Optimal Market Entry Timing for Successive Generations of Technological Innovations

Zhengrui Jiang
School of Business, Nanjing University,
Nanjing 210093 CHINA {zjiang@nju.edu.cn}

Xinxue (Shawn) Qu
Mendoza College of Business, University of Notre Dame,
Notre Dame, IN 46556 50011 U.S.A. {xqu2@nd.edu}

Dipak C. Jain
CEIBS, Shanghai, CHINA 201203 {dipakcjain@ceibs.edu}

Appendix

Interpretation of the GNB Model Equations

Based on the Generalized Norton-Bass Model (Jiang and Jain 2012), the number of units-in-use for the two successive generations can be represented by the following equations:

$$S_1(t) = m_1F_1(t) - m_1F_1(t)F_2(t-\tau_2) = m_1F_1(t)[1-F_2(t-\tau_2)] \quad \text{(A1)}$$

$$S_2(t) = m_2F_2(t-\tau_2) + m_1F_1(t)F_2(t-\tau_2) = [m_2 + m_1F_1(t)]F_2(t-\tau_2) \quad \text{(A2)}$$

The term $m_1F_1(t)$ in Equation (1) represents the cumulative number of adoptions of G1, assuming that G2 were never introduced. Because G2 is introduced at time τ_2, a portion, $m_1F_1(t)F_2(t-\tau_2)$ to be exact, of the G1 adopters will substitute G1 with G2, hence the term is subtracted from $S_1(t)$ and added to $S_2(t)$. The substitution term counts both leapfroggers (those who skip G1) and switchers (those who upgrade from G1 to G2). The leapfrogging multiplier, i.e., the proportion of potential adoteprs of G1 who choose to leapfrog at time *t* is assumed to be $F_2(t-\tau_2)$ in the GNB model. Taking into consideration the leapfroggers, the instantaneous adoption rate for G1 takes the form

$$y_1(t) = m_1f_1(t)[1-F_2(t-\tau_2)] \quad \text{(A3)}$$

In addition to leapfrogging, existing adopters of G1 also switch to G2. The number of switchers take the form of $m_1F_1(t)f_2(t-\tau_2)$. Adding the leapfrogging and switching terms to the adoptions by the G2-specific adopters, we obtain the adoption rate for G2

$$y_2(t) = [m_2 + m_1F_1(t)]f_2(t-\tau_2) + m_1f_1(t)F_2(t-\tau_2) \quad \text{(A4)}$$

From Equations (3) and (4), we can obtain the cumulative numbers of adoptions for the two generations

$$Y_1(t) = m_1F_1(t) - m_1\int_{\tau_2}^{t} f_1(\theta)F_2(\theta-\tau_2)d\theta \quad \text{(A5)}$$

$$Y_2(t) = [m_2 + m_1F_1(t)]F_2(t-\tau_2) \quad \text{(A6)}$$

Proof of Proposition 1

The derivative of the objective function of (19) with respect to τ_2 equals

$$\frac{d\pi(\tau_2)}{d\tau_2} = \frac{d}{d\tau_2}\left\{\pi_1\left[m_1F(D) - m_1\int_{\tau_2}^{D} f(\theta)F(\theta-\tau_2)d\theta\right] + \pi_2[m_2 + m_1F(D)]F(D-\tau_2)\right\}$$

$$= \frac{d}{d\tau_2}\{\pi_1 m_1 F(D)\} - \frac{d}{d\tau_2}\left\{\pi_1 m_1 \int_{\tau_2}^{D} f(\theta)F(\theta-\tau_2)d\theta\right\} + \frac{d}{d\tau_2}\{\pi_2[m_2 + m_1F(D)]F(D-\tau_2)\}$$
$$= 0 - \pi_1 m_1 \frac{d}{d\tau_2}\left\{\int_{\tau_2}^{D} f(\theta)F(\theta-\tau_2)d\theta\right\} - \pi_2[m_2 + m_1F(D)]f(D-\tau_2)$$

Note that that the last term is obtained based on Equation (7) or equivalently $\frac{dF(t)}{dt} = f(t)$.

The derivative of integral term can be derived based on the Leibniz's rule:

$$\frac{d}{d\tau_2}\left\{\int_{\tau_2}^{D} f(\theta)F(\theta-\tau_2)d\theta\right\} = -f(\tau_2)F(\tau_2-\tau_2) + \int_{\tau_2}^{D} \frac{d}{d\tau_2}[f(\theta)F(\theta-\tau_2)]d\theta$$
$$= -0 - \int_{\tau_2}^{D} f(\theta)f(\theta-\tau_2)d\theta$$

The first term equals 0 because $F(\tau_2 - \tau_2) = F(0) = 0$.

Therefore,

$$\frac{d\pi(\tau_2)}{d\tau_2} = \pi_1 m_1 \int_{\tau_2}^{D} f(\theta)f(\theta-\tau_2)d\theta - \pi_2[m_2 + m_1F(D)]f(D-\tau_2) \qquad \text{(A7)}$$

From (A7), we have

$$\frac{d\pi(\tau_2)}{d\tau_2} < \pi_1 m_1 \left[\max_{0\le\theta\le D} f(\theta)\right]^2 (D-\tau_2) - \pi_2[m_2 + m_1F(D)]f(D-\tau_2)$$
$$< \pi_1 m_1 \left[\max_{0\le\theta\le D} f(\theta)\right]^2 (D-\tau_2) - \pi_2[m_2 + m_1F(D)]\min_{0\le\theta\le D} f(\theta)$$
$$= \pi_1 m_1 \left[\max_{0\le\theta\le D} f(\theta)\right]^2 \left[(D-\tau_2) - \frac{\pi_2[m_2 + m_1F(D)]\min_{0\le\theta\le D} f(\theta)}{\pi_1 m_1 \left[\max_{0\le\theta\le D} f(\theta)\right]^2}\right]$$

