Discussion Papers In Economics And Business

Free Entry, Market Diffusion, and Social Inefficiency with Endogenously Growing Demand

Hiroshi Kitamura, Akira Miyaoka and Misato Sato

Discussion Paper 11-04-Rev.

Graduate School of Economics and
Osaka School of International Public Policy (OSIPP)
Osaka University, Toyonaka, Osaka 560-0043, JAPAN

Free Entry, Market Diffusion, and Social Inefficiency with Endogenously Growing Demand

Hiroshi Kitamura, Akira Miyaoka and Misato Sato

Discussion Paper 11-04-Rev.

November 2012

Graduate School of Economics and
Osaka School of International Public Policy (OSIPP)
Osaka University, Toyonaka, Osaka 560-0043, JAPAN

Free Entry, Market Diffusion, and Social Inefficiency with Endogenously Growing Demand*

Hiroshi Kitamura† Akira Miyaoka‡ Misato Sato §

November 23, 2012

Abstract

This paper analyzes market diffusion in the presence of oligopolistic interaction among firms. Market demand is positively related to past market size because of consumer learning, networks, and bandwagon effects. Firms enter the market freely in each period with fixed costs and compete in quantities. We demonstrate that the nature of the inefficiency under free entry can change as the market grows, and more importantly, that S-shaped diffusion can be a signal that the number of firms under free entry is initially insufficient, but eventually excessive.

JEL Classification Codes: D11, L11, L14.
Keywords: Free Entry; Market Diffusion; Intertemporal Externalities; Oligopolistic Interaction; S-shaped Diffusion.

---

*We especially thank Katsuya Takii, Shingo Ishiguro, Roberto Samaniego, Junichiro Ishida, Noriaki Matsushima, Reiko Aoki, David Flath, Makoto Hanazono, Hiroaki Ino, Masayuki Kanazaki, Jong In Lee, Keizo Mizuno, Toshihiro Matsumura, Keiichi Morimoto, Mitsuru Sunada, Ryoji Yoshioka, the conference participants at the Japanese Economic Association, the Japanese Association for Applied Economics, and 2011 Institutions and Economics International Conference, and the seminar participants at Osaka University, Osaka Prefecture University, and Kwansei Gakuin University for helpful discussions and comments. We also thank Hideshi Itoh and two anonymous referees for constructive comments and suggestions. We finally thank Haruna Tsuchiya for outstanding research assistance. The first author gratefully acknowledges financial support from JSPS Grant-in-Aid for Research Activity start-up No. 22830075, for Young Scientists (B) No. 24730220, and for Scientific Research (A) No. 22243022, organized by Masaki Nakabayashi. The second author gratefully acknowledges financial support from Grant-in-Aid for JSPS Fellows No. 24·1593 and the Global COE program "Human Behavior and Socioeconomic Dynamics" of Osaka University. The usual disclaimer applies.

†Corresponding Author: Faculty of Economics, Kyoto Sangyo University, Motoyama, Kamigamo, Kita-Ku, Kyoto-City, Kyoto, 603-8555 Japan. E-mail: hiroshikitamura@cc.kyoto-su.ac.jp Phone: +81-75-705-1584

‡Graduate School of Economics, Osaka University, 1-7 Machikaneyama, Toyonaka, Osaka, 560-0043, Japan. E-mail: jge013ma@mail2.econ.osaka-u.ac.jp

§Graduate School of Economics, George Washington University, 2115 G street, NW Monroe Hall 340 Washington DC 20052, USA. E-mail: smisato@gwmail.gwu.edu

1 Introduction

The number of suppliers changes over the lifetime of a product market. Gort and Klepper (1982) investigate 46 new products in the US from their initial introductions up to 1981.[1] They characterize the evolution of markets as having five stages. In Stage 1, the number of firms in the market is small. In Stage 2, the number drastically grows. In Stage 3, it reaches a maximum. In Stage 4, there is a shakeout of firms, and the number rapidly decreases. In Stage 5, the number stabilizes. In particular, the time pattern in the growth of the number of firms from Stage 1 to Stage 3 is characterized as "S-shaped diffusion." Gort and Klepper (1982) and other papers observe this phenomenon in several markets, such as computers, television products (Gort and Klepper, 1982), and generic drugs (Ching, 2010).[2]

In the related theoretical literature, market diffusion from Stage 1 to Stage 3 is regarded as resulting from intertemporal externalities, such as learning by doing (Jovanovic and Lach, 1989), firms' learning of the market demand (Rob, 1991), or intertemporal consumption externalities (Vettas, 2000; Kitamura, 2010).[3] The common features of these studies are that firms are atomistic price takers whose production levels are exogenously determined, and that the number of firms under free entry is socially insufficient over time.

This paper aims to investigate theoretically market diffusion with oligopolistic interaction among firms. We develop a dynamic model of market diffusion following Kitamura (2010). In his model, the market grows because of intertemporal consumption externalities, through which the market demand depends positively on the previous period's market size. His approach allows us to analyze the diffusion model with oligopolistic interaction by comparing free-entry diffusion, in which the number of firms is determined by the zero-profit condition, and socially optimal diffusion, in which the number of firms maximizes social welfare.[4] The

[1]There are several papers on the growth of markets. See, e.g., Klepper and Graddy (1990), Jovanovic and Macdonald (1994), and Klepper (1997).

[2]Empirical evidence shows that the inter- and intrafirm diffusion of new technology tends to be S-shaped. See Griliches (1957), Mansfield (1968), and Stoneman (2002) for a survey of technological diffusion. In addition, the S-shaped interhousehold diffusion is treated as a stylized fact in the marketing literature and found in color televisions (Karshenas and Stoneman, 1992), fax machines (Economides and Himmelberg, 1995), clothes dryers (Krishnan, Bass, and Jain, 1999), and mobile phones (Gamboa and Otero, 2009).

[3]See also Vettas (1998), who develops a model of two-sided learning to explain S-shaped diffusion.

[4]One of the important elements to develop a model of market diffusion is demand structure. In the models of Rob (1991) and Vettas (2000), the demand curve is a horizontal straight line, and we cannot examine the role

novel dimension here is that the output per firm is endogenously determined.[5]

We demonstrate that the existence of oligopolistic interaction turns out to be crucial in the nature of inefficiency of free entry. In the previous literature, the number of firms under free entry is insufficient over time. In our model, in contrast, the nature of the inefficiency (i.e., too few firms or too many) can change depending on the degree of market maturity, and, more importantly, S-shaped diffusion can be a signal that the number of firms under free entry is initially insufficient, but eventually excessive.

To understand this result, we begin by considering the case where the output level is exogenously determined, as in the previous literature. When the output level of each firm is exogenous, new entry only leads to the demand-shift effect; i.e., new entry today increases tomorrow's demand because of intertemporal consumption externalities. This effect can be regarded as the future benefit of increasing the current number of firms. The previous literature concludes that the number of firms under free entry is socially insufficient over time because firms under free entry do not internalize the future benefit from intertemporal externalities when they enter the market.

In contrast, when oligopolistic interaction exists, new entry today also leads to a "business-stealing effect"; i.e., new entry causes existing firms to reduce their output levels today. This effect can be regarded as the current loss of increasing the current number of firms. Therefore, in our model, the socially optimal number of firms depends on the magnitudes of the future benefit from demand shift and of the current loss from business stealing. Because, as in Kitamura (2010), an S-shaped time pattern of free-entry diffusion arises for the initially stronger but eventually weaker demand-shift effect, S-shaped diffusion can be a signal of the initially insufficient, but eventually excessive number of firms under free entry.

These results provide a new policy implication. According to the previous literature, entry into markets should be encouraged over time. Based on our result, however, entry regulations should be changed depending on the phase of market growth: entry should be initially encouraged, but eventually restricted. In particular, this policy recommendation tends to apply

of oligopolistic interaction. In contrast, the demand curve is downward sloping in Kitamura (2010). This allows us to analyze the model of diffusion with oligopolistic interaction.

[5] Bergemann and Välimäki (1997) analyze market diffusion with strategic behavior. In their model, firm output is endogenously determined, but the number of firms is fixed exogenously.

to the industries where S-shaped diffusion has already been observed in other countries.

