard conversion rate is $\lambda(l_{ag}(t)-p+\delta)$, and the Non-Adopt conversion rate is $\lambda(p-\delta)+(1-\lambda)p$. The conversion rate calculation can be done similarly for $t \in (t_1, t_2]$, $t \in (t_2, t_3]$, and $t \in (t_3, \tau]$, respectively.

We denote $u_1(t)$, $a_1(t)$, $l_1(t)$, $g_1(t)$, and $n_1(t)$ as the cumulative number of consumers who have chosen the Upgrade, Adopt, Leapfrog, Laggard, and Non-Adopt options, respectively. We can formulate the respective conversion rates as follows:

$$f\left(u_1(t)\right) = \begin{cases} \lambda(1-\frac{p-\mu}{\rho-1}) & \text{if } t \leq t_2 \\ \lambda(1-l_{ul}(t)) & \text{if } t_2 < t \leq t_3 \\ 0 & \text{if } t_2 < t < \tau \end{cases} \tag{2}$$

$$f\left(a_1(t)\right) = \begin{cases} \lambda(\frac{p-\mu}{\rho-1}-l_{ag}(t))+(1-\lambda)(1-p) & \text{if } t \leq t_1 \\ \lambda(\frac{p-\mu}{\rho-1}-l_{al}(t))+(1-\lambda)(1-p) & \text{if } t_1 < t \leq t_2 \\ (1-\lambda)(1-p) & \text{if } t_2 < t \leq \tau \end{cases} \tag{3}$$

$$f\left(l_1(t)\right) = \begin{cases} 0 & \text{if } t \leq t_1 \\ \lambda(l_{al}(t)-\frac{\delta}{\rho-1}) & \text{if } t_1 < t \leq t_2 \\ \lambda(l_{ul}(t)-\frac{\delta}{\rho-1}) & \text{if } t_2 < t \leq t_3 \\ \lambda(1-\frac{\delta}{\rho-1}) & \text{if } t_3 < t < \tau \end{cases} \tag{4}$$

$$f\left(g_1(t)\right) = \begin{cases} \lambda(l_{ag}(t)-p+\delta) & \text{if } t \leq t_1 \\ \lambda(\frac{\delta}{\rho-1}-p+\delta) & \text{if } t_1 < t \leq \tau \end{cases} \tag{5}$$

and

$$f\left(n_1(t)\right) = \lambda(p-\delta)+(1-\lambda)p. \tag{6}$$

Assuming that the initial adoption of the first-generation product is 0, we can express the diffusion dynamics for the first-generation product as:

$$\dot{z}_1(t) = f\left[z_1(t)\right](\alpha_1+\beta_1\frac{x_1(t)}{m_1})[m_1-l_1(t)-g_1(t)-n_1(t)-x_1(t)], \tag{7}$$

where $z_1(t) = u_1(t)$, $a_1(t)$, $l_1(t)$, $g_1(t)$, and $n_1(t)$, and $x_1(t) = a_1(t) + u_1(t)$ is the cumulative number of adopters at time t. The term in the square bracket is the remaining market potential at time t; that is, delayed adopters (i.e., $l_1(t)$ and $g_1(t)$) and those who decide not to buy (i.e., $n_1(t)$) are excluded from the price-dependent potential market. Together with the expressions in (2), (3), (4), (5), and (6), the system of differential equations fully specifies the system dynamics when $t < \tau$. Note that a key difference between our specification and the previous multi-generation diffusion models is that our model endogenizes the market segmentation by consumers' strategic choices, whereas the previous models' flow dynamics across the various generations follow exogenously specified fractions for growth and substitution.

4.2.2 Diffusion when $t \geq \tau$

Because technological improvements help firms expand into new markets and lead to additional unique demand for the new-generation product, this new-generation product presumably creates its own market appeal that could not have been achieved by the first-generation technology. Following Norton and Bass (1987), we denote the new market demand as m_2, which is the incremental market for the newer-generation product.[2]

Following the technological innovation literature, the word-of-mouth influence is affected by the installed base. The installed base is defined as the number of products in use for each generation. In the one-generation case, the cumulative sales is equal to the installed base. In the multi-generation setting, the installed base might decrease as a result of replacement by the newer-generation product, so the base might be less than the cumulative sales. At $t = \tau$, we expect to observe a discontinuous jump for both generations because of the waiting consumers—the Upgrade consumers will switch from the first-generation product to the second generation, the Leapfrog consumers will buy the second-generation product, and the Laggard consumers will buy the discounted first-generation product immediately. Denote $u_1(\tau)$, $l_1(\tau)$, and $g_1(\tau)$ as the total number of Upgrade, Leapfrog, and Laggard consumers

[2] Other assumptions regarding the new market demand are also plausible. For example, Mahajan and Muller (1996) represent m_2 as the incremental market expansion after the newer-generation product becomes available, so that both generations of products continue to diffuse and compete in the expanded market.

at time τ. We have initial value $x_{1\tau} = x_1(\tau) + g_1(\tau) - u_1(\tau)$ and $x_{2\tau} = l_1(\tau) + u_1(\tau)$ for the two generations of products.

In the spirit of Norton and Bass (1987), we denote $y_2(t)$ and $n_2(t)$ as the cumulative numbers of Adopters and Non-Adopters from the unique market of the second-generation product, respectively; $\dot{y}_2(t)$ and $\dot{n}_2(t)$ are the corresponding adoption rates. Their initial values are $y_{2\tau} = 0$ and $n_{2\tau} = 0$. When $t > \tau$, the instantaneous adoption rates follow the following system of diffusion dynamics:

$$\begin{cases} \dot{x}_1(t) = (\frac{\delta}{\rho-1} - p + \delta)(\alpha_1 + \beta_1 \frac{x_1(t)+x_2(t)+y_2(t)}{m_1+m_2})(m_1 - n_1(t) - x_1(t) - x_2(t)) & (a) \\ \dot{x}_2(t) = (1 - \frac{\delta}{\rho-1})(\alpha_1 + \beta_1 \frac{x_1(t)+x_2(t)+y_2(t)}{m_1+m_2})(m_1 - n_1(t) - x_1(t) - x_2(t)) & (b) \\ \dot{n}_1(t) = (p - \delta)(\alpha_1 + \beta_1 \frac{x_1(t)+x_2(t)+y_2(t)}{m_1+m_2})(m_1 - n_1(t) - x_1(t) - x_2(t)) & (c) \\ \dot{y}_2(t) = (1 - \frac{p}{\rho})(\alpha_2 + \beta_2 \frac{x_2(t)+y_2(t)}{m_1+m_2})(m_2 - n_2(t) - y_2(t)) & (d) \\ \dot{n}_2(t) = \frac{p}{\rho}(\alpha_2 + \beta_2 \frac{x_2(t)+y_2(t)}{m_1+m_2})(m_2 - n_2(t) - y_2(t)) & (e) \end{cases} \tag{8}$$

Note that the two products, the first generation and second generation, compete for market share because of the substitution effect. (See the diffusion dynamics in 8(a), (b), and (c).) Because the total number of adopters for both products is $x_1(t) + x_2(t) + y_2(t)$, the fraction of adoption in the population for both generations is expressed as $\frac{x_1(t)+x_2(t)+y_2(t)}{m_1+m_2}$. Moreover, the second-generation product has its own unique market potential. The instantaneous adoption rate from this unique market penetration is expressed in 8(d) and (e). Because both $x_2(t)$ and $y_2(t)$ are related to the adoption of the second-generation product among the total population $m_1 + m_2$, the fraction of adopters is $\frac{x_2(t)+y_2(t)}{m_1+m_2}$.[3]

The firm's total discounted profit, denoted as the net present value of sales for the two generations, is expressed as:

[3]Following Mahajan and Muller (1996), an alternative specification is:

$$\begin{cases} \dot{x}_1(t) = (\frac{\delta}{\rho-1} - p + \delta)(\alpha_2 + \frac{\beta_1 x_1(t)+\beta_2 x_2(t)}{m_1+m_2})(m_1 + m_2 - n_2(t) - x_1(t) - x_2(t)) & (a) \\ \dot{x}_2(t) = (1 - \frac{\delta}{\rho-1})(\alpha_2 + \frac{\beta_1 x_1(t)+\beta_2 x_2(t)}{m_1+m_2})(m_1 + m_2 - n_2(t) - x_1(t) - x_2(t)) & (b) \\ \dot{n}_2(t) = (p - \delta)(\alpha_2 + \frac{\beta_1 x_1(t)+\beta_2 x_2(t)}{m_1+m_2})(m_1 + m_2 - n_2(t) - x_1(t) - x_2(t)) & (c) \end{cases}$$

The difference between this alternative specification and Model (8) is whether the expanded market is unique to the second-generation product or is shared between the two generations. The coefficients α_i and β_i, where $i = 1, 2$, could be the same or different, depending on the nature of the technological innovation. Our numerical simulation shows that our major insights are robust across the different model specifications.

$$\begin{aligned}\pi = \textstyle\int_0^\tau e^{-rt} p\dot{x}_1(t)dt + e^{-r\tau}\,(p-\delta)\,g_1(\tau) + e^{-r\tau} p[l_1(\tau)+u_1(\tau)] \\ + \textstyle\int_\tau^T e^{-rt}(p-\delta)\dot{x}_1(t)dt + \int_\tau^T e^{-rt} p[\dot{x}_2(t)+\dot{y}_2(t)]dt]\end{aligned} \tag{9}$$

The first term in the profit function is the profit from sales of the first-generation product before the second generation is released. The next three terms are the profit from selling the discounted first-generation product to Laggards, and the profit from selling the second-generation product to Leapfroggers and to Upgraders at time τ, respectively.[4] The fourth term is the profit from continued sales of the first-generation product after the second-generation product is released until the end of the planning horizon. The fifth and final term is the profit from sales of the second-generation product in both the competing market and the product's own unique market.