Let

$$\varphi = \frac{\pi_2[m_2 + m_1F(D)]\min_{0\le\theta\le D} f(\theta)}{\pi_1 m_1 \left[\max_{0\le\theta\le D} f(\theta)\right]^2}$$

then we have

$$\frac{d\pi(\tau_2)}{d\tau_2} < \pi_1 m_1 \left[\max_{0\le\theta\le D} f(\theta)\right]^2 [(D-\varphi) - \tau_2]$$

Therefore,

$$\frac{d\pi(\tau_2)}{d\tau_2} < 0, \text{if } (D-\varphi) \le \tau_2 \le D$$

Given that $\pi(\tau_2)$ is monotonically decreasing with $\tau_2 \in [D-\varphi, D]$, we conclude that

$\tau_2^* < (D-\varphi)$, if $(D-\varphi) > \alpha_2$ or equivalently, $\varphi < D - \alpha_2$

and $\tau_2^* = \alpha_2$, if $(D-\varphi) \le \alpha_2$ or equivalently, $\varphi \ge D - \alpha_2$

In either case, $\tau_2^* < D$ holds. Therefore, not introducing the second generation during the planning horizon cannot be optimal.

Proof of Proposition 2

To prove that G2 should be introduced as early as possible, we only need to show

$$\frac{d\pi(\tau_2)}{d\tau_2} < 0, \forall \tau_2 \in [0, D]$$

where $\frac{d\pi(\tau_2)}{d\tau_2}$ has the same expression shown in (A7). First, with $\theta \in [\tau_2, D]$, we have

$$0 \le (\theta - \tau_2) \le (D - \tau_2) \le D \le T^* = \ln(q/p)/(p+q)$$

Since $f(t)$ is a monotonically increasing function of t before its peak is reached, i.e., when $t \le T^*$, we have

$$f(\theta - \tau_2) \le f(D - \tau_2) \Rightarrow$$

$$\int_{\tau_2}^{D} f(\theta) f(\theta - \tau_2) d\theta < \int_{\tau_2}^{D} f(\theta) f(D - \tau_2) d\theta = [F(D) - F(\tau_2)] f(D - \tau_2) \le F(D) f(D - \tau_2)$$

From $\pi_1 \le \pi_2$, we further conclude

$$\pi_1 m_1 \int_{\tau_2}^{D} f(\theta) f(\theta - \tau_2) d\theta < \pi_2 m_1 F(D) f(D - \tau_2)$$

which leads to

$$\frac{d\pi(\tau_2)}{d\tau_2} = \pi_1 m_1 \int_{\tau_2}^{D} f(\theta) f(\theta - \tau_2) d\theta - \pi_2 [m_2 + m_1 F(D)] f(D - \tau_2) < 0, \qquad \forall \tau_2 \in [0, D]$$

Derivation of Optimal Market Entry Timing for PTO Products under Total Transition

First, the objective function of (21) can be reorganized as

$$\pi(\tau_2) = \pi_2 m_1 F(D) + (\pi_1 - \pi_2) m_1 F(\tau_2) + \pi_2 [m_2 + m_1 F(\tau_2)] F(D - \tau_2)$$

Taking the derivative of the above function with respect to τ_2 yields

$$\frac{\partial \pi(\tau_2)}{\partial \tau_2} = (\pi_1 - \pi_2) m_1 f(\tau_2) - \pi_2 m_2 f(D - \tau_2) + \pi_2 m_1 f(\tau_2) F(D - \tau_2) - \pi_2 m_1 F(\tau_2) f(D - \tau_2) \tag{A8}$$

After substituting $F(\cdot)$ and $f(\cdot)$, letting $x = e^{(p+q)\tau_2}$ and $\delta = e^{-(p+q)D}$, and some additional algebraic rearrangement (details available from the authors), the above derivative can be expressed as

$$\frac{d\pi(\tau_2)}{d\tau_2} = H(x)(ax^2 + bx + c) \tag{A9}$$

where $H(x) = -\frac{(p+q)^2 x}{p[(q/p)+x]^2[(q/p)\delta x + 1]^2}$

$a = (\pi_2 - \pi_1) m_1 \delta^2 \left(\frac{q}{p}\right)^2 + \pi_2 m_2 \delta + \pi_2 m_1 \delta^2 \frac{q}{p} + \pi_2 m_1 \delta$

$b = 2(\pi_2 - \pi_1) m_1 \delta \frac{q}{p} + 2\pi_2 m_2 \delta \frac{q}{p}$

and

$c = \pi_2 m_2 \delta \left(\frac{q}{p}\right)^2 - \pi_1 m_1 - \pi_2 m_1 \delta \frac{q}{p}$

We next take a closer look at the terms in (A9). Since $x = e^{(p+q)\tau_2} \ge 1$, we must have

$H(x) < 0$

We present below the solution for the most likely scenario, i.e., $a > 0$ and $b > 0$.[1]

We first examine the first-order condition $\frac{d\pi(\tau_2)}{d\tau_2} = 0$. Since $H(x)$ in (A9) is always negative, we conclude that the first-order condition requires

$$ax^2 + bx + c = 0 \qquad \text{(A10)}$$

Note that here $x = e^{(p+q)\tau_2} \geq 1$ for $\forall\tau_2 \geq 0$.

We examine three conditions, based on how the value of c compares with the other parameters.