This paper is related to the literature examining the social inefficiency of free entry. Mankiw and Whinston (1986) and Suzumura and Kiyono (1987) show that the number of firms under free entry is socially excessive because oligopolistic interaction leads to a business-stealing effect.[6] This finding is often called the "excess-entry theorem."[7] Recently, on the other hand, several papers show that free entry can be socially insufficient under different circumstances. Ghosh and Morita (2007a, b) show that if we consider the vertical relationship between firms, then free entry can lead to a socially insufficient number of firms. Mukherjee (2010) also obtains the result of insufficient entry by considering the external economies of scale under which the production cost of each firm decreases as more firms enter the market.[8]

These results show a trade-off between negative and positive externalities of new entry in a static environment and which externality becomes dominant depends on the exogenous parameters in the model. In contrast, this paper focuses on the intertemporal trade-off between the current loss and future benefit of new entry and shows that the relative size of these externalities depends not only on the exogenous parameters, but also on the phase of market growth. In addition, our analysis within a dynamic framework also enables us to explore the relationship between social inefficiency of free entry and the time pattern of market diffusion. Therefore, our approach allows us to shed light on the importance of a dynamic perspective in analyzing the inefficiency of free entry.

This paper is also related to the literature concerned with intertemporal consumption externalities.[9] Some empirical studies actually report that demand can be positively related to past market size. Goolsbee and Klenow (2002) empirically examine the importance of learn-

[6] See also von Weizsacker (1980) and Perry (1984).

[7] According to Suzumura and Kiyono (1987) and Suzumura (2012), there are two types of excess-entry theorem, the "first-best" or "second-best." In the former (latter), a social planner is assumed to be (un)able to control each firm's output or price level in the post-entry stage. In this paper, we assume a "second-best" social planner. Therefore, the socially optimal diffusion in our paper corresponds to a "second-best" social welfare optimum, and the terms "insufficient" or "excessive" are used from a "second-best" social welfare perspective.

[8] Spence (1976) and Dixit and Stiglitz (1977) show that the number of firms under free entry can be socially insufficient if consumers prefer product diversity. In addition, Mukherjee (2011) shows that under the Stackelberg leader–follower structure, free entry leads to a socially insufficient number of firms.

[9] Intertemporal consumption externalities here are also related to rational addiction, where a consumer's utility is positively related to the volume of his/her own past consumption (Becker and Murphy, 1988).

ing and network externalities in the diffusion of home computers. They find that people are more likely to buy their first home computer in areas where a large fraction of households already own computers.[10] In addition, Berndt, Pindyck, and Azoulay (2003) present empirical evidence that the past sales of a drug have a positive effect on both its value to consumers and its rate of diffusion at the brand level.[11] In a (generic) drug market, the qualities of drugs tend to be initially uncertain for doctors and patients. Therefore, the widespread use of the drug in previous periods can be a signal that it is sufficiently safe or effective.[12]

The rest of this paper is organized as follows. Section 2 sets up the model. Section 3 introduces free-entry and socially optimal equilibria. Section 4 analyzes the social inefficiency of free-entry diffusion. Section 5 gives concluding remarks. All proofs are provided in the Appendix.

2 Model

This section develops the model following Kitamura (2010). The new dimension here is the oligopolistic interaction among firms: the output per firm is endogenously determined. This modeling strategy clarifies the importance of oligopolistic interaction among firms.

We characterize the consumers' behavior in 2.1 and the firms' behavior under free entry in 2.2. Then, we introduce the timing of the game in 2.3. We assume that time is discrete and that the horizon is infinite. In this paper, it is also assumed that the market is a perishable goods market or a service market in which the service fee is charged in every period.

2.1 Consumers

There are a number of mass unit consumers for all periods. Each consumer has a different preference for a product. Let θ be the type of consumer, which is stationary for all periods and

[10]For a survey of the literature on network externalities, see Katz and Shapiro (1994).

[11]Caminal and Vives (1996) analyze the importance of past market share as a signal of product quality. Monteiro and Gonzalez (2003) analyze the role of past-sales advertising as a corporate strategy. Grinblatt, Keoharju, and Ikäheimo (2008) find that the purchases of neighbors influence a consumer's purchases of automobiles.

[12]There is also the bandwagon effect (Leibenstein, 1950). Becker (1991) studies restaurant pricing where consumer demand is positively related to market size. Biddle (1991) develops an empirical model of the bandwagon effect and shows that the current demand is positively related to past demand levels.

is uniformly distributed on the interval $[0, 1]$. We also assume that the number of consumers is a/b, where $a > 0, b > 0$. The market size, defined as the number of consumers who purchase the product, at period t is denoted by Q_t. The consumers' willingness to pay depends on the previous period's market size because of the intertemporal consumption externality. We assume the following reservation price for consumers of type θ at periods $t = 1, 2, ..., u_t(\theta)$:

Assumption 1.

$$u_t(\theta) = U(\theta, Q_{t-1}) = a\theta + \sigma(Q_{t-1}), \tag{1}$$

where $\sigma(Q_{t-1}) > 0$ *represents the intertemporal consumption externality and has the following properties:* $\sigma(0) = 0$, $\sigma'(Q_{t-1}) > 0$, $\sigma''(Q_{t-1}) < 0$, $\sigma'''(Q_{t-1}) > 0$, $\lim_{Q_{t-1}\to 0} \sigma'(Q_{t-1}) = \infty$, *and* $\lim_{Q_{t-1}\to\infty} \sigma'(Q_{t-1}) = 0$.

From Assumption 1, it is easy to see that the intertemporal consumption externality has the following two properties: it is strictly increasing in the previous market size; however, its degree, or equivalently the benefit of the externality, is strictly decreasing. This assumption guarantees that the market size converges to a finite number.

A consumer of type θ pays p_t for the product and enjoys consumer surplus of $u_t(\theta) - p_t$. The consumer purchases the product if and only if the consumer surplus is nonnegative, i.e., $u_t(\theta) - p_t \geq 0$. Then, the inverse demand function at period t, $P(Q_{t-1}, Q_t)$, becomes:

$$P(Q_{t-1}, Q_t) = \begin{cases} a + \sigma(Q_{t-1}) - bQ_t & 0 \leq Q_t \leq \frac{a}{b}, \\ 0 & Q_t > \frac{a}{b}, \end{cases} \tag{2}$$

for all $t = 1, 2, ...,$ and $0 \leq Q_{t-1} \leq a/b$. It is easy to see that the inverse demand is strictly increasing in the previous period's market size, but strictly decreasing in the current period's market size. Note that our demand structure differs from that of Rob (1991) and Vettas (2000). In their models, the demand is perfectly elastic: the demand curve is a horizontal straight line. This demand structure does not allow us to analyze the role of oligopolistic interaction, where firms compete in quantities. In contrast, the downward-sloping demand here allows us to introduce oligopolistic interaction to the model of market diffusion.[13]

[13] In addition to the perfectly elastic demand, Rob (1991) and Vettas (2000) assume demand uncertainty and firms' learning. If we introduced demand uncertainty into our model, an additional externality associated with learning the demand would emerge and the analysis would be considerably more complicated.

2.2 Firms

In this paper, in contrast to Vettas (2000) and Kitamura (2010), firms compete in quantities, and the output per firm is endogenously determined. For every period, there are incumbents and a large (infinite) number of identical potential entrants. When potential entrants decide to enter the market, they must incur a setup cost $f > 0$, which is the initial investment in purchases such as machines. We assume that machines are perfectly durable and can be operated in an environment of constant returns to scale for all periods. To simplify the analysis, we assume that the scrap value of machines is zero. Let $c > 0$ be the marginal production cost and $\beta \in [0, 1)$ denote the discount factor.

2.3 Timing

For each period, a period game consists of a two-stage game, as follows.

2.3.1 Stage 1: Entry Decision

In Stage 1, potential entrants decide whether to enter the market.[14] If they enter the market with setup cost f, they compete with incumbents in Stage 2 and earn profits. If they do not, their profits for the period are zero. Let N_t be the number of incumbents at period t and let n_t be the number of new entrants in Stage 1 of period t. By definition, we have $n_t = N_t - N_{t-1}$ for all $t = 1, 2, ...$ Assuming that $N_0 = 0$, we have $N_t = \sum_{\tau=1}^{t} n_\tau$. If the demand is small and the fixed cost is high or the discount factor is low, then entry may not occur in the first period. The following assumption guarantees first-period entry:

Assumption 2.

$$\frac{(a-c)^2}{4b} > (1-\beta)f. \tag{3}$$

Assumption 2 implies that the number of the first-period entrants is larger than one, i.e., $N_1 > 1$. If inequality (3) does not hold, the first-period entry is not profitable and does not occur. Because this condition holds for all following periods, the entry never occurs.

[14]Because the setup cost is not recoverable, and the equilibrium operating profits become nonnegative in the environment of linear demand and constant marginal cost, exit never occurs in this model.

2.3.2 Stage 2: Production

In Stage 2, firms in the market compete in quantities to maximize operating profits for the period. We assume that the equilibrium in this stage is symmetric. Let q_t^i be the equilibrium output of firm i at period t. We also define p_t as the equilibrium price in period t and $\pi_t^i = [p_t - c]q_t^i$ as the equilibrium operating profits of firm i in period t.