The firm's total discounted profit depends on a few key parameters—especially the pricing- and timing-related decisions (p, δ, τ). Because of the complicated sales dynamics involved, the effects of price and release time τ are non-linear. The price variables (p, δ) are often strategic decisions to which the firm commits in advance, and they can be treated as fixed during the short product life cycle. The introduction timing variable (τ) is a tactical decision that the firm can control. We next examine how the introduction timing and pricing strategies affect the firm's total discount profits, and whether the presence of strategic consumers might alter the the conventional "now-or-never" (Wilson and Norton, 1989) or "now or maturity" introduction timing rules (Mahajan and Muller, 1996).

5 Numerical Optimization and Managerial Insights

In this section, we take an incremental approach to understanding the effect of strategic consumers' behavior on the firm's sales and profit. First, under a given firm's pricing and timing strategies, we compare the diffusion dynamics with and without consideration of strategic consumers. Next, for a given price and discount schedule, we examine a firm's optimal timing strategy. Finally, we determine the optimal entry timing and pricing strategy

[4] Note that although the selling price to Upgraders is $(p-\mu)$, we assume the first-generation product has salvage value μ to the firm. So the unit net profit from the upgrade consumers is still p.

simultaneously through numerical optimization.[5]

5.1 Impact of Strategic Consumer Behavior on Diffusion Dynamics

Given the firm's decision about price, discount, and release time (p, δ, τ), we first illustrate how product diffusion curves are affected in the presence of strategic consumers. Although the shapes of the sales diffusion curves are highly dependent on the parameter values, our main purpose is to show how the curves shift as the proportion of strategic consumers in the population increases.

The parameter values in Figure 3 are as follows. The pricing policy used is $p = 0.25$ and $\delta = 0.04$, with $\mu = 0.16$. We assume $T = 6$, and the firm plans to introduce the second generation at $\tau = 3$. The discount rate is $r = 0.1$, and the performance improvement ratio $\rho = 1.1$. Furthermore, $\alpha_1 = \alpha_2 = 0.3$, and $\beta_1 = \beta_2 = 0.6$. We set $m_1 = 100$ and $m_2 = 20$. We take as the benchmark the case in which consumers are non-strategic (i.e., $\lambda = 0$). We compare the benchmark with two scenarios: one in which a consumer is equally likely to be either myopic or strategic (i.e., $\lambda = 0.5$) and one in which all consumers are strategic (i.e., $\lambda = 1$).

The left panel in Figure 3 compares the diffusion of total sales for both generations of product under the three scenarios. The right panel breaks down the total sales into the first-generation sales and second-generation sales, respectively, under the three scenarios. We look first at the dotted curve, which represents the myopic benchmark. Because the case involves no strategic waiting, the cumulative total sales curve is smooth and increasing. The division of sales in the right panel shows that the second generation starts its own diffusion process at $\tau = 3$ from 0. The sales growth of the first-generation product is significantly slowed after this point because of the cannibalization of sales from the second-generation product. The substitution of sales from the first-generation product and the unique market expansion of the second-generation product contribute to the relatively large sales of the

[5] In the multi-generation product diffusion literature, analytical tractability is a known challenge. As is widely recognized in the literature, the analysis of a multi-generation product diffusion process using a micromodeling approach in general has no closed-form results because of its inherent complexity (Chatterjee and Eliashberg, 1990). We therefore resort to numerical optimization to derive the optimal solutions.

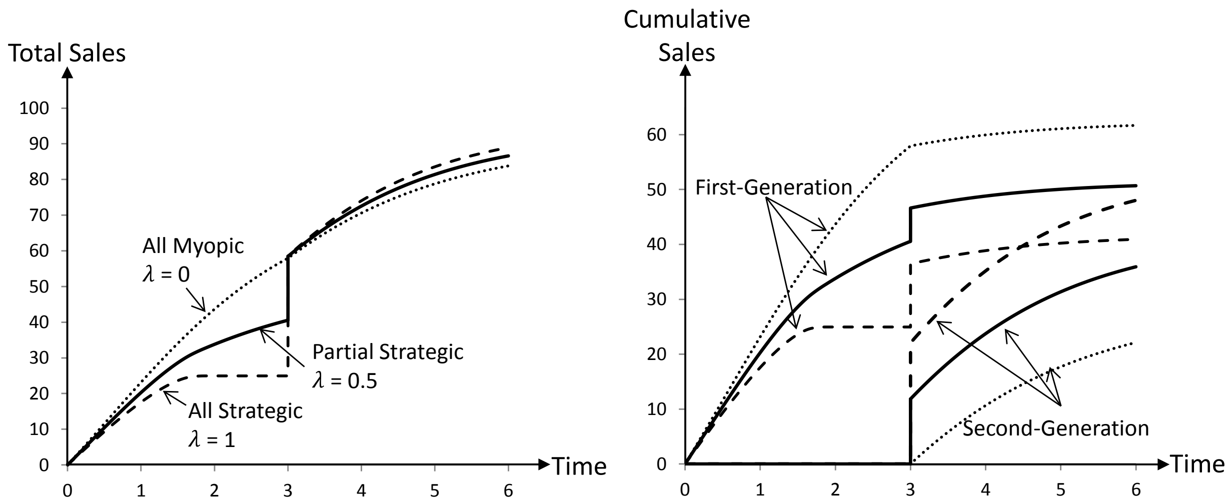


Figure 3: Comparison of Sales Diffusion Dynamics: Myopic vs. Strategic Consumers

second-generation product.

In contrast, the solid curve, which represents the general case where some of the consumers are strategic and some are myopic, replicates the Apple sales pattern we observed in Figure 1. At the beginning, the sales curve has a slower increasing rate compared to the myopic case. As time goes by, more and more strategic consumers prefer to wait—either for the better second-generation product (Leapfrog) or for the price cut for the first-generation product (Laggard). At time $\tau = 3$, we observe a sales jump for both generations. (See the two solid curves in the right panel.) The sales increase for the first generation is 6.1: an addition of 8.6 Laggards and a deduction of 2.5 Upgraders. Because this jump represents the net effect of Laggards minus Upgraders, the change to the first-generation product adopters (the installed base) could be either upward or downward, depending on the relative magnitude of these two segments of consumers. The sales increase for the second-generation product is 11.8: 9.3 Leapfroggers and 2.5 Upgraders from the first generation. This jump is always upward. The final sales for the first and second generations are 50.7 and 35.9, respectively, which is smaller and larger than the myopic case of 61.7 and 22.1. The total sales is 86.6, which is greater than the myopic case of 83.8. Hence, the myopic model underestimates the sales of the second-generation product.

The dashed curve represents another extreme case, in which all consumers are strategic.

As the proportion of strategic consumers gets larger in the population, sales of the first-generation product is further slowed. We observe a flat growth rate in the time interval (1.7,3), where almost all consumers wait until the introduction of the second generation at $\tau = 3$; then we observe a large sales increase for both generations (the dashed curves in the right panel). A higher proportion of strategic consumers leads to bigger jumps at the time of the second generation's introduction. The final second-generation sales hits 48, and the total sales for both generations is 89; both of these figures are higher than in the other two cases.

Conventional wisdom tells us that, everything else being equal, the firm's sales and profits decrease in the presence of strategic consumers because strategic consumers might prefer to delay their purchase (Besanko and Winston, 1990; Levin et al., 2009; Liu and Zhang, 2013). In sharp contrast, our results show that the seller could be better off in the presence of strategic consumers. The total discounted profit for the all-strategic case ($\lambda = 1$) is the highest; at 17.4, it is greater than the profit for the partial strategic case (17.2 when $\lambda = 0.5$) and the myopic case (16.9 when $\lambda = 0$). The superior performance can be explained as follows: As the number of strategic consumers increases, the number of initial adopters of the second-generation product increases. This initial seeding helps to generate stronger word-of-mouth influence, which speeds up the market penetration of the second-generation product and leads to faster product diffusion of the second generation. The seller benefits from the higher volume of second-generation product sales that can be reached at the end of the planning horizon. Ultimately, the gain from the total sales increase of the second-generation product outweighs the loss of the total sales decrease of the first-generation product. Our main insight is that the presence of strategic consumers does not necessarily hurt the seller, largely because of the initial "seeding" effect of the strategic consumers who choose to wait. Our findings underscore the importance of considering the effect of strategic consumers on the multi-generation product diffusion process. In contrast with the current practice of offering free trial software for a limited amount of time (Jiang, 2010; Cheng and Liu, 2012; Dou et al., 2012), our results suggest that an alternative promising strategy is to educate strategic consumers. Although the two approaches have the same effect of jump-starting the diffusion, our approach has the advantage of not compromising profits because strategic consumers eventually pay for the product they purchase.

That strategic consumers might hurt the seller under some circumstances is also entirely possible because of the lower volume of sales for the first generation and the delayed profit realization resulting from consumers' strategic waiting behavior. The influence of strategic consumers on profit and sales depends on, among other things, the discount-to-price ratio of the first generation relative to the performance improvement in the second generation. When the relative discount is small, the "seeding" effect on the second-generation product dominates. When the relative discount is large, the "cannibalization" effect on the first-generation product dominates. Overall, our model is flexible enough to generate a wide range of sales patterns. It recognizes the possibility of a sharp jumpstart of sales for the second-generation product when the product is introduced. The jump is well explained by strategic consumers' waiting and is supported by the recent empirical sales data in the high-tech industry (e.g., iPhone sales data in Figure 1).