(1) $\boldsymbol{c > b^2/(4a)}$. This condition implies $b^2 - 4ac < 0$. Under this scenario, (A10) has no real solution. From $a > 0$, we have

$$ax^2 + bx + c = 0 > 0,\ \forall x$$

Since $H(x) < 0$, we conclude

$$\frac{d\pi(\tau_2)}{d\tau_2} < 0,\ \forall\tau_2 \geq 0$$

Because the net profit decreases monotonically as the introduction of G2 is delayed, it is optimal to introduce generation 2 as early as possible, that is,

$$\tau_2^* = \alpha_2$$

(2) $\boldsymbol{0 \leq c \leq b^2/(4a)}$. Under this scenario, $b^2 - 4ac \geq 0$, hence (A10) has a real-number solution:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Since $c \geq 0$, we must have $(b^2 - 4ac) \leq b^2$. Under this condition, the root(s) of (A10) are non-positive. Hence,

$$ax^2 + bx + c = 0 > 0,\ \forall x \geq 1$$

$$\Rightarrow \frac{d\pi(\tau_2)}{d\tau_2} < 0,\ \forall\tau_2 \geq 0$$

Therefore,

$$\tau_2^* = \alpha_2$$

Based on scenarios (1) and (2), we conclude

$$c \geq 0 \Rightarrow \tau_2^* = \alpha_2$$

(3) $\boldsymbol{c} < 0$. This condition leads to $b^2 - 4ac \geq b^2$. In this case, the two roots of (A10) are

$$\begin{cases} x_1 = \dfrac{-b - \sqrt{b^2 - 4ac}}{2a} < 0 \\ x_2 = \dfrac{-b + \sqrt{b^2 - 4ac}}{2a} > 0 \end{cases}$$

In this case, if $x_2 \leq e^{(p+q)\alpha_2}$, or equivalently,

$$c \geq -ae^{2(p+q)\alpha_2} - be^{(p+q)\alpha_2}$$

[1]For most products, we expect the unit contribution margin for G2 to be at least as high as that for G1 (i.e., $\pi_2 \geq \pi_1$), hence $a > 0$ and $b > 0$ must hold. Even if π_2 is slightly lower than π_1, numerical analysis based on common parameter values show both a and b are likely to remain positive.

we still have

$$\frac{d\pi(\tau_2)}{d\tau_2} \leq 0, \forall x \geq e^{(p+q)\alpha_2}, \text{ or } \forall \tau_2 \geq \alpha_2$$

Therefore,

$$\tau_2^* = \alpha_2$$

On the other hand, if $x_2 > e^{(p+q)\alpha_2}$, or equivalently,

$$c < -ae^{2(p+q)\alpha_2} - be^{(p+q)\alpha_2}$$

we have

$$\frac{d\pi(\tau_2)}{d\tau_2} \geq 0, \forall x \leq x_2$$

And

$$\frac{d\pi(\tau_2)}{d\tau_2} < 0, \forall x > x_2$$

Hence,

$$\tau_2^* = \frac{\text{Ln}(x_2)}{p+q}$$

or

$$\tau_2^* = \text{Ln}(\frac{-b+\sqrt{b^2-4ac}}{2a})/(p+q)$$

Optimal Solution: Taking into consideration all three scenarios, the optimal solution for Scenario II is

$$\tau_2^* = \begin{cases} \text{Ln}\left(\dfrac{-b+\sqrt{b^2-4ac}}{2a}\right)/(p+q), & \text{if } c < -ae^{2(p+q)\alpha_2} - be^{(p+q)\alpha_2} \\ \alpha_2, & \text{otherwise} \end{cases}$$

Proof of Proposition 3

(1). We first prove that with $\pi_1 \leq \pi_2$, we have $\tau_2^* < D$.

Again, taking the derivative of the objective function of (21) with respect to τ_2 yields the results shown in (A8):

$$\frac{\partial\pi(\tau_2)}{\partial\tau_2} = (\pi_1 - \pi_2)m_1 f(\tau_2) - \pi_2 m_2 f(D-\tau_2) + \pi_2 m_1 f(\tau_2)F(D-\tau_2) - \pi_2 m_1 F(\tau_2) f(D-\tau_2)$$

It is easy to show that the sum of first two terms is less than zero, that is,

$$(\pi_1 - \pi_2)m_1 f(\tau_2) - \pi_2 m_2 f(D-\tau_2) < 0 \tag{A11}$$

Now let us take a look at the third and fourth terms.

$$\begin{aligned} &\pi_2 m_1 f(\tau_2)F(D-\tau_2) - \pi_2 m_1 F(\tau_2) f(D-\tau_2) \\ &= \pi_2 m_1 [f(\tau_2)F(D-\tau_2) - F(\tau_2) f(D-\tau_2)] \end{aligned} \tag{A12}$$

In (A12), as $\tau_2 \to \infty$, $F(D-\tau_2) \to 0$, while all other terms remain well above zero. As a result, (A12) necessarily becomes negative. This result, together with (A11), lead to the conclusion that

$$\frac{\partial \pi(\tau_2)}{\partial \tau_2} < 0, \text{ as } \tau_2 \to \infty$$

Therefore, we must have $\tau_2^* < D$.

(2). We now prove that with $\pi_1 > \pi_2$, it is possible to have $\tau_2^* > D$, implying that not introducing the second generation during the planning horizon could be an optimal solution.

Again, by looking at the expression in (A8), it can be shown that regardless of the value of the other parameters, if π_1 is sufficiently large, we can have $\frac{\partial \pi(\tau_2)}{\partial \tau_2} > 0,\ \forall \tau_2 \in [0, D]$. Then, not introducing the second generation is indeed an optimal solution.

Proof of Proposition 4

We separately examine Scenarios III and IV.