3 Equilibrium

This section provides the characterization of free-entry equilibrium and socially optimal diffusion. We first characterize the post-entry equilibrium at period t given the number of incumbents N_{t-1} and the number of new entrants n_t in 3.1. Then, we characterize free-entry equilibrium and socially optimal diffusion in 3.2 and 3.3, respectively.

3.1 Post-Entry Equilibrium

Given the number of firms in Stage 1, firms in Stage 2 compete in quantities and choose their output levels to maximize their profits for the period, π_t^i.[15] The post-entry equilibrium is determined by the market-clearing condition, the firms' profit-maximizing behavior, and the symmetry property. Now, we define the post-entry equilibrium as follows:

Definition 1. *Given N_{t-1} and n_t, the post-entry equilibrium consists of sequences $\{Q_t, p_t, q_t^i\}$ that simultaneously satisfy the following conditions:*

1. *Firm output is symmetric for all $t = 1, 2, ...$:*

$$q_t^i = q_t \ \text{ for all } i. \tag{4}$$

2. *The market clears for all $t = 1, 2, ...$:*

$$Q_t = N_t q_t. \tag{5}$$

[15]If we assumed that firms decide their output levels to maximize the discounted sum of future operating profits, then the analysis of socially optimal diffusion would become considerably more complicated while the results in the free-entry equilibrium are unchanged. See also footnotes 18 and 20.

3. *The market price is determined by the inverse demand for all $t = 1, 2, ...$:*

$$p_t = P\Big(Q_{t-1}, Q_t\Big). \tag{6}$$

4. *Each firm's output is the best response to other firms' outputs for all $t = 1, 2, ...$:*

$$q_t^i = \arg\max_{q_t^i \geq 0} \Big[P\Big(N_{t-1}q_{t-1}, (N_t - 1)q_t + q_t^i\Big) - c\Big]q_t^i \;\; \textit{for all } i. \tag{7}$$

According to the above definition, we identify the properties of the post-entry stage equilibrium. From equation (2) and equilibrium conditions (4)–(7), the output per firm becomes:

$$q_t = \frac{1}{N_t + 1}\frac{a + \sigma(N_{t-1}q_{t-1}) - c}{b}. \tag{8}$$

Then, we obtain the post-entry equilibrium price and operating profits per firm, respectively:

$$p_t = c + \frac{a + \sigma(N_{t-1}q_{t-1}) - c}{N_t + 1}, \tag{9}$$

and

$$\pi_t = \frac{1}{(N_t + 1)^2}\frac{(a + \sigma(N_{t-1}q_{t-1}) - c)^2}{b}. \tag{10}$$

It is easy to see that the post-entry equilibrium has the standard properties of the Cournot–Nash equilibrium under linear demand and constant marginal cost. For the analysis in the following sections, we summarize these properties of post-entry equilibrium as follows:

Lemma 1. *The post-entry equilibrium has the following properties:*

1. *Aggregate output is strictly increasing in the number of firms and is bounded:*

$$\frac{\partial N_t q_t}{\partial N_t} > 0 \textit{ and } \lim_{N_t \to \infty} N_t q_t = \frac{a + \sigma(N_{t-1}q_{t-1}) - c}{b}. \tag{11}$$

2. *Output per firm is strictly decreasing in the number of firms and converges to zero:*

$$\frac{\partial q_t}{\partial N_t} < 0 \textit{ and } \lim_{N_t \to \infty} q_t = 0. \tag{12}$$

3. *Equilibrium prices (operating profits per firm) are strictly decreasing in the number of firms and converge to the marginal cost (zero):*

$$\frac{\partial p_t}{\partial N_t} < 0, \lim_{N_t \to \infty} p_t = c \textit{ and } \frac{\partial \pi_t}{\partial N_t} < 0, \lim_{N_t \to \infty} \pi_t = 0. \tag{13}$$

One of the significant properties of the post-entry equilibrium is that the entry generates an externality effect called a business-stealing effect; i.e., the new entry decreases the incumbents' output levels. As proved in Mankiw and Whinston (1986) and Suzumura and Kiyono (1987), this effect makes the free-entry equilibrium socially excessive in the static model. In the previous literature on market diffusion, the output per firm is exogenously determined under free entry. Therefore, the second property of Lemma 1 does not hold. This leads to different welfare implications for market diffusion.

3.2 Free-Entry Diffusion

In this subsection, we characterize the free-entry diffusion given the post-entry equilibrium outcome derived in 3.1. Let n_t^e be the free-entry equilibrium number of new entrants at period t, and let N_t^e be the free-entry equilibrium number of firms at period t. For simplicity, we treat N as a continuous variable. We define $R(N_{t-1}, n_t)$ as the discounted sum of future operating profits at period t, which is composed of the direct operating profits at period t and the discounted future operating profits,[16] i.e.:

$$R(N_{t-1}, n_t) = \pi_t + \beta R(N_t, n_{t+1}), \tag{14}$$

for all $t = 1, 2, ...$ In Stage 1 of each period, potential entrants enter the market as long as the present value of net profits is positive, i.e., $R(N_{t-1}, n_t) > f$. Therefore, in each period, the number of new entrants satisfies the zero-profit condition, defined as follows:[17]

Definition 2. *Given the post-entry equilibrium outcome, the free-entry equilibrium consists of the sequence $\{n_t^e\}_0^\infty$, which satisfies the zero-profit condition for all periods, i.e.:*

$$f \geq R(N_{t-1}^e, n_t^e), \text{ with equality if } n_t^e > 0. \tag{15}$$

[16]Because entrants and incumbents in Stage 2 are symmetric at each period, and the horizon is infinite, they have the same present value of their future revenue streams.

[17]One of the important factors for the existence of market diffusion is that the horizon is infinite. When the horizon is infinite, the zero-profit condition holds with a positive number of entrants at each period. In contrast, if the horizon is finite, incumbents and entrants are not symmetric. To hold the zero-profit condition, the entrants need to achieve higher profits than the incumbents, and the number of incumbents should decrease. Therefore, under a finite horizon, the zero-profit condition does not lead to a positive number of entrants at each period.

According to the above definition, the properties of the free-entry equilibrium under the transition process are identified. From the zero-profit condition (15), the operating profits in the free-entry equilibrium become $\pi_t^e = (1-\beta)f$ for each period. In addition, from equations (9) and (10), the price in the free-entry equilibrium becomes $p_t^e = c + \sqrt{(1-\beta)bf}$ for all periods. These properties imply that both the equilibrium profits and the equilibrium price are constant. By substituting equation (10) into the zero-profit condition (15), the market diffusion under free-entry equilibrium is summarized as follows:[18]

Proposition 1. *Let N^e be the steady-state number of firms in the free-entry equilibrium that satisfies $N_t^e = N_{t+1}^e = N^e$. Suppose that $N_0^e = 0$. Then, for all $t = 1, 2, ...$, there exists a unique $n_t^e > 0$ that satisfies (15), while the free-entry equilibrium output per firm, q_t^e, and the number of firms satisfy the following conditions:*

1. *The output per firm is constant over time:*

$$q_t^e = \sqrt{\frac{(1-\beta)f}{b}}, \text{ for all } t = 1, 2, ... \tag{16}$$

2. *The number of firms is an increasing function of the number in the previous period:*

$$N_t^e = \frac{a + \sigma\left(N_{t-1}^e \sqrt{\frac{(1-\beta)f}{b}}\right) - c}{\sqrt{(1-\beta)bf}} - 1, \text{ for all } t = 1, 2, ..., \tag{17}$$

and it satisfies $N_t^e \in [0, N^e]$ for all $t = 1, 2, ...$, and monotonicity, $N_0^e = 0$ and $N_t^e \to N^e$ as $t \to \infty$.

The dynamical system of equation (17) is summarized in Figure 1. Figure 1 shows that we have $\sigma'(N^e\sqrt{(1-\beta)f}/\sqrt{b})/b < 1$ in the steady state.[19] This indicates that the number of firms under free entry does not reach the steady state as long as the degree of consumption externality is strong enough.

[Figure 1 about here.]

[18]In this paper, we assume that firms choose their output levels to maximize their profits for the period. However, even if we assumed that firms maximize the discounted sum of future operating profits, our results in the free-entry equilibrium are unchanged. This is because the free-entry equilibrium continuation profit is always equal to f due to the zero-profit condition and this does not depend on today's output level.

[19]Note that by differentiating (17) with respect to N_{t-1}^e, we have $\partial N_t^e/\partial N_{t-1}^e = \sigma'(N_{t-1}^e\sqrt{(1-\beta)f}/\sqrt{b})/b$.