5.2 Optimal Product Release Time with Pre-Announced Prices

Normative guidelines in the literature suggest that a firm should either introduce a new generation as soon as it is available or delay its introduction until the maturity stage (or end) of the preceding generation (Mahajan and Muller, 1996; Wilson and Norton, 1989). Although a too early or premature release of the second-generation product results in forgone sales of and profits from the first-generation model, late release of the second generation delays its own market expansion and profit realization. Because the release time of the second-generation product influences both the product's own diffusion and the diffusion of its preceding generations resulting from consumer switching and substitution, firms need to consider the effects on demand of both generations simultaneously. In general, we observe three profit patterns when we vary the introduction time of the second-generation product, as shown in Figure 4.

Figure 4 shows how the firm's profit (left panel) and total sales of the two generations of the product (middle and right panels) change as the introduction time varies from the

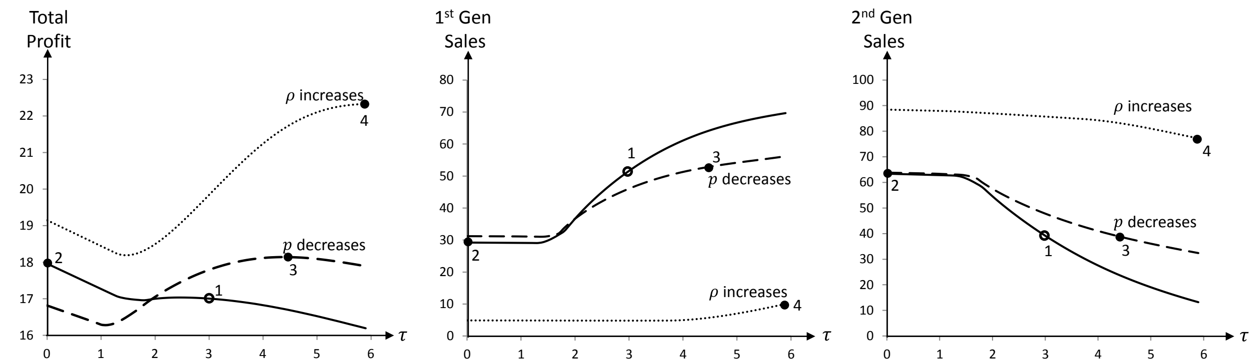


Figure 4: Effect of Entry Timing on Total Profit and Sales

beginning (a strategy of simultaneous introduction of the two generations) to the end of the planning horizon (a strategy of sequential introduction of two generations of the product). The solid curves are based on the benchmark values $p = 0.25$, $\delta = 0.05$, and $\mu = 0.16$, assuming that all consumers are strategic ($\lambda = 1$). The dashed and dotted curves contrast the scenarios when the product price p is higher and when the second-generation product has a larger performance improvement ratio ρ.

Several interesting observations are worth highlighting. First, note that the sales dynamics presented in Figure 3 are associated with Node 1 on the solid black curve in Figure 4, with the corresponding introduction time at $\tau = 3$. We see that introducing the second generation at $\tau = 3$ is not optimal. In fact, the seller can make a higher profit by adopting a simultaneous introduction strategy (Node 2 at $\tau = 0$), earning a higher profit of 17.96. Simultaneous release of both products is similar to a versioning strategy in which the firm offers high-valuation consumers a high-quality product at a higher price and low-valuation consumers a low-quality product at a lower price.

In fact, the seller can further increase its profit by reducing the price p from 0.25 to 0.23. As shown by Node 3 on the dashed curves, the highest profit of 18.14 is obtained at $\tau = 4.4$; the final sales for the first generation is 52.6 and for the second generation is 38.9. Compared with Node 2, reducing the price has two effects. First, the lower selling price results in a relatively higher conversion rate, which helps develop the first-generation market in an initial stage of the planning horizon. Second, a late introduction time for the second-generation product discourages strategic consumers' waiting, making high valuation consumers less willing to postpone their purchase. Both effects contribute to the first-generation product sales and the seller's early profit realization, resulting in a higher total discounted profit. The solid and dashed profit curves demonstrate the "now" and "maturity" optimal introduction timing strategies identified in the literature (Mahajan and Muller, 1996).

A third type of profit curve is shown as the dotted curves in Figure 4. In this case, the optimal introduction time (Node 4) is at the end of the planning horizon $\tau = 5.9$, which represents the "never" optimal introduction timing strategy (Wilson and Norton, 1989). The curve is obtained when we increase the quality improvement ratio ρ from 1.1 to 1.2. This improvement implies that the second generation is a much better product. Higher valuations from consumers make the second-generation product more attractive and therefore generate a higher risk of sales cannibalization of the first-generation product. The firm thus postpones the introduction time as late as possible, so that more consumers would prefer to adopt the first-generation product at the beginning and upgrade to the newer generation when it is available. Ultimately, our model supports various introduction timing strategies discussed in the literature. Because consumers are strategic in nature, overlooking consumers' strategic behavior would lead to non-optimal timing and inaccurate predictions of the firm's sales and profits.

Next, we compare the effect of strategic consumers on the firm's optimal introduction time under different price and discount policies. Under the base scenario, we focus on four price values $p = \{0.22, 0.23, 0.24, 0.25\}$ and choose three price discount levels $\delta = \{0.02, 0.03, 0.05\}$.[6] The left panel in Figure 5 contrasts the different optimal introduction times under the different price pairs (p, δ) for strategic consumers and non-strategic consumers. We plot the corresponding profits and total sales in the middle and right panels.

Comparing the upper left panel of Figure 5 with the lower left panel, we see that a wider range of optimal timings is possible when strategic consumers are taken into consideration.

[6] Justification of the parameter values is as follows: If $p \geq 0.26$, then $\frac{p-\mu}{\rho-1} \geq 1$, and no strategic consumers would choose to upgrade. The no-arbitrage condition $p - \delta \leq \mu$ requires that $p > 0.21$ for $\delta = 0.05$. These boundary values define the price range. In addition, if $\delta \leq 0.02$, no one would buy the discounted first-generation product. This condition imposes the lower bound on the discount.

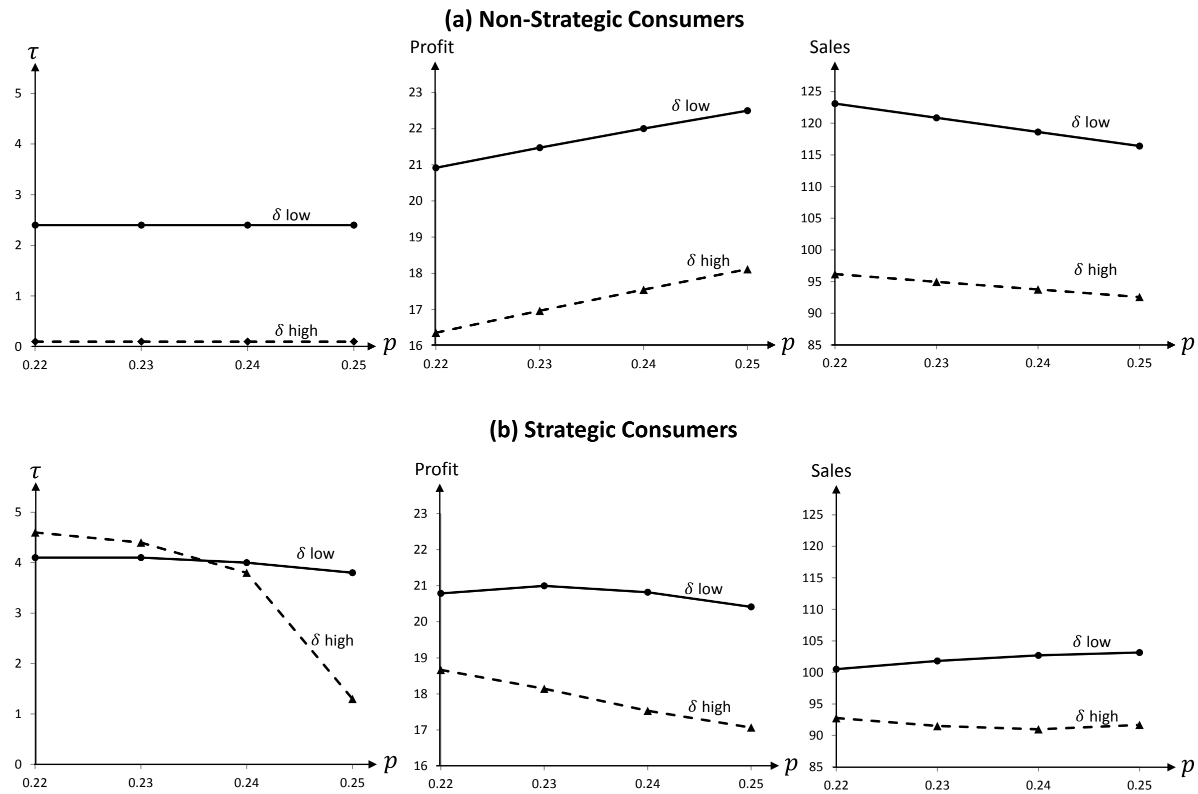


Figure 5: Optimal Timing Under Given Pricing Strategies

If we do not consider strategic consumers, the optimal timing is relatively stable against price increases. When the discount is relatively small, the firm prefers to delay the introduction of the second generation to avoid the cannibalization of sales. When the discount is relatively large, the firm prefers to use a simultaneous introduction strategy. However, in the presence of strategic consumers, the optimal introduction time of the second-generation product becomes earlier as the product price increases. Moreover, the preference for an earlier introduction time becomes stronger when the price discount for the first-generation product becomes larger.