(1) Under Scenario III (STU products and phase-out transition), the objective function of (24) can be rearranged to

$$\pi(\tau_2) = \varphi_1 m_1 \int_0^D F(\theta)d\theta + (\varphi_2 - \varphi_1)m_1 \int_{\tau_2}^D F(\theta)F(\theta - \tau_2)d\theta + \varphi_2 m_2 \int_{\tau_2}^D F(\theta - \tau_2)d\theta.$$

Therefore,

$$\begin{aligned}\frac{d\pi(\tau_2)}{d\tau_2} &= (\varphi_2 - \varphi_1)m_1 \left[-F(\tau_2)F(\tau_2 - \tau_2) - \int_{\tau_2}^D F(\theta)f(\theta - \tau_2)d\theta\right] + \varphi_2 m_2 \left[-F(\tau_2 - \tau_2) - \int_{\tau_2}^D f(\theta - \tau_2)d\theta\right] \\ &= -(\varphi_2 - \varphi_1)m_1 \textstyle\int_{\tau_2}^D F(\theta)f(\theta - \tau_2)d\theta - \varphi_2 m_2 F(D - \tau_2)\end{aligned} \tag{A13}$$

If $\varphi_1 \le \varphi_2$, we have

$$\begin{cases} \dfrac{d\pi(\tau_2)}{d\tau_2} < 0, & if\ \tau_2 < D \\ \dfrac{d\pi(\tau_2)}{d\tau_2} = 0, & if\ \tau_2 = D \end{cases}$$

Since $\pi(\tau_2)$ decreases monotonically as τ_2 increases, G2 should be introduced to the market as early as possible, that is, $\tau_2^* = \alpha_2$.

(2) Under Scenario IV (STU products and total transition), the objective function of (25) can be rearranged to

$$\begin{aligned}\pi(\tau_2) &= \varphi_1 m_1 \int_0^{\tau_2} F(\theta)d\theta + \varphi_1 \int_{\tau_2}^D m_1 F(\tau_2)[1 - F(\theta - \tau_2)]d\theta + \varphi_2 \int_{\tau_2}^D \{[m_2 + m_1 F(\tau_2)]F(\theta - \tau_2) + m_1[F(\theta) - F(\tau_2)]\}d\theta \\ &= \varphi_1 m_1 \int_0^D F(\theta)d\theta - \varphi_1 m_1 \int_{\tau_2}^D F(\theta)d\theta + \int_{\tau_2}^D \{(\varphi_1 - \varphi_2)m_1 F(\tau_2) + \varphi_2 m_2 F(\theta - \tau_2) \\ &\quad +(\varphi_2 - \varphi_1)m_1 F(\tau_2)F(\theta - \tau_2) + \varphi_2 m_1 F(\theta)\}d\theta \\ &= \varphi_1 m_1 \int_0^D F(\theta)d\theta + \int_{\tau_2}^D \{(\varphi_1 - \varphi_2)m_1 F(\tau_2) + \varphi_2 m_2 F(\theta - \tau_2) \\ &\quad +(\varphi_2 - \varphi_1)m_1 F(\tau_2)F(\theta - \tau_2) + (\varphi_2 - \varphi_1)m_1 F(\theta)\}d\theta\end{aligned}$$

Then,

$$\begin{aligned}\frac{d\pi(\tau_2)}{d\tau_2} &= -\{(\varphi_1 - \varphi_2)m_1 F(\tau_2) + \varphi_2 m_2 F(\tau_2 - \tau_2) + (\varphi_2 - \varphi_1)m_1 F(\tau_2)F(\tau_2 - \tau_2) + (\varphi_2 - \varphi_1)m_1 F(\tau_2)\} \\ &\quad + \int_{\tau_2}^D \{(\varphi_1 - \varphi_2)m_1 f(\tau_2) - \varphi_2 m_2 f(\theta - \tau_2) \\ &\quad -(\varphi_2 - \varphi_1)m_1 F(\tau_2)f(\theta - \tau_2) + (\varphi_2 - \varphi_1)m_1 f(\tau_2)F(\theta - \tau_2)\}d\theta \\ &= (\varphi_1 - \varphi_2)m_1 f(\tau_2)(D - \tau_2) - \varphi_2 m_2 F(D - \tau_2) - (\varphi_2 - \varphi_1)m_1 F(\tau_2)F(D - \tau_2) +\end{aligned}$$

$$(\varphi_2 - \varphi_1)m_1 f(\tau_2)\int_{\tau_2}^{D} F(\theta - \tau_2)d\theta \tag{A14}$$

If $\varphi_1 \leq \varphi_2$, we have

$$\begin{gathered}\frac{d\pi(\tau_2)}{d\tau_2} < (\varphi_1 - \varphi_2)m_1 f(\tau_2)(D - \tau_2) - \varphi_2 m_2 F(D - \tau_2) - (\varphi_2 - \varphi_1)m_1 F(\tau_2)F(D - \tau_2) + (\varphi_2 - \varphi_1)m_1 f(\tau_2)\int_{\tau_2}^{D} d\theta \\ = (\varphi_1 - \varphi_2)m_1 f(\tau_2)(D - \tau_2) - \varphi_2 m_2 F(D - \tau_2) - (\varphi_2 - \varphi_1)m_1 F(\tau_2)F(D - \tau_2) + (\varphi_2 - \varphi_1)m_1 f(\tau_2)(D - \tau_2) \\ = -\varphi_2 m_2 F(D - \tau_2) - (\varphi_2 - \varphi_1)m_1 F(\tau_2)F(D - \tau_2)\end{gathered}$$

Hence,

$$\begin{cases}\dfrac{d\pi(\tau_2)}{d\tau_2} < 0, & if\ \tau_2 < D \\ \dfrac{d\pi(\tau_2)}{d\tau_2} = 0, & if\ \tau_2 = D\end{cases}$$

Therefore, it is optimal to release G2 as early as possible, that is, $\tau_2^* = \alpha_2$.

Proof of Proposition 5

We separately examine Scenarios III and IV.