The constant values of the equilibrium output, the profits per firm, and the equilibrium price have several implications. First, the time pattern of the free-entry equilibrium number of firms coincides with that of the aggregate output. In addition, the constant equilibrium price implies that the firms under free entry act as if they were price takers whose output levels were exogenously determined. Therefore, the free-entry equilibrium in this paper has basically the same properties as in the previous literature on market diffusion.

3.3 Socially Optimal Diffusion

In this subsection, we derive the socially optimal diffusion. We consider a "second-best" social planner, who can control the number of firms entering the market, but not their post-entry output or price levels. Therefore, a social planner sets the number of firms to maximize the social welfare, given the post-entry equilibrium outcome derived in 3.1.[20] Let n_t^o be the number of new entrants and N_t^o be the number of firms set by the planner at period t, and let q_t^o be the output per firm under socially optimal planning at period t. Now, we define the social optimum as follows:

Definition 3. *Given the post-entry equilibrium outcome, socially optimal planning satisfies the following Bellman equation:*

$$\begin{aligned} V(N_{t-1}) = \max_{n_t \geq 0} \Bigg\{ \int_0^{(N_{t-1}+n_t)q_t} & [a + \sigma(N_{t-1}q_{t-1}) - bQ]dQ \\ & - (N_{t-1} + n_t)q_t c - n_t f + \beta V(N_{t-1} + n_t) \Bigg\}, \end{aligned} \tag{18}$$

subject to equation (8).

The interpretation of equation (18) is as follows. The present value of the sum of future welfare is current welfare plus the discounted next-period value of the sum of future welfare. We now characterize the socially optimal diffusion as follows:

[20] If firms were assumed to choose their output levels to maximize the discounted sum of future operating profits, instead of current profits as we assume in this paper, then the social planner's problem would be considerably complicated. Because forward-looking firms take into account how their output levels affect the level of new entry set by the social planner in the future periods, the firms' equilibrium outputs in each period would be different from the stage game Nash equilibrium represented by equation (8). To keep the model tractable and facilitate a clear comparison between free-entry and socially optimal diffusion, we assume that firms act myopically in the quantity-setting stage.

Proposition 2. *Suppose that $N^o_0 = 0$. Then, the optimal diffusion path satisfies the following second-order difference equation:*

$$\begin{aligned} f =& R(N^o_{t-1}, n^o_t) \\ &+ [p^o_t - c]N^o_t \frac{\partial q^o_t}{\partial n_t} \\ &+ \beta\sigma'(N^o_t q^o_t)\Big[q^o_t + N^o_t \frac{\partial q^o_t}{\partial n_t}\Big] N^o_{t+1} q^o_{t+1} \\ &+ \beta[p^o_{t+1} - c]N^o_{t+1}\Big[\frac{\partial q^o_{t+1}}{\partial N_t} - \frac{\partial q^o_{t+1}}{\partial n_{t+1}}\Big], \end{aligned} \tag{19}$$

where $R(N_{t-1}, n_t) = \pi_t + \beta f$.

Equation (19) shows that the marginal expansion cost is equal to the marginal social benefit, which is composed of four elements. The first term on the right-hand side of equation (19) is the present value of future operating profits. The second term is the welfare loss from the business-stealing effect, which is captured as $\partial q^o_t / \partial n_t$: the new entry reduces the current output per firm. This term is regarded as the current loss: increasing current entry reduces the social welfare. The third term is the future benefit from the demand-shift effect, in which an increase in the number of firms directly raises the demand in the subsequent period. The last term is the future benefit following the business-creating effect, $\partial q^o_{t+1}/\partial N_t - \partial q^o_{t+1}/\partial n_{t+1}$: the new entry indirectly raises the output per firm in the subsequent period through the demand-shift effect.

The last three terms represent the intertemporal trade-off. The current loss gives the planner an incentive to restrict the number of firms. However, the planner has a competing incentive to raise the number of firms because of future benefits. The optimal planning is determined by the magnitudes of these losses and benefits.[21]

Note that the new dimension here beyond the previous literature is the existence of the intertemporal trade-off in the optimal planning. This trade-off has not been addressed in the previous literature concerned with market diffusion. In past works, the business-stealing and

[21] In socially optimal diffusion, in contrast to free-entry diffusion, the number of firms in the first period depends on, and is positively related to, the degree of intertemporal consumption externality. Therefore, the strong externality effect leads to an initially large number of new entrants in socially optimal diffusion. This makes the S-shaped time pattern more difficult to obtain than in free-entry diffusion.

business-creating effects are not considered, although there exists a benefit from internalizing intertemporal externalities that are engines of the market growth.[22]

4 Social Inefficiency and Time Dependence

This section analyzes the social inefficiency of free-entry diffusion by comparison with socially optimal diffusion. Note that the only difference between free-entry diffusion and socially optimal diffusion is whether the intertemporal trade-off exists. We first explore the social inefficiency of free-entry diffusion when the market is in the growing phase in 4.1. Then, we examine the case when the market is in the mature phase in 4.2.

4.1 Social Inefficiency in the Growing Phase

To begin the analysis, we rewrite equation (19). Let the current loss of increasing the number of firms at period t be $\mu(N^o_{t-1}, N^o_t)$, and let the future benefit of increasing the number of firms at period t be $\lambda(N^o_{t-1}, N^o_t, N^o_{t+1})$. Then, we rewrite equation (19) with the linear demand function as follows:

$$(1-\beta)f = \pi(N^o_{t-1}, N^o_t) - \mu(N^o_{t-1}, N^o_t) + \lambda(N^o_{t-1}, N^o_t, N^o_{t+1}), \tag{20}$$

where

$$\pi(N_{t-1}, N_t) = \frac{1}{(N_t+1)^2}\frac{(a+\sigma(N_{t-1}q_{t-1})-c)^2}{b}, \tag{21}$$

$$\mu(N_{t-1}, N_t) = \frac{N_t}{N_t+1}\pi(N_{t-1}, N_t), \tag{22}$$

$$\lambda(N_{t-1}, N_t, N_{t+1}) = \frac{\beta\sigma'(N_t q_t)}{b}\frac{N_{t+1}q_{t+1}}{N_t q_t}\frac{N_{t+1}+2}{N_{t+1}+1}\mu(N_{t-1}, N_t). \tag{23}$$

Note that the free-entry diffusion satisfies $(1-\beta)f = \pi(N^e_{t-1}, N^e_t)$ for all $t = 1, 2, ...$ Therefore, the difference in the number of firms depends on the magnitude of the current loss, $\mu(N^o_{t-1}, N^o_t)$, and the magnitude of the future benefit, $\lambda(N^o_{t-1}, N^o_t, N^o_{t+1})$, which depends on the demand-shift and business-creating effects.

[22] Because the output per firm is exogenously determined in this literature, $\partial q^o_t/\partial n_t = 0$ and $\partial q^o_t/\partial N_{t-1} = 0$ for all $t = 1, 2, ...$ It is easy to see that the business-stealing and business-creating effects in equation (20) are absent, but the future benefit from the demand-shift effect still exists.

Furthermore, the future benefit at period t is determined by the degree of discounting, β, the degree of consumption externality, $\sigma'(N_t^o q_t^o)$, the growth rate of the market size, $[N_{t+1}^o q_{t+1}^o - N_t^o q_t^o]/N_t^o q_t^o$, and the number of firms at period $t+1$, N_{t+1}^o. By comparing equations (22) and (23), it is seen that future benefits are produced by a higher discount factor, a stronger degree of consumption externality, higher market growth, and a smaller number of firms. In this environment, free entry leads to a socially insufficient number of firms:

Proposition 3. *Suppose that* $\mu(N_{t-1}^o, N_t^o) \gtrless \lambda(N_{t-1}^o, N_t^o, N_{t+1}^o)$*, i.e.:*

$$1 \gtrless \frac{\beta\sigma'(N_t^o q_t^o)}{b} \frac{N_{t+1}^o q_{t+1}^o}{N_t^o q_t^o} \frac{N_{t+1}^o + 2}{N_{t+1}^o + 1}. \tag{24}$$

Then, for all $N_{t-1}^e \gtreqless N_{t-1}^o$*, we have* $N_t^e \gtrless N_t^o$*.*

Note that we have $N_0^e = N_0^o = 0$. Therefore, Proposition 3 implies that, at Period 1, the free-entry equilibrium number of firms becomes socially insufficient, $N_1^e < N_1^o$, when the future benefit is larger than the current loss, $\mu(N_0^o, N_1^o) < \lambda(N_0^o, N_1^o, N_2^o)$.