In terms of profit, the upper middle panel in Figure 5 shows that, without the consideration of strategic consumers, the profit always increases in price. Correspondingly, the upper right panel shows that total sales decrease as the price increases. However, in the presence of strategic consumers, the effect of price on profit and sales is not linear. For example, when the discount is low, the profit first increases and then decreases as the price increases.

The interdependence of the nonlinear price effect and the introduction timing effect must be recognized. A relatively early release time not only cannibalizes the first-generation product sales, but also diminishes the strategic consumers' market seeding effect, which negatively affects the second-generation product sales. Therefore, either insufficient market seeding of the second generation or the early discount of the first generation can lead to lower overall profits. Comparing profits with and without the consideration of strategic consumers, we observe that firms tend to overestimate profits when the selling price is relatively high and to underestimate profits when the selling price is relatively low.

Overall, we find that firms are more likely to use a simultaneous introduction strategy when they overlook the consumers' strategic purchasing behavior. This finding indicates that ignorance of consumers' strategic behavior can lead firms to release products at the wrong time, which often results in an overestimation of sales and either overestimation or underestimation of the firm's profitability.

5.3 Optimal Pricing and Timing Strategies

In the previous section, we investigated the introduction timing of the second-generation product on the seller's profit implications under given price and discount schedules. In this section, we perform numerical optimization to jointly optimize the pricing and timing decisions.

We start with the base scenario values and then vary the key performance improvement parameter from $\rho = 1.1$, to 1.2 and 1.3, representing three scenarios of increasing performance improvement of the second-generation product. The three scenarios are presented in Figure 6. Each scenario is examined using a series of numerical searches in the three-dimensional parameter space (p, δ, τ). We thoroughly explore the price parameter space by varying the values of p and δ in the allowable price range, using an increment of 0.01, and by varying the introduction time of the second generation from 0 to T using an increment of 0.1. We plot a series of curves to illustrate the different cases in Figure 6. The left panels show the general case, in which all five segments of strategic consumers exist, and the right panels show the special case of complete substitution, in which no consumers buy the discounted first-generation product after the second generation is released.

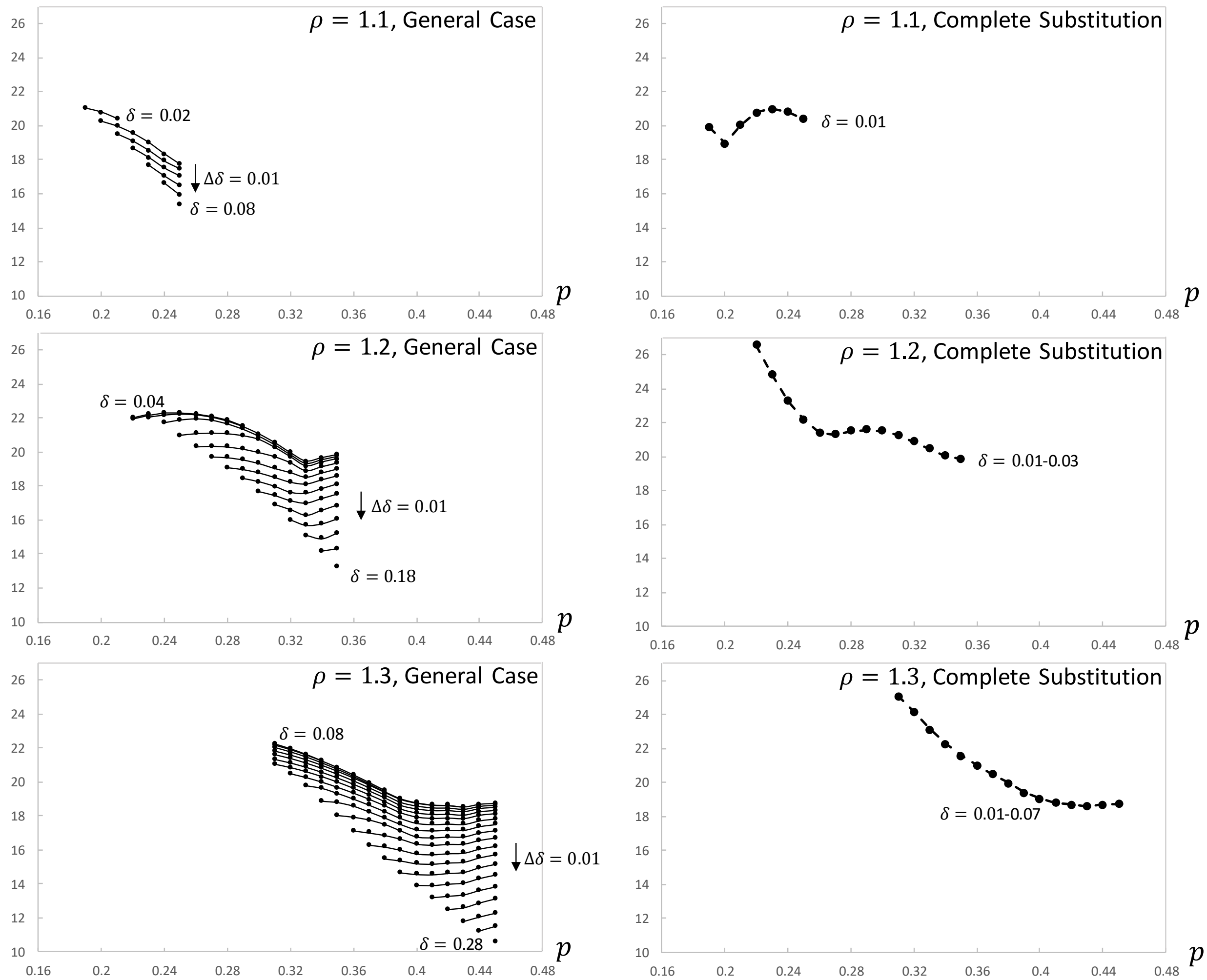


Figure 6: Comparison of Optimal Profit: General Case vs. Complete Substitution

Each dot in Figure 6 represents the highest profit that can be obtained under a specific (p, δ) configuration, and each curve corresponds to a specific discount δ. Furthermore, each curve is cut off by lower and upper bounded prices. No arbitrage condition requries $p > \mu+\delta$, which gives the lower bound for each curve in the figure. In order to ensure Upgrade option is attractive, we must have $\frac{p-\mu}{\rho-1} < 1$. In the base scenario, $\mu = 0.16$. When $\rho = 1.1$, 1.2, and 1.3, respectively, $p = 0.26$, 0.36, and 0.46 define the upper bound of the price. If the right-hand-side of the curve is cut off before it reaches the upper bound of the price in the general case, the case degenerates to the special case of complete substitution (as in Figure 2(a)). In the case of complete substitution, the price discount does not play a role, because no consumers would be interested in purchasing the discounted first-generation product after the second generation is released. For example, when $\rho = 1.1$ and $\delta = 0.02$, if $0.22 \leq p < 0.26$, then the scenario under the general case degenerates to the complete substitution case. The profit curve in this price range overlaps with the dashed line in the upper right panel in Figure 6.

Comparing the general case in the three scenarios (the left panels in Figure 6), we see that the seller can more feasibly charge a higher price and offer a larger discount as the performance improvement becomes larger. We see that, for a given discount on the first-generation product, the effect of price on profit is non-linear. The profit curves can exhibit many different shapes—such as strictly decreasing; first decreasing and then increasing; or first increasing, then decreasing, and then increasing again. In contrast, the effect of the price discount on the firm's profit is monotonic. Holding the price constant, the firm's optimal profit decreases as the discount increases.

The optimal prices are found by identifying the dot that indicates the highest profit in each subfigure. Because older generation products often co-exist with the new-generation product in consumer durables, the general case is more realistic and interesting than the complete substitution case. However, we observe that the firm can earn higher profit by pricing strategically to induce complete substitution in some cases. In our example, although the firm cannot yield a higher profit than the general case when $\rho = 1.1$, it can achieve a higher profit by pricing at $p = 0.22$ and $p = 0.31$ when $\rho = 1.2$ and $\rho = 1.3$, respectively. As long as the discount δ is no higher than 0.03 and 0.07, respectively, consumers are not interested in buying the discounted first-generation product after the second generation becomes available, resulting in complete substitution of first-generation product demand. The result seems to suggest that a firm can induce complete substitution to maximize its profit when the performance improvement of the product is relatively large.

To see the effect of other key model parameters on the firm's optimal pricing and timing strategies, the following table summarizes a few interesting scenarios and presents the jointly optimal decisions (p, δ, τ) and the resulting optimal total sales and total discounted profits. We focus on the more interesting general case.