(1) Under Scenario III (STU products and phase-out transition), if $\varphi_1 > \varphi_2$, from (A13) we have

$$\begin{aligned}\frac{d\pi(\tau_2)}{d\tau_2} &= (\varphi_1 - \varphi_2)m_1\int_{\tau_2}^{D} F(\theta)f(\theta - \tau_2)d\theta - \varphi_2 m_2 F(D - \tau_2) \\ &< (\varphi_1 - \varphi_2)m_1\int_{\tau_2}^{D} F(D)f(\theta - \tau_2)d\theta - \varphi_2 m_2 F(D - \tau_2) \\ &= [(\varphi_1 - \varphi_2)m_1 F(D) - \varphi_2 m_2]F(D - \tau_2)\end{aligned}$$

Therefore,

$$\frac{d\pi(\tau_2)}{d\tau_2} < 0 \text{ if } (\varphi_1 - \varphi_2)m_1 F(D) \leq \varphi_2 m_2$$

or equivalently,

$$\frac{d\pi(\tau_2)}{d\tau_2} < 0 \text{ if } m_2 \geq \frac{(\varphi_1 - \varphi_2)F(D)}{\varphi_2}m_1$$

Thus

$$\tau_2^* = \alpha_2 \quad \text{if } m_2 \geq \frac{(\varphi_1 - \varphi_2)F(D)}{\varphi_2}m_1$$

(2) Under Scenario IV (STU products and total transition), if $\varphi_1 > \varphi_2$, from (A14) we have

$$\begin{gathered}\frac{d\pi(\tau_2)}{d\tau_2} = -\varphi_2 m_2 F(D - \tau_2) - (\varphi_2 - \varphi_1)m_1 F(\tau_2)F(D - \tau_2) + (\varphi_1 - \varphi_2)m_1 f(\tau_2)\int_{\tau_2}^{D}[1 - F(\theta - \tau_2)]d\theta \\ \frac{d\pi(\tau_2)}{d\tau_2} \leq 0 \Leftrightarrow -\varphi_2 m_2 F(D - \tau_2) + (\varphi_1 - \varphi_2)m_1 F(\tau_2)F(D - \tau_2) + (\varphi_1 - \varphi_2)m_1 f(\tau_2)\int_{\tau_2}^{D}[1 - F(\theta - \tau_2)]d\theta \leq 0 \\ \Leftrightarrow \varphi_2 m_2 F(D - \tau_2) \geq (\varphi_1 - \varphi_2)m_1 F(\tau_2)F(D - \tau_2) + (\varphi_1 - \varphi_2)m_1 f(\tau_2)\int_{\tau_2}^{D}[1 - F(\theta - \tau_2)]d\theta\end{gathered}$$

$$\Leftrightarrow m_2 \geq \frac{(\varphi_1 - \varphi_2)m_1}{\varphi_2}\left[F(\tau_2) + \frac{f(\tau_2)}{F(D-\tau_2)}\int_{\tau_2}^{D}[1 - F(\theta - \tau_2)]d\theta\right]$$

The second term on the RHS of the equation is always positive, that is,

$$F(\tau_2) + \frac{f(\tau_2)}{F(D-\tau_2)}\int_{\tau_2}^{D}[1 - F(\theta - \tau_2)]d\theta > 0, \forall \tau_2 \in [\alpha_2, D]$$

We let

$$\nu = \max_{0 \leq \tau_2 \leq D} f\left[F(\tau_2) + \frac{f(\tau_2)}{F(D-\tau_2)}\int_{\tau_2}^{D}[1 - F(\theta - \tau_2)]d\theta\right]$$

It can be shown that

$$\frac{d\pi(\tau_2)}{d\tau_2} \leq 0, \qquad \text{if } m_2 \geq \frac{(\varphi_1 - \varphi_2)m_1}{\varphi_2}\nu$$

Extension Model with Network Effects, Compatibility, and Switching Cost Considered

In the main text we have implicitly assumed that network effects, product compatibility, and switching cost are either exogenous or negligible. Here, we develop an extension model to examine how these additional economic factors can affect the new generation's optimal market entry timing and the total profit. We consider only PTO products under phase-out transition in the following analysis.

Network effects refer to the phenomenon that the utility a consumer receives from using a product increases with the number of users of that product (Katz and Shapiro 1985). In the presence of two product generations, it is possible that the valuation of one generation can benefit from its own network as well as the other generation's network. In other words, there exist both *within-generation network effects* and *cross-generation network effects*. Cross-generation network effects exist because of backward and forward compatibility. Specifically, *backward (forward) compatibility* allows the new (old) product generation to benefit from the old (new) product generation's network (Choi 1994).

To differentiate the multiple sources of network effects, we denote the intensity of within-generation network effects (i.e., the incremental utility resulting from one more users joining the focal generation's network) by α, the intensity of forward compatibility by β_f, and the intensity of backward network effects by β_b. Because forward compatibility is more difficult to achieve and backward compatibility is more critical to the success of a product line, firms typically consider backward compatibility to be of higher strategic importance (Choi 1994, Kretschmer and Claussen 2016). Hence we assume that the intensity of forward compatibility is no higher than that of backward compatibility. In addition, the intensity of cross-generation network effects cannot be larger than that of within-generation network effects. Given these two assumptions, we have $0 \leq \beta_f \leq \beta_b \leq \alpha$. The network size of each generation at time *t* is simply the number of units-in-use at that time, that is, $S_1(t)$ or $(S_2(t)$.