In addition, by interpreting Proposition 3 differently, we see that the free-entry equilibrium number of firms is more likely to be socially insufficient at early periods. At early periods: (a) the degree of consumption externality is strong;[23] (b) the growth rate of the market size is high;[24] and (c) the number of firms is small. These elements increase the value of the right-hand side of inequality (24).

4.2 Social Inefficiency in the Mature Phase

Next, we turn to the analysis of the mature phase. Let N^o (q^o) denote the steady-state number of firms (output per firm) in the socially optimal diffusion that satisfies $N_t^o = N_{t+1}^o = N^o$ ($q_t^o = q_{t+1}^o = q^o$). Then, in the steady state, we can rewrite equations (20)–(23) as follows:

$$(1-\beta)f = \pi(N^o) - \mu(N^o) + \lambda(N^o), \tag{25}$$

where

$$\pi(N) = \frac{(a + \sigma(Nq) - c)^2}{(N+1)^2 b}, \tag{26}$$

[23]Note that the right-hand side of (24) is larger than 1 as long as $\beta\sigma'(N_t^o q_t^o)/b \geq 1$.

[24]Note that insufficient entrants appear even if $\beta\sigma'(N_t^o q_t^o)/b < 1$. This occurs if the growth rate of the market is high enough and the number of firms is small enough.

$$\mu(N) = \frac{N}{N+1}\pi(N), \tag{27}$$

$$\lambda(N) = \frac{\beta\sigma'(Nq)}{b}\frac{N+2}{N+1}\mu(N). \tag{28}$$

Note that $(1-\beta)f = \pi(N^e)$ holds in free-entry diffusion. Thus, as in the analysis of the growing phase, the difference in the number of firms depends on the magnitudes of the current loss, $\mu(N^o)$, and the future benefit, $\lambda(N^o)$. As the market becomes mature, the degree of consumption externality and the growth rate of the market size become lower, and the number of firms operating in the market increases. Compared with the growing phase, these changes in the market environment lower the magnitude of the future benefit relative to the current loss, i.e., $\lambda(N^o_{t-1}, N^o_t, N^o_{t+1})/\mu(N^o_{t-1}, N^o_t) \geq \lambda(N^o)/\mu(N^o)$.[25] More importantly, the following lemma shows that the future benefit becomes smaller than the current loss in the steady state:

Lemma 2. *In the steady state, the future benefit becomes smaller than the current loss, i.e., $\mu(N^o) > \lambda(N^o)$. More precisely, we have:*

$$1 > \frac{\beta\sigma'(N^o q^o)}{b}\frac{N^o+2}{N^o+1}. \tag{30}$$

Lemma 2 implies that, in the steady state, the social benefit of increasing the number of firms, $\pi(N) - \mu(N) + \lambda(N)$, becomes smaller than the private benefit of firms entering the market, $\pi(N)$. Therefore, it is optimal for the social planner to slow down the rate of new firms entering as the market becomes mature. The following proposition shows that, in the steady state, the free-entry equilibrium number of firms is socially excessive:

Proposition 4. *In the mature phase (steady state), the free-entry equilibrium number of firms is larger than the socially optimal number of firms, i.e., $N^e > N^o$.*

From Propositions 3 and 4, we conclude that the free-entry equilibrium number of firms tends to be initially insufficient but eventually excessive.[26] This is in contrast to the results of

[25]By comparing equations (23) and (28), it is easy to see that:

$$\frac{\lambda(N^o_{t-1}, N^o_t, N^o_{t+1})}{\mu(N^o_{t-1}, N^o_t)} = \frac{\beta\sigma'(N^o_t q^o_t)}{b}\frac{N^o_{t+1}q^o_{t+1}}{N^o_t q^o_t}\frac{N^o_{t+1}+2}{N^o_{t+1}+1} \geq \frac{\beta\sigma'(N^o q^o)}{b}\frac{N^o+2}{N^o+1} = \frac{\lambda(N^o)}{\mu(N^o)}, \tag{29}$$

because $\sigma''(\cdot) < 0$, $\partial N_t q_t/\partial N_t > 0$, and $N^o_t \leq N^o_{t+1} \leq N^o$.

[26]If we assumed a "first-best" social planner, who could control both the number of firms and their output or price levels, the socially optimal number of firms would be one because of economies of scale. In contrast, the equilibrium number of firms under free entry is larger than one by Assumption 2. Therefore, the "first-best excess-entry theorem" in Suzumura and Kiyono (1987) still holds in our dynamic setting.

previous studies, where the number of firms in the free-entry diffusion is socially insufficient over time. The second and third columns of Table 1 and Figure 2 present a numerical example in which the number of firms under free entry is initially insufficient (from Period 1 to Period 3) but eventually excessive (from Period 4 onward).[27]

[Table 1 about here.]

[Figure 2 about here.]

4.3 Time Pattern and Social Inefficiency

In this subsection, we first examine the time pattern of free-entry diffusion and derive the condition under which S-shaped diffusion arises under free entry. Then we discuss the relationship between the nature of inefficiency and time pattern of free-entry diffusion.

First, we show that the time pattern of free-entry diffusion becomes S-shaped (initially convex and eventually concave) when the externality effect is initially strong enough. Note that the degree of consumption externality decreases as the market size increases. This makes the number of firms under free entry monotonically converge to the steady state, and the time pattern eventually becomes concave. Therefore, the time pattern of free-entry diffusion becomes S-shaped if and only if the number of new entrants initially increases.

Note that under free entry, the numbers of new entrants in the first and second periods are $n_1^e = N_1^e = [a - c]/\sqrt{(1-\beta)bf} - 1$ and $n_2^e = N_2^e - N_1^e = \sigma(N_1^e q_1^e)/\sqrt{(1-\beta)bf}$, respectively. While the former does not depend on the consumption externality, the latter does. This implies that the strong consumption externality effect leads to the high market growth from Period 1 to Period 2 and the convex time pattern of market diffusion in the early periods:

Proposition 5. *The time pattern of free-entry diffusion becomes S-shaped if and only if:*

$$a - c - \sqrt{(1-\beta)bf} < \sigma\Big(\frac{a - c - \sqrt{(1-\beta)bf}}{b}\Big). \tag{31}$$

The mechanism of the initial convexity of free-entry diffusion here is the same as that in Vettas (2000) and Kitamura (2010), in which firms are small atomistic price takers. Therefore,

[27] In this example, the steady state of the socially optimal diffusion path is locally saddle stable.

our result establishes that the fundamental mechanism of S-shaped diffusion clarified in those papers is robust to considerations of oligopolistic interaction among firms.

From inequality (31) and the properties of $\sigma(\cdot)$, it is easy to see that low initial market size contributes to the initial convexity of free-entry diffusion. More importantly, from Figure 1, we have $\sigma'(N^e_{t-1}q^e_{t-1})/b > n^e_{t+1}/n^e_t$ for all $t = 1, 2, ...$ This implies that the free-entry diffusion up to Period $t + 1$ has a convex time pattern only if a degree of consumption externality is sufficiently strong such that $\sigma'(N^e_{t-1}q^e_{t-1})/b > 1$.[28]

Next, we explore the relationship between time pattern and social inefficiency of free-entry diffusion. Proposition 3 shows that the number of firms under free entry in the initial periods is socially insufficient for a strong degree of consumption externality. In addition, Proposition 5 shows that the initial convexity of free-entry diffusion arises for a strong degree of consumption externality. The following proposition summarizes the relationship between the convex time pattern and insufficient entry in the initial phase of free-entry diffusion:

Proposition 6. *If the time pattern of free-entry diffusion up to Period 3 is sufficiently convex so that:*

$$\frac{n^e_3}{n^e_2} \geq \frac{1}{\beta}, \tag{32}$$

the number of firms under free entry in the first period is socially insufficient, i.e., $N^e_1 < N^o_1$.

Note that the condition (31) alone (i.e., the convexity from Period 0 to Period 2) does not guarantee the insufficient entry in the initial phase of free-entry diffusion. This is because the large value of $\sigma(N^e_1 q^e_1)$ does not necessarily imply the large value of $\sigma'(N^e_1 q^e_1)/b$. Recalling that we eventually observe concave time pattern and excessive entry under free-entry diffusion, the result of Proposition 6 implies that S-shaped diffusion can be a signal that the number of firms under free entry is initially insufficient, but eventually excessive.[29]

[28]From Figure 1, it can also be found that $n^e_t/n^e_{t-1} > \sigma'(N^e_{t-1}q^e_{t-1})/b$ for all $t = 2, 3, ...$ This implies that the free-entry diffusion from Period $t-2$ to Period t has a convex time pattern if $\sigma'(N^e_{t-1}q^e_{t-1})/b \geq 1$.