ρ	Scenarios	Parameter values	p	δ	τ	Total Sales	Total Profit
1.1	Base	$T = 6, \mu = 0.16, r = 0.1$	0.19	0.02	4.3	93.45	21.04
	Longer Horizon	$T = 12$	0.19	0.02	4.6	99.48	21.63
	Lower Salvage	$\mu = 0.1$	0.18	0.01	3.9	115.48	17.46
	Lower Discount	$r = 0.05$	0.19	0.01	4.9	92.26	25.80
1.2	Base	$T = 6, \mu = 0.16, r = 0.1$	0.25	0.05	5.9	87.22	22.33
	Longer Horizon	$T = 12$	0.25	0.05	6.1	95.69	23.36
	Lower Salvage	$\mu = 0.1$	0.22	0.04	5.9	90.52	18.89
	Lower Discount	$r = 0.05$	0.31	0.06	5.9	68.16	17.74
1.3	Base	$T = 6, \mu = 0.16, r = 0.1$	0.31	0.08	5.9	80.53	22.20
	Longer Horizon	$T = 12$	0.30	0.07	7.6	92.26	24.70
	Lower Salvage	$\mu = 0.1$	0.25	0.06	5.9	87.13	20.09
	Lower Discount	$r = 0.05$	NA	NA	NA	NA	NA

Table 1: Optimal Pricing and Timing Strategies Under the General Case

In terms of the optimal selling strategies in the base scenario, we see that $p = 0.19$ and $\delta = 0.02$ yield the highest profit of 21.04 when $\rho = 1.1$; that $p = 0.25$ and $\delta = 0.05$ yield the highest profit of 22.33 when $\rho = 1.2$; and that $p = 0.31$ and $\delta = 0.08$ yield the highest profit of 22.20 when $\rho = 1.3$. Having a very high level of performance improvement is not necessary to make a higher profit. Examining the optimal timing, we see that $\tau = 4.3$ when $\rho = 1.1$, and $\tau = 5.9$ when $\rho = 1.2$ and 1.3. These findings support the "maturity" (Mahajan and Muller, 1996) or "never" (Wilson and Norton, 1989) optimal timing discussed in the literature. Ultimately, Table 1 reveals three primary findings for the base scenario: (1) Higher performance improvement in the second-generation product supports a higher optimal price and discount; (2) The optimal introduction time tends to be late in the product life cycle to mitigate the negative cannibalization effect; and (3) Total profit is not linearly increasing in performance improvement.

If the majority of market penetration has been achieved across a short time horizon (more than 80, compared with the potential market size of 120 in all base secnarios), then extending the selling horizon has little effect on the optimal prices and marginally delays the optimal introduction time. An introduction timing strategy that aims for "maturity" might emerge as the optimal one if the planning horizon is longer.

If the salvage value decreases from 0.16 to 0.1, we see that both the optimal price and the discount tend to decrease in all three scenarios. The optimal introduction time is earlier when the performance improvement is relatively small ($\rho = 1.1$). In this case, even though the total sales increase from 93.45 to 115.48, the total profit decreases. The additional sales seem unable to compensate for the loss that results from the lower profit margin.

If the consumers have a lower discount rate on their utility, then the firm should avoid a higher performance improvement. In fact, in the case of $\rho = 1.3$, no feasible price exists to support the general case in which all five segments of strategic consumers exist. Intuitively, when consumers are very patient (i.e., they have a low discount rate), they would prefer to wait for the better second-generation product rather than adopting the first-generation product immediately. Hence, the Adopt option is dominated. When the performance improvement is relatively small ($\rho = 1.1$), the optimal prices are relatively insensitive to parameter changes. This relatively small range of best pricing options generally favors a later introduction time (from 4.3 to 4.6) and yields a higher total discounted profit. When the performance improvement is relatively large ($\rho = 1.2$), the firm tends to charge a higher retail price because of the better performance of the second-generation product. However, the higher profit margin hurts the sales of both generations of the product (total sales decline from 87.22 to 68.16), leading to a lower total profit. Thus, the firm is better off when the performance improvement rate is relatively small and worse off when it is relatively large.

In summary, higher performance improvement and lower salvage value would generally support a higher optimal price, a larger discount, and a later introduction time. If the planning horizon is longer, the "maturity" introduction timing would emerge more often as optimal. In addition, the firm can benefit from patient consumers when the performance improvement is relatively small, and it can induce complete substitution to maximize profit when the performance improvement is relatively large.

6 Conclusion

Rapid product innovation is common in many industries today. As the product life cycle becomes increasingly shorter and new technologies are more frequently introduced into the market, understanding how consumers strategically respond to such trends is becoming an urgent need. This paper studies the effect of strategic consumers on a monopolistic firm's multi-generation product diffusion strategies. We extend the original Norton and Bass (1987) model, which captures diffusion and substitution at the aggregate level, to the consideration of adoption decisions at the individual level, in which strategic choices by consumers determine the dynamics of the interdependent demand growth of two generations of a product. By incorporating an enriched behavior model of consumers, our micromodeling approach better captures the multi-generation product sales interactions and allows for market segmentation in a way that has been impossible in traditional multi-generation product diffusion models. Understanding the segmentation of the target population in terms of consumers' adoption preferences enables firms to better predict their sales trajectory prior to a product launch. In addition, better understanding the interrelationality of products and its underlying dynamics can inform marketers for more accurate product forecasting and more effective allocation of marketing resources.

Our results underscore the importance of taking into account strategic consumers' purchase behavior when managing successive generations of product diffusion. We provide several major insights into optimal pricing and timing strategies that take into account the demand dependency between the two generations of the product. First, we find that the seemingly undesirable delayed purchase from strategic consumers could actually be beneficial to a firm. The reason is that consumers who have waited for the second-generation product are automatically converted to initial adopters of the new generation, which is essential to boost word-of-mouth sales in the early stage of the new product diffusion process. Second, we show that the influence of strategic consumers on profit and sales depends largely on the discount-to-price ratio of the first generation relative to the performance improvement in the second generation. When the relative discount is very small, the "seeding" effect on the second-generation product dominates. When the relative discount is large, the "cannibalization" effect on the first-generation product dominates. When a firm does not consider strategic consumers, it tends to overestimate (underestimate) profit when the selling price is relatively high (low), which causes the firm to introduce the second generation earlier than it should. Third, we demonstrate that various optimal entry timings recommended in the literature (i.e., “now,” “maturity,” or “never”) can occur under different market conditions. Our model is flexible enough to predict various sales patterns that can be fitted with the empirical data. Fourth, we show that higher performance improvement and lower salvage value generally would support a higher optimal price, a larger discount, and a later introduction time. If the planning horizon is longer, more “maturity” introduction timing would emerge as optimal. In addition, the firm can benefit from patient consumers when the performance improvement is relatively small, and it can induce complete substitution to maximize profit when the performance improvement is relatively large. Finally, as seen in the Online Appendix, we find that product development costs and production costs also affect the optimal pricing and timing strategy. When costs are considered, firms prefer extreme introduction timing (either at the beginning or the end of the first-generation product life cycle) in a larger range of market conditions.

This study has a few limitations. First, we account only for the demand-side dynamics without considering the supply-side capacity rationing. That is, the firm in our model does not manipulate product availability to influence consumers' purchase decisions. Second, we do not incorporate valuation uncertainty into our model. Third, we use a static markdown-pricing strategy. Future work might relax the static pricing strategy and allow strategic consumers to change their behavior by developing price expectations under dynamic pricing. Future work also might examine the effect of market competition on a firm's strategies in managing its multi-generation product diffusion. In addition, empirical testing of the model using data from high-tech firms is desirable. These extensions present interesting future research opportunities.

Acknowledgment

The authors thank the senior editor, the associate editor, and the anonymous referees for many valuable comments and helpful suggestions.

Appendix

Proof of Lemma 1

Proof. (a) Strategic consumers who adopt the first generation choose to upgrade if and only if $v - p + e^{-r(\tau-t)}[(\rho-1)v - p + \mu] \geq v - p$; that is, if $v \geq \frac{p-\mu}{\rho-1}$. Because $v \in [0,1]$, no consumers upgrade if $\rho - 1 \leq p - \mu$.

(b) Strategic consumers prefer Leapfrog to Laggard if and only if $e^{-r(\tau-t)}(\rho v - p) \geq e^{-r(\tau-t)}(v - p + \delta)$; that is, if $v \geq \frac{\delta}{\rho-1}$. Because $v \in [0,1]$, if $\delta \geq \rho - 1$, Leapfrog is dominated by Laggard.

(c) Consumers may choose Leapfrog only if $v \geq \frac{p}{\rho}$ and may choose Laggard only if $v \geq p - \delta$. Based on the proof in (b), the consumer with $v = \frac{\delta}{\rho-1}$ is indifferent between Leapfrog and Laggard options. Therefore, the proportion of strategic consumers who prefer Laggard to Leapfrog is in the interval $\left[(p-\delta), \max\{\frac{\delta}{\rho-1}, p-\delta\}\right]$. If $\frac{\delta}{\rho-1} \leq p - \delta$, or, equivalently, if $\delta \leq \frac{p(\rho-1)}{\rho}$, the interval is empty and no strategic consumers choose Laggard. □

Proof of Proposition 1

Proof. We define the indifferent curve between Leapfrog and Laggard options as $l_{lg}(t)$, which is determined by $e^{-r(\tau-t)}(\rho v - p) = e^{-r(\tau-t)}[v - (p-\delta)]$. We thus have $l_{lg} = \frac{\delta}{\rho-1}$, and consumers with $v > l_{lg}$ prefer Leapfrog than Laggard. Similarly, the indifferent curve between Upgrade and Adopt options $l_{ua} = \frac{p-\mu}{\rho-1}$, and consumers with $v \geq \frac{p-\mu}{\rho-1}$ prefer Upgrade than Adopt.