To capture the impact of these economic factors, it is necessary to consider the value of each product generation to potential adopters. First, the new product generation is expected to have a higher quality than that of the older generation. We denote the product quality of G1 (G2), measured by its value to adopters, by δ_1 (δ_2).[2] With the network effects induced benefits considered, the value of G1 at time *t* can be expressed as

$$V_1(t) = \delta_1 + \alpha S_1(t) + \beta_f S_2(t) \tag{A15}$$

Similarly, the value of G2 at time *t* take the form

$$V_2(t) = \delta_2 + \alpha S_2(t) + \beta_b S_1(t) \tag{A16}$$

Therefore, the difference in value between the two product generations is

$$\Delta V(t) = V_1(t) - V_2(t) = (\delta_1 - \delta_2) + \alpha[S_1(t) - S_2(t)] + \beta_f S_2(t) - \beta_b S_1(t) \tag{A17}$$

[2]For simplicity, we assume the price difference is negligible. In case the difference is significant, we can redefine product quality to capture this difference, and the rest of the analysis and findings remain valid.

For potential adopters of G1 who have not adopted G1 yet, they have the options of adopting either G1 or G2 after the latter is released. Using the widely adopted logit choice model (Guadagni and Little 1983, Morgan et al. 2001), the percentage of potential adopters who choose G2 instead of G1 at time *t*, which is essentially the *leapfrogging multiplier*, equals

$$u_2(t) = \frac{e^{V_2(t)}}{e^{V_2(t)} + e^{V_1(t)}} = \frac{1}{1+e^{\Delta V(t)}},\ t \geq \tau_2 \tag{A18}$$

With respect to those customers who have already adopted G1, the switching cost, denoted by ω, comes into play. As a result, the percentage of them who are willing to switch to G2 at time *t*, i.e., the *switching multiplier*, takes the form

$$w_2(t) = \frac{e^{V_2(t)-\omega}}{e^{V_2(t)-\omega} + e^{V_1(t)}} = \frac{1}{1+e^{\Delta V(t)+\omega}},\ t \geq \tau_2 \tag{A19}$$

With the leapfrogging and switching multipliers defined, we can revise the GNB model to capture the effect of network effects, backward-forward compatibility, and switching cost on the diffusion of the two product generations. Because of the complex dependencies among leapfrogging/switching, units-in-use, and adoptions, it is not possible to develop a continuous-time multigeneration diffusion model. Therefore, we revise a discrete version of the GNB model based on the leapfrogging and switching multipliers shown in Equations (A18) and (A19).

Under a discrete time model, we consider discrete time instances $t \in \{0, 1, 2, \ldots, \tau_2, \ldots, D\}$, where τ_2 still denotes the release time of G2 and the length of planning horizon is still *D*. The noncumulative adoption rate $y_G(t)$ for generation *G* is now interpreted as the number of adoptions occurring during time interval (*t*–1, *t*]. $S_G(t)$ and $Y_G(t)$ represent the number of units-in-use and the cumulative number of adoptions, respectively, for generation *G* at time *t*. Based on these definitions, prior to the release of G2 at time τ_2, the adoptions of G1 can be characterized by

$$y_1(t) = m_1[F_1(t) - F_1(t-1)],\ \ 1 \leq t \leq \tau_2 \tag{A20}$$

$$S_1(t) = Y_1(t) = m_1 F_1(\tau_2),\ 0 \leq t \leq \tau_2 \tag{A21}$$

After the release of G2, due to leapfrogging and switching, the two functions change to

$$y_1(t) = m_1[F_1(t) - F_1(t-1)][1 - u_2(t)],\ t \geq \tau_2 + 1 \tag{A22}$$

$$S_1(t) = S_1(t-1)[1 - w_2(t)]\ + m_1[F_1(t) - F_1(t-1)]\{1 - u_2(t)\},\ t \geq \tau_2 + 1 \tag{A23}$$

In addition, the following equation holds throughout the planning horizon:

$$Y_1(t) = Y_1(t-1) + y_1(t),\ \ t \geq 1 \tag{A24}$$

For the new generation G2, we have

$$\begin{aligned} &y_2(t) = m_2[F_2(t-\tau_2) - F_2(t-\tau_2-1)] \\ &+ S_1(t-1)w_2(t) + m_1[F_1(t) - F_1(t-1)]u_2(t),\ t \geq \tau_2 + 1 \end{aligned} \tag{A25}$$

$$S_2(\tau_2) = Y_2(\tau_2) = 0,\ t = \tau_2 \tag{A26}$$

$$S_2(t) = Y_2(t) = Y_2(t-1) + y_2(t),\ t \geq \tau_2 + 1 \tag{A27}$$

Equations (A15) – (A27) jointly constitute the *discrete-time diffusion model with network effects*. We refer to it as the DNE model in the rest of the discussion.

Based on the DNE model, the optimal entry time for G2 can be formulated as

$$\underset{\alpha_2 \leq \tau_2 \leq D}{\text{MAX}}\ \pi(\tau_2) = \pi_1 \sum_{t=1}^{D} y_1(t) + \pi_2 \sum_{t=\tau_2}^{D} y_2(t) \tag{A28}$$

Given the complexity of the model, analytical findings are unattainable. Hence, similar to prior studies (e.g., Mahajan and Muller 1996, Koca, Souza, and Druehl 2010, and Joshi, Reibstein, and Zhang 2009), we conduct numerical analysis to examine how the network effects, backward-forward compatibility, and switching cost affect the new generation's optimal market entry timing and the total profit. We retain some parameter values used in the previous numerical analysis as default values, i.e., $m_1 = m_2 = 10$ million, $\pi_1 = \pi_2 = \$100$, and D=10 years (or 120 months). The other default values are set at $p = 0.02$, $q = 0.2$, $\delta_1 - \delta_2 = -10$; α=2, $\beta_f = \beta_b = 1$, and ω=20.