[29]There is also a similar relationship between consumer surplus at each period and S-shaped diffusion. In our model, the consumer surplus at each period is calculated as $b(N_t q_t)^2/2$. Because the consumer surplus at each period is increasing in the number of firms at each period, the consumer surplus under free-entry diffusion tends to be smaller (larger) than under socially optimal diffusion if the number of firms under free entry is socially insufficient (excessive). Therefore, the S-shaped diffusion can also be a signal that the consumer surplus under free-entry diffusion is initially smaller, but eventually larger than under socially optimal diffusion.

The results in this section imply that oligopolistic interaction is an important factor for the discussion of entry regulation policies in new industries where the S-shaped diffusion has already been observed in other countries. The entry regulation policies in such industries should be changed depending on the phase of market growth: entry should be initially encouraged, but eventually discouraged. Therefore, the regulatory authority can improve social welfare by subsidizing early entry and taxing late entry. For the numerical example used above, the optimal entry subsidy/tax schedule is computed as shown in the fourth column of Table 1 and Figure 3.[30] According to them, we can find that while a subsidy should be given in Periods 1 and 2 for entry promotion, a tax should be imposed from Period 3 onward for entry restriction.[31]

[Figure 3 about here.]

5 Concluding Remarks

This paper models market diffusion in the presence of oligopolistic interaction. In contrast to the previous literature on market diffusion, not only the number of new entrants but also the output per firm is endogenously determined by the oligopolistic interaction. The major result reported here is that the nature of the inefficiency under free entry can depend on the degree of market maturity and that an S-shaped time pattern can be a signal that the number of firms under free entry is initially insufficient but eventually excessive.

Our result provides important policy implications for entry regulations in new industries where S-shaped diffusion has already been observed in other countries. Our result implies that entry should be initially encouraged but eventually discouraged. It may be possible to

[30]The optimal entry subsidy or tax is equal to the difference between the right-hand side of equation (15) and the right-hand side of equation (19), i.e., equal to the last three terms of equation (19). If these terms are positive (negative), then a subsidy (tax) should be given (imposed).

[31]From Table 1, it can be seen that, in Period 3, although the number of firms under free entry is socially insufficient, a tax is imposed on new entrants under an optimal subsidy/tax schedule. Note that, under the optimal subsidy/tax schedule, when potential entrants decide whether to enter the market at the beginning of Period 3, the number of incumbent firms is not N_2^e, but N_2^o. Given $N_2 = N_2^o$, the free-entry equilibrium number of firms in Period 3 becomes 72.11, which is larger than $N_3^o = 71.59$. Therefore, under the optimal subsidy/tax scheme, a social planner has an incentive to impose an entry tax to discourage new entry in Period 3.

improve social welfare by giving subsidies to early entrants but taxing late entrants without violating intertemporally balanced budget constraints.

The result of this paper might be important for the Japanese generic drug market.[32] In Japan, the use of generic drugs is substantially lower than in other developed countries, and the Japanese government has recently taken various policy measures for promoting the use of generic drugs.[33] Our result suggests that, while the government's recent policy is reasonable for the moment, it should be cautious of the possibility for excess entry in the future.

There are several issues requiring future research. First, as the number of firms becomes stable, there is a possibility of market restructuring. An example of market restructuring is horizontal mergers. If there exist cost synergies, market maturity may lead to horizontal mergers that reduce the number of firms and improve welfare (Davidson and Mukherjee, 2007). This may explain the shakeout of firms corresponding to Stage 3 to Stage 5 in Gort and Klepper (1982).[34] Second, there is concern about other intertemporal externalities. Although we use only intertemporal consumption externalities here, we predict that our results would hold even under other intertemporal externalities.

Finally, we assume that firms act myopically in the quantity-setting stage; i.e., firms choose their output levels to maximize their period profits. This assumption makes the model very tractable and enables us to provide a clear comparison between free-entry and socially optimal diffusion. Although our approach seems like a natural first step to analyze market diffusion with oligopolistic interaction among firms, it would be very important to allow for forward-looking firms that decide their output levels to maximize the discounted sum of fu-

[32] In the (generic) drug markets of the United States, several features are observed that are consistent with the focus of this paper. First, an S-shaped time pattern of the number of generic entrants is actually observed (Ching, 2010). Second, empirical evidence is found for the importance of intertemporal consumption externalities (Berndt, Pindyck, and Azoulay, 2003). Finally, some papers point out the significance of entry costs and the possible existence of the "business-stealing effect" in this market. Scott Morton (1999) mentions that the costs of obtaining entry approval are significant for generic entrants. In addition, using the data of drugs that went off-patent during 1976–1987, Caves, Whinston, and Hurwitz (1991) empirically show the relatively small increases in the overall generic market share achieved as the number of generic competitors increases, and mention the possibility of excess entry in this market.

[33] In 2009, the quantity-based share of generic drugs in Japan was only 20.3% (Japan Generic Medicines Association, 2012), while those for the United States, Canada, the United Kingdom, and Germany were above 60% (based on analysis of the IMS Health MIDAS Market Segmentation data by Sawai Pharmaceutical (2011)).

[34] Several papers develop theoretical models that explain the shakeout of firms. See, e.g., Jovanovic and Macdonald (1994), Horvath, Schivardi, and Woywode (2001), and Hanazono and Yang (2009).

ture operating profits. We hope that our study helps researchers to address these issues.

Appendix

Proof of Proposition 1

We first guess that $n_t^e > 0$ for all $N_{t-1}^e \in [0, N^e)$. Then, $R(N_{t-1}^e, n_t^e) = f$ for all $N_{t-1}^e \in [0, N^e)$. By solving $(1-\beta)f = \pi(N_{t-1}^e, N_t^e)$ with respect to N_t^e, we have:

$$N_t^e = N(N_{t-1}^e) = \frac{a + \sigma(N_{t-1}^e q_{t-1}^e) - c}{\sqrt{(1-\beta)bf}} - 1, \tag{33}$$

for all $t = 1, 2, ...$ Together with equation (8), we obtain equations (16) and (17). From the properties of $\sigma(\cdot)$, we have $N_1^e = N(N_0^e) = [a-c]/\sqrt{(1-\beta)bf} - 1 > 0$, $N'(N_{t-1}^e) > 0$, $N''(N_{t-1}^e) < 0$, $\lim_{N_{t-1}^e \to 0} N'(N_{t-1}^e) = \infty$, and $\lim_{N_{t-1}^e \to \infty} N'(N_{t-1}^e) = 0$. Therefore, $N(N_{t-1}^e)$ crosses the $N_t^e = N_{t-1}^e$ line only once, and there is a unique steady state, N^e. We finally verify that $n_t^e > 0$. From Figure 1, it is easy to see that we have $n_t^e > 0$ for all $N_{t-1}^e \in [0, N^e)$.

Q.E.D.

Proof of Proposition 2

Differentiating the right-hand side of (18) with respect to n_t and rearranging, we have the following first-order condition:

$$f = [p_t - c]\left[q_t + N_t \frac{\partial q_t}{\partial n_t}\right] + \beta V'(N_t). \tag{34}$$

Using the envelope theorem, we obtain:

$$\begin{aligned} V'(N_{t-1}) =& \sigma'(N_{t-1}q_{t-1})\left[q_{t-1} + N_{t-1}\frac{\partial q_{t-1}}{\partial N_{t-1}}\right] N_t q_t \\ &+ [p_t - c]\left[q_t + N_t \frac{\partial q_t}{\partial N_{t-1}}\right] + \beta V'(N_t). \end{aligned} \tag{35}$$

Then, from equations (34) and (35), we can express $V'(N_{t-1})$ as follows:

$$\begin{aligned} V'(N_{t-1}) =& \sigma'(N_{t-1}q_{t-1})\left[q_{t-1} + N_{t-1}\frac{\partial q_{t-1}}{\partial N_{t-1}}\right] N_t q_t \\ &+ [p_t - c]N_t\left[\frac{\partial q_t}{\partial N_{t-1}} - \frac{\partial q_t}{\partial n_t}\right] + f. \end{aligned} \tag{36}$$

Taking equation (36) one period forward, substituting into the right-hand side of the first-order condition (34), and rearranging terms, we finally obtain the equation (19).

Q.E.D.