We can verify that $l_{ag}(t)$ is increasing in t, and thus $t_1 = l_{ap}^{-1}(\frac{\delta}{\rho-1}) < l_{ap}^{-1}(\frac{p-\mu}{\rho-1}) = t_2$ because of the assumption $\frac{\delta}{\rho-1} < \frac{p-\mu}{\rho-1}$. We can also verify that $l_{al}(t)$ is increasing in t. In addition, the curves $l_{ua}(t)$ and $l_{al}(t)$ interact at $\frac{p-\mu}{\rho-1}$. Therefore, $t_2 = l_{ua}^{-1}(\frac{p-\mu}{\rho-1}) = l_{al}^{-1}(\frac{p-\mu}{\rho-1}) < l_{ul}^{-1}(1) = t_3$ because of the assumption that $\frac{p-\mu}{\rho-1} < 1$.

Notice that consumers might choose Leapfrog if $v \geq \frac{p}{\rho}$, and consumers might choose Laggard if $v \geq p - \delta$. Consumers with $v \geq p$ prefer Adopt to Non-Adopt, and consumers with $v \geq \frac{p-\mu}{\rho-1}$ prefer Upgrade to Adopt. Because $\min\left\{p-\delta, \frac{p}{\rho}\right\} < p < \frac{p-\mu}{\rho-1}$, among all consumers who choose options other than Non-Adopt, low-value consumers only choose Leapfrog or

Laggard, rather than Adopt or Upgrade. Medium-value consumers might choose Adopt, and high-value consumers might choose Upgrade.

We consider different value segments as follows.

(1) For $v \in [\frac{p-\mu}{\rho-1}, 1]$, Upgrade dominates Adopt because $l_{ua} < \frac{p-\mu}{\rho-1}$, and Leapfrog dominates Laggard because $l_{lg} < \frac{p-\mu}{\rho-1}$. Therefore, the options that might appear to be optimal are Upgrade and Leapfrog. By the definition of $l_{ul}(t)$, when $t < t_2$, consumers choose Upgrade and when $t > t_3$, consumers choose Leapfrog. When $t_2 < t < t_3$, consumers with $v > l_{ul}(t)$ choose Upgrade and the others choose Leapfrog.

(2) For $v \in [0, \frac{p-\mu}{\rho-1}]$, we distinguish two cases.

(2.1) If $\frac{\delta}{\rho-1} \leq \frac{p}{\rho}$, Leapfrog dominates Laggard for all consumers who derive positive utility from the Laggard option. Therefore, the options that might appear to be optimal are Adopt and Leapfrog. By the definition of $l_{al}(t)$, when $t < t_2$, consumers with $v > l_{ap}(t)$ choose Adopt, and the others choose Leapfrog. When $t > t_2$, consumers choose Leapfrog.

(2.2) If $\frac{\delta}{\rho-1} > \frac{p}{\rho}$, which implies $\frac{p}{\rho} > p - \delta$, low-value consumers might choose the Laggard option. (i) For $v \in [p-\delta, \frac{\delta}{\rho-1}]$, the options that might appear to be optimal are Laggard and Adopt. If $l_{al}(0) > \frac{\delta}{\rho-1}$, because of the monotonicity of $l_{al}(t)$, Adopt is dominated by Leapfrog for all the consumers in this region, and Leapfrog is dominated by Laggard. Therefore, all the consumers choose Laggard. If $l_{al}(0) < \frac{\delta}{\rho-1}$, noticing that $l_{al}(t)$ and $l_{ag}(t)$ intersect at $\frac{\delta}{\rho-1}$, we have $t_1 = l_{al}^{-1}(\frac{\delta}{\rho-1}) = l_{ag}^{-1}(\frac{\delta}{\rho-1})$. Therefore, when $t < t_1$, consumers with $v > l_{ag}(t)$ choose Adopt and the others choose Laggard. When $t > t_1$, all the consumers choose Laggard.

(ii) For $v \in [\frac{\delta}{\rho-1}, \frac{p-\mu}{\rho-1}]$, the options that might appear to be optimal are Adopt and Leapfrog. By the definition of $l_{al}(t)$, when $t \leq \max\{0, t_1\}$, consumers choose Adopt, and when $t > t_2$, consumers choose Leapfrog. When $\max\{0, t_1\} < t < t_2$, consumers with $v > l_{al}(t)$ choose Adopt, and the others choose Leapfrog.

Organizing consumers' choice along the time dimension, we conclude consumer segmentation in the proposition. □

References

Apple. 2012. iPhone 5 pre-orders top two million in first 24 hours. *Apple Press Info* September 17.

Aviv, Y., A. Pazgal. 2008. Optimal pricing of seasonal products in the presence of forward-looking consumers. *Manufacturing and Service Operations Management* **10**(3) 339–359.

Bala, R., S. Carr. 2009. Pricing software upgrades: The role of product improvement and user costs. *Production and Operations Management* **18**(5) 560–580.

Balakrishnan, P.V., S. Pathak. 2014. Impeding the juggernaut of innovation diffusion: A production-constrained model. *Production and Operations Management* **23**(7) 1183–1197.

Bass, F. M. 1969. A new-product growth model for consumer durables. *Management Science* **15**(1) 215–227.

Besanko, D., W. L. Winston. 1990. Optimal price skimming by a monopolist facing rational customers. *Management Science* **36**(5) 555–567.

Bhattacharya, S., V. Krishnan, V. Mahajan. 2003. Operationalizing technology improvements in product development decision-making. *European Journal of Operational Research* **149**(1) 102–130.

Chatterjee, R., J. Eliashberg. 1990. The innovation diffusion process in a heterogeneous population: A micromodeling approach. *Management Science* **36**(9) 1057–1079.

Cheng, H.K., Y. Liu. 2012. Optimal software free trial strategy: The impact of network externalities and consumer uncertainty. *Information Systems Research* **23**(2) 488–504.

Chu, C.P., J.G. Pan. 2008. The forecasting of the mobile internet in Taiwan by diffusion model. *Technological Forecasting and Social Change* **75** 1054–1067.

Danaher, P.J., B.G.S. Hardie, W.P. Putsis Jr. 2001. Marketing-mix variables and the diffusion of successive generations of a technological innovation. *Journal of Marketing Research* **38** 501–514.

Dolan, R.J., A.P. Jeuland. 1981. Experience curve and dynamic demand models: Implications for optimal pricing strategies. *Journal of Marketing* **45** 52–73.

Dou, Y., M.F. Niculescu, D.J. Wu. 2012. Engineering optimal network effects via social media features and seeding in markets for digital goods and services. *Information Systems Research* forthcoming.

Feichtinger, G. 1982. Optimal pricing in a diffusion model with concave price-dependent market potential. *Operations Research Letters* **1**(6) 236–240.

Goldenberg, J., B. Linai, E. Muller. 2002. Riding the saddle: How cross-market communications can create a major slump in sales. *Journal of Marketing* **66** 1–16.

Golder, P.N., G.J. Tellis. 1997. Will it ever fly? modeling the takeoff of really new consumer durables. *Marketing Science* **3** 256–270.

Hiebert, L.D. 1974. Risk, learning and the adoption of fertilizer responsive seed varieties. *American Journal of Agricultural Economics* **56** 764–768.

Ho, T., S. Savin, C. Terwiesch. 2002. Managing demand and sales dynamics in new product diffusion under supply constraint. *Management Science* **48** 187–206.

Islam, T., N. Meade. 1997. The diffusion of successive generations of a technology: A more general model. *Technological Forecasting and Social Change* **56** 49–60.

Jain, D.C., V. Mahajan, E. Muller. 1991. Innovation of diffusion in presence of supply restrictions. *Marketing Science* **10**(1) 83–90.

Jensen, R. 1982. Adoption and diffusion of an innovation of uncertain profitability. *Journal of Economic Theory* **27** 182–193.

Jiang, Z. 2010. How to give away software with successive versions. *Decision Support Systems* **49**(4) 430–441.

Jiang, Z., D. Jain. 2012. A generalized Norton-Bass model for multigeneration diffusion. *Management Science* **58**(10) 1887–1897.

Jun, D.B., 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.

Kalish, S. 1985. A new product adoption model with price, advertising, and uncertainty. *Management Science* **31**(12) 1569–1585.

Ke, T.T., Z.M. Shen, S. Li. 2013. How inventory cost influences introduction timing of product line extensions. *Production and Operations Management* **22**(5) 1214–1231.

Kim, N., R.K. Srivastava, J.K. Han. 2001. Consumer decision-making in a multi-generational choice set context. *Journal of Business Research* **53** 123–136.

Kim, W.J., J.D. Lee, 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** 825–849.

Krankel, R.M., I. Duenyas, R. Kapuscinski. 2006. Timing successive product introductions with demand diffusion and stochastic technology improvement. *Manufacturing and Serive Operations Management* **8** 119–135.

Kurawarwala, A., H. Matsuo. 1996. Forecasting and inventory management of short life-cycle products. *Operations Research* **44** 131–150.

Levin, Y., J. McGill, M. Nediak. 2009. Dynamic pricing in the presence of strategic consumers and oligopolistic competition. *Management Science* **55**(1) 32–46.

Liu, Q., G. van Ryzin. 2008. Strategic capacity rationing to induce early purchases. *Management Science* **54**(6) 1115–1131.

Liu, Q., D. Zhang. 2013. Dynamic pricing competition with strategic customers under vertical product differentiation. *Management Science* **59**(1) 84–101.