Within-Generation Network Effects

We first examine how the intensity of within-generation network (α) affects the optimal market entry timing of G2 and the firm's profitability. From Equation (A17), we can tell that when the two product generations are completely compatible, i.e., $\alpha = \beta_f = \beta_b$, having a larger network size of its own does not help a particular product generation, because the other generation can benefit equally from its larger network size. In fact, the larger is the difference between α and $\beta_f(\beta_b)$, the more advantageous it is to have a large network size. For this reason, in our first analysis, we change the value of α while keeping the values of β_f, β_b, and other parameters fixed at their default values. As shown in Figure A1(a), as the value of α increases, the firm's total profit exhibits a monotonically increasing pattern, showing that the firm can benefit directly from stronger within-generation network effects.

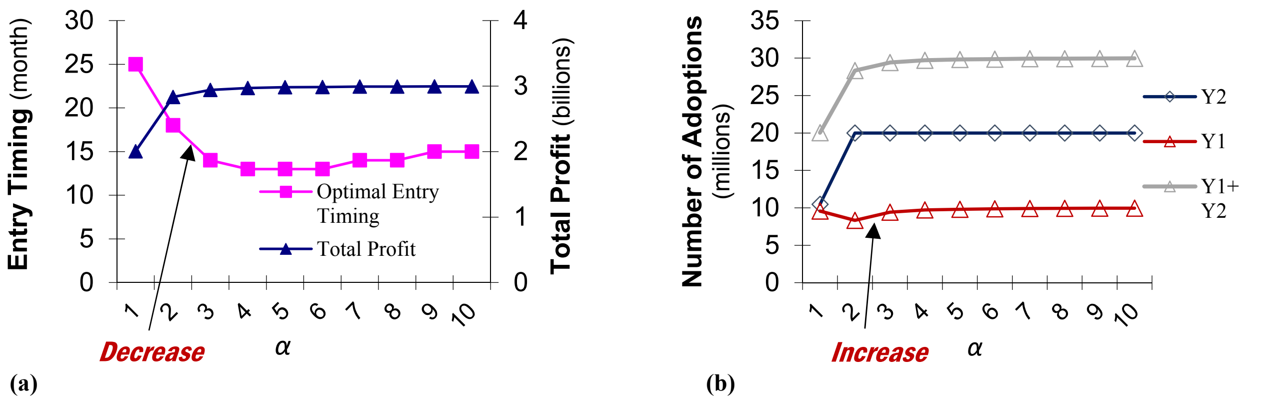


Figure A1. Impact of Intensity of Within-Generation Network Effects (α)

What is more interesting is that the optimal entry timing shows a non-monotonic pattern, which first decreases (shorter time to market) and then increases (longer time-to-market) after a bottom is reached. In order to better understand this interesting pattern, we plot the numbers of adoptions of G1 and G2, as well as the total number of adoptions, in Figure A1(b). From this figure, we can see that Y_1, denoting the total number of adopters of G1, drops when the value of α changes from 1 to 2; this is because, as shown in Figure A1(a), G2 enters the market earlier, hence taking away more potential adopters from G1. This reduction in the number of G1 adoptions, however, is more than compensated by the sharp increase in the number of G2 adoptions, which is evident in Figure A1(b). In sum, although the number G1 adoptions drops as G2 enter the market earlier, the total number of adoptions of G1 and G2 increases, leading to a higher profit.

However, as α changes from 2 to 3, the optimal market entry timing of G2 and Y_1 are no longer in sync, and the results are not expected — the number of G1 adoptions increases although G2 enters the market earlier. Upon further examination, we find that this is a result of within-network effects. Specifically, with a higher intensity of within-network effect, the proportion of potential adopters who are willing to leapfrog from G1 to G2 becomes smaller; this is because upon release G2 has a much smaller network size than G1. In other words, the cannibalization of G1 adoptions by the earlier release of G2 will be limited if the within-generation network effects are sufficiently high.

As show in Figure A1(a), the optimal market entry timing reaches the bottom when α equals 4, 5, 6. We next explain why it is optimal to delay the release of G2 as α increases beyond 6. It is worth noting that with a constant market size, the key to improve the total profit is to increase the number of cross-generation repeat adoptions, i.e., the number of adopters who buy both G1 and G2. As explained earlier, a large within-generation network effects, resulting from a large network size of G1 in relation to that of G2, can decrease the canalization effect. Therefore, delaying the release of G2 can help increase the number of G1 adopters in two ways. First, it increases the number of G1 adoptions before the release. Second, a larger network size upon the release of G2 can further reduce the canalization effect. Since the vast majority of G1 adopters will eventually switch to G2, a large number of G1 adoptions implies more cross-generation repeat adoptions, and hence a higher total profit.

Backward/Forward Compatibility

We also conduct numerical analysis to examine the impact of backward and forward compatibility on the optimal market entry timing and the total profit. To allow a large range of parameter values, we set the intensity of with-generation network effects to $\alpha = 10$. Then, with the intensity of forward-compatibility (β_f) fixed at 0, we vary the intensity of backward-compatibility (β_b) from 0 to 10, and record the results in Figure A2. Subsequently, we fix β_b at 8, and vary the value of β_f from 0 to 8; the results are summarized in Figure A3. The breakdown of the adoptions of the two generations are not shown because they add limited additional insight. From Figure A2, we can tell that with a higher backward compatibility, a firm can delay the release of the second generation. The profit increases dramatically as β_b increases from 0 to 1, but remains relatively flat afterward. The impact of forward compatibility is almost the opposite. The new generation should be released earlier with a higher forward compatibility. The profit remains little changed as β_f changes from 0 to 7, but drops sharply when β_f=7=β_b. Considering that fact that profit is also at the lowest when in Figure A3 when β_b=0=β_f, we conclude that it is less profitable when the intensities of forward and backward compatibility are equal.