Proof of Proposition 3

We prove the first case. Let $1 > \frac{\beta\sigma'(N_t^o q_t^o)}{b}\frac{N_{t+1}^o q_{t+1}^o}{N_t^o q_t^o}\frac{N_{t+1}^o+2}{N_{t+1}^o+1}$ and $N_{t-1}^e \geq N_{t-1}^o$. Suppose in negation that $N_t^e \leq N_t^o$. Then, from the properties of $\pi(N_{t-1}, N_t)$, we would have the following inequalities:

$$\pi(N_{t-1}^e, N_t^e) \geq \pi(N_{t-1}^e, N_t^o) \geq \pi(N_{t-1}^o, N_t^o). \tag{37}$$

Because $\pi(N_{t-1}^e, N_t^e) = (1-\beta)f$ in the free-entry equilibrium, inequalities (37) imply that $\pi(N_{t-1}^o, N_t^o) \leq (1-\beta)f$. This is a contradiction to equation (20) because $\mu(N_{t-1}^o, N_t^o) > \lambda(N_{t-1}^o, N_t^o, N_{t+1}^o)$. In the same way, we can prove the second case.

Q.E.D.

Proof of Lemma 2

Suppose in negation that:

$$1 \leq \frac{\beta\sigma'(N^o q^o)}{b}\frac{N^o+2}{N^o+1}. \tag{38}$$

Then, we would have the following inequalities:

$$N^o q^o \leq \frac{\beta\sigma'(N^o q^o)N^o q^o}{b}\frac{N^o+2}{N^o+1} < \frac{\sigma(N^o q^o)}{b}\frac{N^o+2}{N^o+1}, \tag{39}$$

where the last inequality follows from the properties of $\sigma(\cdot)$. By substituting $q^o = (a + \sigma(N^o q^o) - c)/(b(N^o+1))$ into inequality (39), we have

$$N^o q^o < \frac{\sigma(N^o q^o)}{b}\frac{N^o+2}{N^o+1} \Leftrightarrow N^o\left[\frac{a+\sigma(N^o q^o)-c}{b(N^o+1)}\right] < \frac{\sigma(N^o q^o)}{b}\frac{N^o+2}{N^o+1} \tag{40}$$

$$\Leftrightarrow \frac{N^o[a-c]}{2} < \sigma(N^o q^o). \tag{41}$$

By using inequality (41), we have:

$$\begin{aligned}\pi(N^o) = \frac{(a+\sigma(N^o q^o)-c)^2}{b(N^o+1)^2} &> \frac{(a+\frac{N^o[a-c]}{2}-c)^2}{b(N^o+1)^2}\\ &= \left[\frac{N^o+2}{N^o+1}\right]^2\frac{(a-c)^2}{4b} > \frac{(a-c)^2}{4b}.\end{aligned} \tag{42}$$

Note that inequality (38) implies that $\mu(N^o) \leq \lambda(N^o)$. Then, together with inequality (42) and Assumption 2 (inequality (3)), we have:

$$\pi(N^o) - \mu(N^o) + \lambda(N^o) > \frac{(a-c)^2}{4b} > (1-\beta)f. \tag{43}$$

However, this contradicts the equilibrium condition (25). Therefore, inequality (30) holds.

Q.E.D.

Proof of Proposition 4

Let $q(N)$ be the steady-state output per firm in the post-entry equilibrium given N. As mentioned in Subsection 3.2, we have $\sigma'(N^e q^e)/b < 1$ in the steady state of free-entry diffusion, where q^e is the equilibrium level of steady-state output per firm in the free-entry diffusion. Hence, we have $N^e > \hat{N}$, where $\hat{N}$ is such that $\sigma'(\hat{N}q(\hat{N}))/b = 1$. To begin the proof of Proposition 4, we first prove the following lemma:

Lemma 3. *Let $Q(N) = Nq(N)$. Then, for $N \in (\hat{N}, \infty)$,*

1. $Q'(N) \to 0$ *and* $Q(N) \to m \in (0, \infty)$ *as* $N \to \infty$.
2. $\pi(N)$ *is strictly decreasing in* N *and approaches zero as* N *becomes larger:* $\pi'(N) < 0$ *and* $\lim_{N\to\infty} \pi(N) = 0$.

Proof of Lemma 3

1. By differentiating $Q(N)$ with respect to N, we have:

$$Q'(N) = \frac{[1 - \frac{N}{N+1}]q(N)}{1 - \frac{N}{N+1}\frac{\sigma'(Q(N))}{b}} > 0, \tag{44}$$

for $N \in (\hat{N}, \infty)$. It is easy to see that $Q'(N) \to 0$ as $N \to \infty$. Then, by using L'Hôpital's rule, we obtain:

$$\begin{aligned}
\lim_{N\to\infty} Q(N) &= \lim_{N\to\infty} \Big\{\frac{a + \sigma(Q(N)) - c}{b} + \frac{\sigma'(Q(N))NQ'(N)}{b}\Big\} \\
&= \lim_{N\to\infty} \Big\{\frac{a + \sigma(Q(N)) - c}{b}\Big\} + \lim_{N\to\infty} \Big\{\frac{\sigma'(Q(N))}{b} \frac{[1 - \frac{N}{N+1}]Q(N)}{1 - \frac{N}{N+1}\frac{\sigma'(Q(N))}{b}}\Big\} \\
&= \lim_{N\to\infty} \Big\{\Big[\frac{a + \sigma(Q(N)) - c}{b}\Big]\Big[1 + \frac{\sigma'(Q(N))}{b} \frac{[1 - \frac{N}{N+1}]\frac{N}{N+1}}{1 - \frac{N}{N+1}\frac{\sigma'(Q(N))}{b}}\Big]\Big\} \\
&= \lim_{N\to\infty} \Big\{\frac{a + \sigma(Q(N)) - c}{b}\Big\}.
\end{aligned} \tag{45}$$

Since $\sigma(\cdot)$ satisfies the Inada condition, and $a > b$, there exists a unique $m \in (0, \infty)$ such that $m = [a + \sigma(m) - c]/b$.

2. By the properties of $Q(N)$, it is easy to see that $\pi(N) \to 0$ as $N \to \infty$. Next, we show that $\pi(N)$ is strictly decreasing in N. By differentiating $\pi(N)$, we have:

$$\pi'(N) = -\frac{2\pi(N)[b - \sigma'(Q(N))]}{N[b - \sigma'(Q(N))] + b} < 0, \tag{46}$$

for $N \in (\hat{N}, \infty)$. Thus, $\pi(N)$ is strictly decreasing in N.

This completes the proof of Lemma 3.

Q.E.D.

Now, we turn to the proof of Proposition 4. Suppose in negation that $N^e \leq N^o$. Then, because $\pi(N)$ is strictly decreasing in $N \in (\hat{N}, \infty)$ by Lemma 3, we have the following inequality:

$$\pi(N^e) \geq \pi(N^o). \tag{47}$$

Since N^e satisfies $\pi(N^e) = (1-\beta)f$, we have $\pi(N^o) \leq (1-\beta)f$. However, this is a contradiction to (25) because $\mu(N^o) > \lambda(N^o)$ by Lemma 2. Therefore, we have $N^e > N^o$.

Q.E.D.

Proof of Proposition 5

From Figure 1, the time pattern of free entry eventually becomes concave. Therefore, it becomes S-shaped if and only if $n_1^e < n_2^e$. From equation (17), we obtain inequality (31).

Q.E.D.

Proof of Proposition 6

Note that, from Figure 1 and the strict concavity of $\sigma(\cdot)$, we have

$$\frac{\sigma'\left(N_{t-1}^e q_{t-1}^e\right)}{b} > \frac{n_{t+1}^e}{n_t^e} \tag{48}$$

for all $t = 1, 2, ...$ Then, if the time pattern of free-entry diffusion up to Period 3 is sufficiently convex so that inequality (32) holds, we obtain

$$\frac{\sigma'\left(N_1^e q_1^e\right)}{b} > \frac{n_3^e}{n_2^e} \geq \frac{1}{\beta}. \tag{49}$$

Now, we show that $N_1^e < N_1^o$ using a proof by contradiction. Suppose in negation that $N_1^e \geq N_1^o$. Then, because $N_1 q_1$ is increasing in N_1 from Lemma 1 and $N_0^e = N_0^o = 0$ by assumption, we have $N_1^e q_1^e \geq N_1^o q_1^o$. Using the concavity of $\sigma(\cdot)$ and inequality (49), this implies that

$$\frac{\sigma'\left(N_1^o q_1^o\right)}{b} \geq \frac{\sigma'\left(N_1^e q_1^e\right)}{b} > \frac{1}{\beta}. \tag{50}$$

From this inequality, and the fact that both $N_2^o q_2^o / N_1^o q_1^o$ and $(N_2^o + 2)/(N_2^o + 1)$ are larger than 1, we have

$$\frac{\beta\sigma'\left(N_1^o q_1^o\right)}{b} \frac{N_2^o q_2^o}{N_1^o q_1^o} \frac{N_2^o + 2}{N_2^o + 1} > 1. \tag{51}$$

Then, together with $N_0^e = N_0^o = 0$, Proposition 3 implies that $N_1^e < N_1^o$. However, this is a contradiction. Therefore, we have $N_1^e < N_1^o$.