Luckerson, V. 2014. Apple's $649 iPhone 6 reportedly costs $200 to make. *Time Inc. Network* September 24.

Mahajan, V., 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.

Mahajan, V., E. Muller, F.M. Bass. 1990. New product diffusion models in marketing: A review and directions for research. *Journal of Marketing* **54** 1–26.

Mehra, A., A. Seidmann, P. Mojumder. 2014. Product life-cycle management of packaged software. *Production and Operations Management* **23**(3) 366–378.

Moorthy, K., I. Png. 1992. Market segmentation, cannibalization, and the timing of product introductions. *Management Science* **38** 345–359.

Newton, C. 2012. iPad sales miss estimates as rate of adoption slowed. *CNET News.com*

October 25.

Niculescu, M.F., H. Shin, S. Whang. 2012. Underlying consumer heterogeneity in markets for subscription-based it services with network effects. *Information Systems Research* **23**(4) 1322–1341.

Norton, J. A., F. M. 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., R.G. Schwartz. 1988. Diffusion of new products in risk-sensitive markets. *Journal of Forecasting* **7** 273–287.

Padmanabhan, V., F.M. Bass. 1993. Optimal pricing of successive generations of product advances. *International Journal of Research in Marketing* **10** 185–207.

Peres, R., E. Muller, V. Mahajan. 2010. Innovation diffusion and new product growth models: A critical review and research directions. *International Journal of Research in Marketing* **27** 91–106.

Roberts, J.H., G.L. Urban. 1988. Modeling multiattribute utility risk, and belief dynamics for new consumer durable brand choice. *Management Science* **34**(2) 167–185.

Robinson, B., C. Lakhani. 1975. Dynamic price models for new-product planning. *Management Science* **21** 1113–1122.

Roger, E.M. 1983. *The Diffusion of Innovation*. Free Press, New York.

Seitz, P. 2015. Many consumers waiting for apple watch version 2. *Investors.com* September 30.

Song, I., P.K. Chintagunta. 2003. A micromodel of new product adoption with heterogeneous and forward-looking consumers: Application to the digital camera category. *Quantitative Marketing and Economics* **1** 371–407.

Stokey, N.L. 1981. Rational expectations and durable goods pricing. *Bell Journal of Economics* **12**(1) 112–128.

Su, X., F. Zhang. 2008. Strategic consumer behavior, commitment, and supply chain performance. *Management Science* **54**(10) 1759–1773.

Wilson, L.O., J.A. Norton. 1989. Optimal entry time for a product line extension. *Marketing*

Science **8** 1–17.

Online Appendix

A. Different Prices for the Two Generations

In the baseline model, we assume that the price of the second generation is the same as the original price of the first generation p. Such a pricing strategy has been widely observed in the consumer electronics industry (e.g., the prices for different generations of iPhones). In this extension, we consider a more general pricing strategy. We assume the price of the second-generation product is higher than the original price of the first generation, possibly to reflect the performance improvement. The reverse case can be similarly analyzed.

We denote the price of the second generation as $p+h$, where the price increment is such that $h \geq 0$. Similar to the baseline model, consumers have five options before the release of the second generation. The main difference is now that the price of the second generation is higher than the price considered in the baseline case. Consumers' payoffs under Adopt, Laggard, and Non-Adopt remain the same as in the baseline case: $v-p$, $e^{-r(\tau-t)}(v-p+\delta)$, and 0, respectively. Under Leapfrog, consumers wait and buy the second-generation product, and the expected payoff becomes $e^{-r(\tau-t)}(\rho v-p-h)$. Under Upgrade, consumers buy the first generation and upgrade to the second generation later, and the expected payoff becomes $v-p+e^{-r(\tau-t)}[(\rho-1)v-p-h+\mu]$.

We next illustrate the market segmentation by comparing it with the most complicated scenario in the baseline model, presented in Figure 2(c). We can easily replicate all the analyses for the other scenarios and show that all the results qualitatively remain. The indifference curve between Upgrade and Adopt is determined by $l_{ua} = \frac{p+h-\mu}{\rho-1} > \frac{p-\mu}{\rho-1}$. Compared to the baseline case, fewer consumers choose to upgrade because of the price increase of the second generation.

Similarly, the indifference curve l_{ul} is defined by $v-p = e^{-r(\tau-t)}(v-\mu)$, l_{al} is defined by $v-p = e^{-r(\tau-t)}(\rho v-p-h)$, and l_{ag} is defined by $v-p = e^{-r(\tau-t)}(v-p+\delta)$. In fact, only l_{al} shifts to the right. Compared to the baseline case, more consumers choose Adopt, rather than Leapfrog, because of the price increase associated with the Leapfrog option.

Also, compared to the baseline case, the Laggard option becomes relatively more attrac-

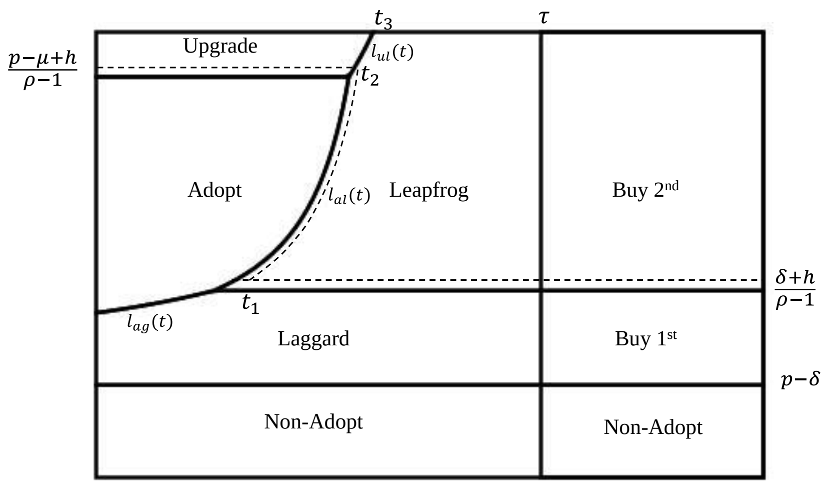


Figure 7: When Two Generations Have Different Prices

tive than the Leapfrog option, and more consumers choose the Laggard option because of the price increase associated with Leapfrog. The indifference curve now becomes $l_{lg} = \frac{\delta+h}{\rho-1}$. The consumers who are indifferent between Laggard and Non-Adopt remain the same because the price difference h has no effect on either option.

Figure 7 illustrates the market segments under the general pricing strategy. The dashed lines represent the shifts of indifference curves. Compared to the baseline case, the proportion of consumers who choose the Upgrade and Leapfrog options shrinks, while the proportion of consumers who choose the Adopt and Laggard options expands.

As we can see, the price difference h now plays a role in consumers' adoption choice and thus also affects the diffusion process. Using a similar computation, we find that all the main insights carry over to this extension, as long as the price increase is not dramatic. The overall effect on profit is unclear. The price increase in the second-generation product reduces the total number of adopters of the second generation, and it increases the number of adopters of the first generation because of the inter-generational substitution. Because the second-generation product has its own unique market potential, compared with the baseline case, the smaller number of Leapfroggers may have a significant effect on the diffusion rate of the second-generation product, resulting in more sales loss than can be compensated for by the larger number of adopters of the first-generation product. Therefore, the final effect on profit depends on whether the profit gain from a higher unit price in second-generation product sales can compensate for the profit loss from the cannibalization of second-generation sales from the first-generation product and the profit loss from being unable to attract new sales from the new market. Through our numerical simulation, we find that charging a higher second-generation price usually is unjustified. Although we are not able to analytically characterize the conditions under which the general pricing strategy is optimal, we find that pricing the two products the same usually gives the firm either optimal or close-to-optimal profit. This insight supports the current practice of pricing the new generation of product the same as the previous generation, while discounting the previous generation product price to pick up the remaining market potential.

B. Uncertain Release Time

Now consider the uncertainty involved in the release time of the second-generation product. For simplicity, we assume that with probability θ, the second-generation product will be released at an earlier time $\tau - \varepsilon$, and with probability $1 - \theta$, it will be launched at a later time $\tau + \varepsilon$. The probability of early release, θ, is determined by factors such as a firm's R&D capabilities. We assume θ is ex ante unknown to consumers, and it is uniformly distributed on the interval $[0, 1]$. Accordingly, the expected value of θ is $E\theta = \frac{1}{2}$. Without any information shared between the firm and consumers, and among consumers themselves, consumers make their wait-or-buy decisions based on their expectation of θ, and the expected release time is $\frac{1}{2}(\tau - \varepsilon) + \frac{1}{2}(\tau + \varepsilon) = \tau$. This is the same expected release time as in the baseline case.