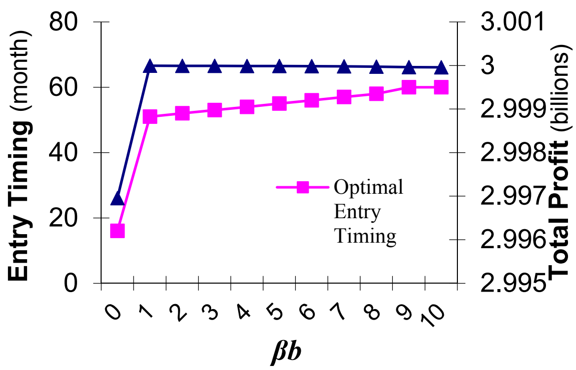


Figure A2. Impact of Intensity of Backward Compatibility (β_b)

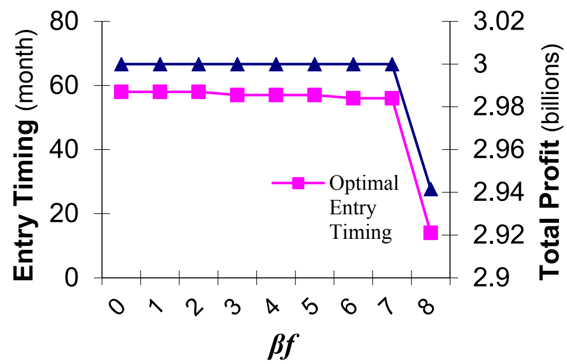
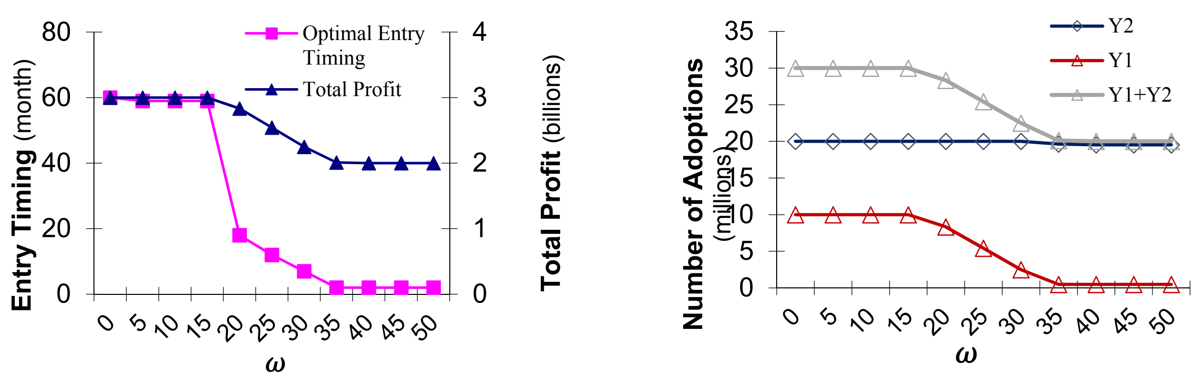


Figure A3. Impact of Intensity of Forward Compatibility (β_f)

Switching Cost

We also examine how the switching cost affects the new generation's optimal market entry timing and the firm's profitability. As shown in Figure A4(a), a larger switching cost leads to a lower profit and an earlier optimal market entry timing for G2. From Figure A4(b), we can see that the number of G2 adoptions remains little changed, while the number of G1 adoptions drops significantly. Our explanations for this trend is as follows. A higher switching cost reduces the attractiveness of the new generation to those who have already adopted G1. As a result, delaying the release of G2 to increase the number of cross-generation repeat purchases becomes less justifiable. Although releasing G2 earlier also leads to more cannibalization, such cannibalization is not a concern because the profit per adoption is the same for G1 and G2. With both factors considered, it makes sense for the firm to release G2 earlier.



(a) (b)

Figure A4. Impact of Switching Cost (ω)

In sum, within-generation network effects, backward-forward compatibility, and switching cost all have influence on the new generation's optimal market entry timing and the firm's total profit. While the directions of their impact on the total profit are as generally expected, their impact on market entry timing are not as straightforward.

References

Choi, J. P. 1994. "Network Externality, Compatibility Choice, and Planned Cbsolescence," *The Journal of Industrial Economics* (42:2), pp. 167-182.

Guadagni, P. M., and Little, J. D. 1983. "A Logit Model of Brand Choice Calibrated on Scanner Data," *Marketing Science* (2:3), pp. 203-238.

Jiang, Z., and Jain, D. C. 2012. "A Generalized Norton-Bass Model for Multigeneration Diffusion," *Management Science* (58:10), pp. 1887-1897.

Joshi, Y. V., Reibstein, D. J., and Zhang, Z. J. 2009. "Optimal Entry Timing in Markets with Social Influence," *Management Science* (55:6), pp. 926-939.

Katz, M. L. and., Shapiro, C. 1985. "Network Externalities, Competition, and Compatibility," *The American Economic Review* (75:3) pp. 424-440.

Koca, E., Souza, G. C., and Druehl, C. T. 2010. "Managing Product Rollovers," *Decision Sciences* (41:2), pp. 403-423.

Kretschmer, T., and Claussen, J. 2016. "Generational Transitions in Platform Markets—The Role of Backward Compatibility," *Strategy Science* (1:2), pp. 90-104.

Mahajan V., and Muller, E. 1996. "Timing, Diffusion, and Substitution of Successive Generations of Technological Innovations: The IBM Mainframe Case," *Technological Forecasting and Social Change* (51:2), pp. 109-132.

Morgan, L. O., Morgan, R. M., and Moore, W. L. 2001. "Quality and Time-to-Market Trade-offs When There Are Multiple Product Generations," *Manufacturing & Service Operations Management* (3:2), pp. 89-104.