The effect of uncertain release time on market segmentation is illustrated in Figure 8. When the strategic consumers do not have any information about the release time, we have the indifference curves indicated by the solid lines, which is the same as the base model. If the actual release time is $\tau - \varepsilon$ (or $\tau + \varepsilon$), and if the strategic consumers know the actual release time, we have the indifference curves indicated by the dotted lines (or the dashed lines). Note that only in the region where the release time is between the parallel bands defined by the dashed and the dotted lines do consumers make strategic errors because of their lack of information. For example, consider the case that $\theta = 0$; that is, the actual release time is $\tau + \varepsilon$. Consumers who become aware of the first-generation product between

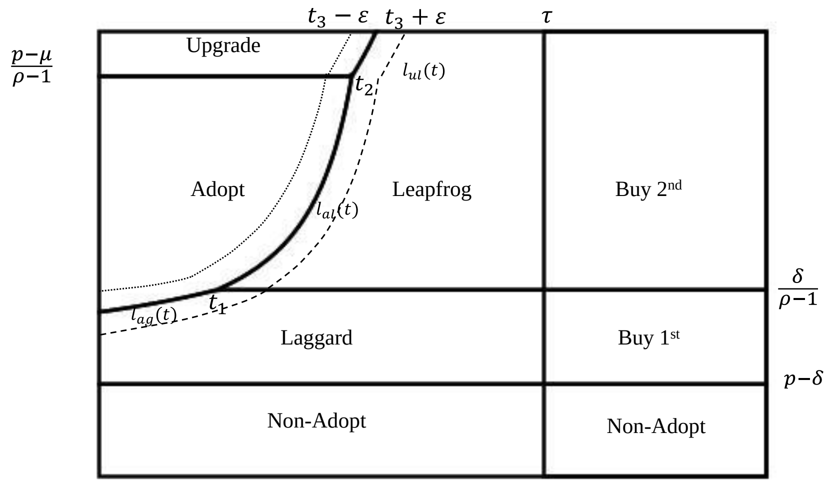


Figure 8: Uncertain Release Time

the solid lines and the dashed lines (covered in the Leapfrog region) should buy immediately. However, they choose Leapfrog in the absence of true information. In contrast, consumers who become aware of the first-generation product between the dotted lines and the solid lines (covered in the Upgrade and Adopt regions) are not affected.

In contrast to the price commitment, which is purely a strategic decision of the firm, the release time of the newer-generation product is affected by the uncertain R&D process and the newer-generation production technology. Technically, the firm might not be able to fully commit to the release time at the beginning of the planning horizon. In addition, although a commitment to the release time allows the firm to coordinate strategic consumers' actions by synchronizing the expected release time, whether commitment to the release time is beneficial is less clear because its profit consequence is ambiguous. As a result, the firm might choose to commit to its pricing strategy but keep the release time private. This strategy partly reflects the current practices by firms in managing their new product and software releases.

C. Effects of Product Development Cost and Production Cost on Optimal Strategies

New product introduction requires investments in product development and production. In this extension, we look at the effect of product development cost and production cost on the firm's optimal pricing and timing strategies.

When the selling horizon starts, the development cost associated with the first-generation product is sunk. However, the firm incurs some development cost to develop its second-generation product, which can be assumed as a convex function of the product quality improvement ρ. We express the total development cost as $k(\rho-1)^{\gamma}$, where $\gamma \geq 1$ and k is the sensitivity parameter for quality improvement.[7]

To understand the effect of production cost on the firm's strategies and profitability, we make several simplified assumptions. From an operational perspective, especially considering the short product life cycle and the widely adopted practice of international outsourcing, we assume the firm adopts a one-replenishment ordering policy, similar to Ke et al. (2013). Under this policy, the firm places the first order (or completes the production) of a certain amount of the first-generation products to satisfy all demand for the first-generation products at time 0. Then, right before the introduction of the second-generation product (i.e., at time τ), the firm makes the second order (or completes the production) of a certain amount of the second-generation products to satisfy all future demand for the second generation. We let c_1 be the production cost for the first generation. We assume the cost of producing the higher quality second-generation product is more expensive, and the unit production cost is proportional to its quality improvement. For simplicity, let $c_2 = \rho c_1$. To focus on the insights related to the procurement (or production) cost, we ignore the inventory holding cost because of the short product life cycle. We denote $x_1(T)$ as the total sales of the first-generation product, and $x_2(T)$ and $y_2(T)$ as the total sales of the second-generation product from the competing market and the unique market, respectively. The firm's discounted total profit expression can be modified as:

$$\begin{aligned}\pi_p = \ & \int_0^{\tau} e^{-rt} p\dot{x}_1(t)dt + e^{-r\tau}(p-\delta)\, g_1(\tau) + e^{-r\tau} p\,[l_1(\tau)+u_1(\tau)] + \int_{\tau}^{T} e^{-rt}(p-\delta)\dot{x}_1(t)dt \\ & + \int_{\tau}^{T} e^{-rt} p[\dot{x}_2(t)+\dot{y}_2(t)]dt - c_1 x_1(T) - e^{-r\tau} c_2\,[x_2(T)+y_2(T)] - k(\rho-1)^{\gamma}\end{aligned}$$

Building on the same parameter values as in the main text, we seek to understand how the unit production cost parameters c_1 and c_2 and the quality sensitivity parameter k affect

[7] We thank the AE and the anonymous reviewer for suggesting this modeling approach.

the firm's optimal entry timing and pricing strategies and profit. According to Luckerson (2014), Apple's \$649 iPhone 6 costs \$200 to make. We assume c_1 =0.05 and 0.08, which account for 20%∼35% of the product selling prices in our numerical studies, to represent the low unit production cost and the high unit production cost. For simplicity, let $\gamma = 2$. We assume k =200 and 500 to assess the cost effect of quality improvement, and $\rho = 1.1$ and 1.2, representing a small quality improvement and a large quality improvement.

δ	p	τ (no cost)		τ (with cost)					
		$\rho = 1.1$	$\rho = 1.2$	$\rho = 1.1$				$\rho = 1.2$	
				$c_1 = 0.05$		$c_1 = 0.08$		$c_1 = 0.05$	$c_1 = 0.08$
				$k = 200$	$k = 500$	$k = 200$	$k = 500$	$k = 200$	$k = 500$
0.02	0.22	4.1	5.6	+0.2	+0.2	+0.4	+0.4	+0	+0
	0.23	4.1	4.6	+0.2	+0.2	+0.3	+0.3	+0	+0
	0.24	4	4.1	+0.1	+0.1	+0.2	+0.2	+0	+0
	0.25	3.8	3.8	+0.1	+0.1	+0.2	+0.2	+0	+0
0.05	0.22	4.6	5.3	+0.5	+0.6	+1	+0.9	+0.5	Never
	0.23	4.4	5.5	+0.6	+0.6	+1	+1	+0.4	Never
	0.24	3.8	5.7	Now	Now	Now	Now	Never	Never
	0.25	Now	Never	Now	Now	Now	Now	Never	Never

Table 2: Effect of Quality Improvement and Cost Parameters on Optimal Timing Strategies

Table 2 shows the optimal introduction time of the second generation under different price-discount schedules (p, δ). The third and fourth columns present the benchmark case without considering the development cost and production cost. We see that the optimal introduction time decreases as the price increases. As the price discount becomes larger, the rate at which the introduction time decreases in price is faster. Moreover, higher quality improvement favors a later introduction time.

When the cost is considered, the effect of cost parameters on the optimal introduction time is different. A higher unit production cost generally delays the introduction time, while the effect of the quality development sensitivity parameter is insignificant. When both the price and discount are relatively high, a great quality improvement favors a "Never" strategy, while a small quality improvement favors a "Now" strategy, as we see from the last two rows of the table.

Table 3 shows the the revenue (without considering costs) and the net profit (after taking into account the product development and production costs) corresponding to the optimal

δ	p	Revenue (no cost)		Net Profit (with cost)					
		$\rho = 1.1$	$\rho = 1.2$	$\rho = 1.1$				$\rho = 1.2$	
				$c_1 = 0.05$		$c_1 = 0.08$		$c_1 = 0.05$	$c_1 = 0.08$
				$k = 200$	$k = 500$	$k = 200$	$k = 500$	$k = 200$	$k = 500$
0.02	0.22	20.79	26.54	14.90	11.90	8.12	11.12	10.80	12.89
	0.23	21.00	24.81	14.97	11.97	8.37	11.37	11.42	13.47
	0.24	20.82	23.26	14.66	11.66	8.62	11.62	11.57	13.72
	0.25	20.41	22.16	14.14	11.14	8.81	11.81	11.65	13.83
0.05	0.22	18.66	22.01	12.78	9.78	8.00	11.03	11.16	12.99
	0.23	18.14	22.19	12.18	9.18	8.17	11.17	11.1	12.96
	0.24	17.53	22.29	11.93	8.93	7.54	10.54	11.09	12.94
	0.25	17.96	22.33	12.51	9.51	7.52	10.52	11.07	12.91

Table 3: Effect of Quality Improvement and Cost Parameters on Revenue and Profit

introduction times presented in Table 2. When the quality improvement is relatively small (i.e., $\rho = 1.1$) and costs are not considered, the highest revenue is 21 under pricing strategies $p = 0.23$, $\delta = 0.02$, and the introduction time $\tau = 4.1$. When costs are considered, the firm delays the introduction time to $\tau = 4.3$. The optimal pricing strategies remain the same when the unit production cost is relatively low; when the unit production cost is relatively high, the firm charges a higher price, $p = 0.25$, to try to protect its profit margin.

When the quality improvement is relatively large (i.e., $\rho = 1.2$) and costs are not considered, the firm postpones the introduction time of the second generation to near the end of the first-generation life cycle ($\tau = 5.6$ vs. 4.1) and charges a lower price ($p = 0.22$ vs. 0.23). However, if costs are considered, the firm prefers to introduce the second generation earlier ($\tau = 3.8$) and charges a higher price ($p = 0.25$).

In summary, considering cost can significantly affect the firm's optimal pricing and introduction timing. Although the optimal prices might be lower than they are when cost is not considered, the optimal introduction timing tends to be more extreme; that is, more "Now" or "Never" strategies can be found to be optimal.