Q.E.D.

References

Becker, G., 1991. A note on restaurant pricing and other examples of social influences on price. Journal of Political Economy 99, 1109–1116.

Becker, G., Murphy, K., 1988. A theory of rational addiction. Journal of Political Economy 96, 675–700.

Bergemann, D., Välimäki, J., 1997. Market diffusion with two-sided learning. RAND Journal of Economics 28, 773–795.

Berndt, E.R., Pindyck, R.S., Azoulay, P., 2003. Consumption externalities and diffusion in pharmaceutical markets: antiulcer drugs. Journal of Industrial Economics 51, 243–270.

Biddle, J., 1991. A bandwagon effect in personalized license plates? Economic Inquiry 29, 375–388.

Caminal, R., Vives, X., 1996. Why market shares matter: an informational-based theory. RAND Journal of Economics 27, 221–239.

Caves, R.E., Whinston, M.D., Hurwitz, M.A., 1991. Patent expiration, entry, and competition in the U.S. pharmaceutical industry. Brookings Papers on Economic Activity, Microeconomics 1, 1–48.

Ching, A.T., 2010. Consumer learning and heterogeneity: dynamics of demand for prescription drugs after patent expiration. International Journal of Industrial Organization 28, 619–638.

Davidson, C., Mukherjee, A., 2007. Horizontal mergers with free entry. International Journal of Industrial Organization 25, 157–172.

Dixit, A., Stiglitz, J., 1977. Monopolistic competition and optimum product diversity. American Economic Review 67, 297–308.

Economides, N., Himmelberg, C., 1995. Critical mass and network size with application to the US fax market. Discussion Paper, School of Business, NYU, EC-95-11.

Gamboa, L.F., Otero, J., 2009. An estimation of the pattern of diffusion of mobile phones: the case of Colombia. Telecommunications Policy 33, 611–620.

Ghosh, A., Morita, H., 2007a. Free entry and social efficiency under vertical oligopoly. RAND Journal of Economics 38, 541–554.

Ghosh, A., Morita, H., 2007b. Social desirability of free entry: a bilateral oligopoly analysis. International Journal of Industrial Organization 25, 925–934.

Goolsbee, A., Klenow, P.J., 2002. Evidence on learning and network externalities in the diffusion of home computers. Journal of Law and Economics 45, 317–343.

Gort, M., Klepper, S., 1982. Time paths in the diffusion of product innovations. Economic Journal 92, 630–652.

Griliches, Z., 1957. Hybrid corn: an exploration in the economics of technological change. Econometrica 25, 501–522.

Grinblatt, M., Keloharju, M., Ikäheimo, S., 2008. Social influence and consumption: evidence from the automobile purchases of neighbors. Review of Economics and Statistics 90, 735–753.

Hanazono, M., Yang, H., 2009. Dynamic entry and exit with uncertain cost positions. International Journal of Industrial Organization 27, 474–487.

Horvath, M., Schivardi, F., Woywode, M., 2001. On industry life-cycles: delay, entry, and shakeout in beer brewing. International Journal of Industrial Organization 19, 1023–1052.

Japan Generic Medicines Association, 2012. Heisei nijusan nendo jenerikku iyakuhin shea ni tsuite (On the Share of Generic Pharmaceuticals in Fiscal 2011). Tokyo: Japan Generic Medicines Association, http://www.jga.gr.jp/pdf/H23_GE_Share.pdf (accessed 13 August, 2012).

Jovanovic, B., Lach, S., 1998. Entry, exit, and diffusion with learning by doing. American Economic Review 79, 690–699.

Jovanovic, B., Macdonald, E., 1994. The life cycle of a competitive industry. Journal of Political Economy 102, 322–347.

Karshenas, M., Stoneman, P., 1992. Flexible model of technological diffusion incorporating economic factors with an application to the spread of colour television ownership in the UK. Journal of Forecasting 11, 577–601.

Katz, M., Shapiro, C., 1994. Systems competition and network effects. Journal of Economic Perspective 8, 93–115.

Kitamura, H., 2010. Capacity expansion in markets with inter-temporal consumption externalities. Australian Economic Papers 49, 127–148.

Klepper, S., 1997. Industry life cycles. Industrial and Corporate Change 6, 145–182.

Klepper, S., Graddy, E., 1990. The evolution of new industries and the determinants of market structure. RAND Journal of Economics 21, 27–44.

Krishnan, T., Bass, F., Jain, D., 1999. Optimal pricing strategy for new products. Marketing Science 45, 1650–1663.

Leibenstein, H., 1950. Bandwagon, snob, and Veblen effects in the theory of consumers' demand. Quarterly Journal of Economics 64, 183–207.

Mankiw, N.G., Whinston, M.D., 1986. Free entry and social inefficiency. RAND Journal of Economics 17, 48–58.

Mansfield, E., 1968. Industrial Research and Technological Innovation. W.W. Norton, New York.

Monteiro, P., Gonzalez, J., 2003. We sold a million units – the role of advertising past-sales. Revista Brasileira de Economia 57, 401–419.

Mukherjee, A., 2010. External economies of scale and insufficient entry. Journal of Industry, Competition and Trade 10, 365–371.

Mukherjee, A., 2012. Social efficiency of entry with market leaders. Journal of Economics & Management Strategy 21, 431–444.

Perry, M.K., 1984. Scale economies, imperfect competition, and public policy. Journal of Industrial Economics 32, 313–333.

Rob, R., 1991. Learning and capacity expansion under demand uncertainty. Review of Economic Studies 58, 655–675.

Sawai Pharmaceutical, 2011. Annual Report 2011. Tokyo: Sawai Pharmaceutical Co., Ltd, http://www.sawai.co.jp/en/wp-content/uploads/2011/12/annual_2011.pdf.

Scott Morton, F.M., 1999. Entry decisions in the generic pharmaceutical industry. RAND Journal of Economics 30, 421–440.

Spence, A., 1976. Product selection, fixed costs, and monopolistic competition. Review of Economic Studies 43, 217–235.

Stoneman, P., 2002. The Economics of Technological Diffusion. Blackwell, Oxford.

Suzumura, K., 2012. Excess entry theorems after 25 years. Japanese Economic Review 63, 152–170.

Suzumura, K., Kiyono, K., 1987. Entry barriers and economic welfare. Review of Economic Studies 54, 157–167.

Vettas, N., 1998. Demand and supply in new markets: diffusion with bilateral learning. RAND Journal of Economics 29, 215–233.

Vettas, N., 2000. Investment dynamics in markets with endogenous demand. Journal of Industrial Economics 48, 189–203.

von Weizsacker, C.C., 1980. A welfare analysis of barriers to entry. Bell Journal of Economics 11, 399–420.

Figure 1: Dynamical system of free-entry diffusion

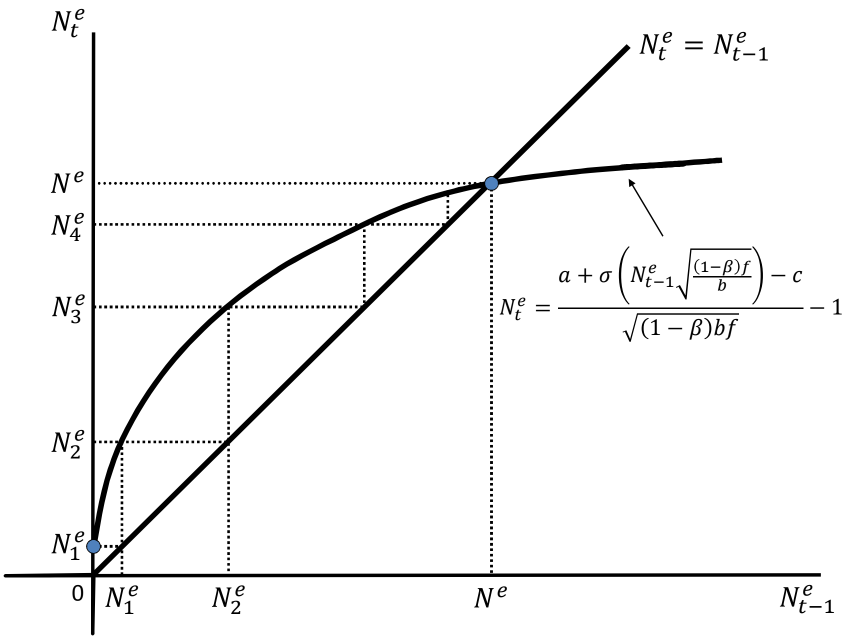

Figure 2: The free-entry and the socially optimal paths for a numerical example

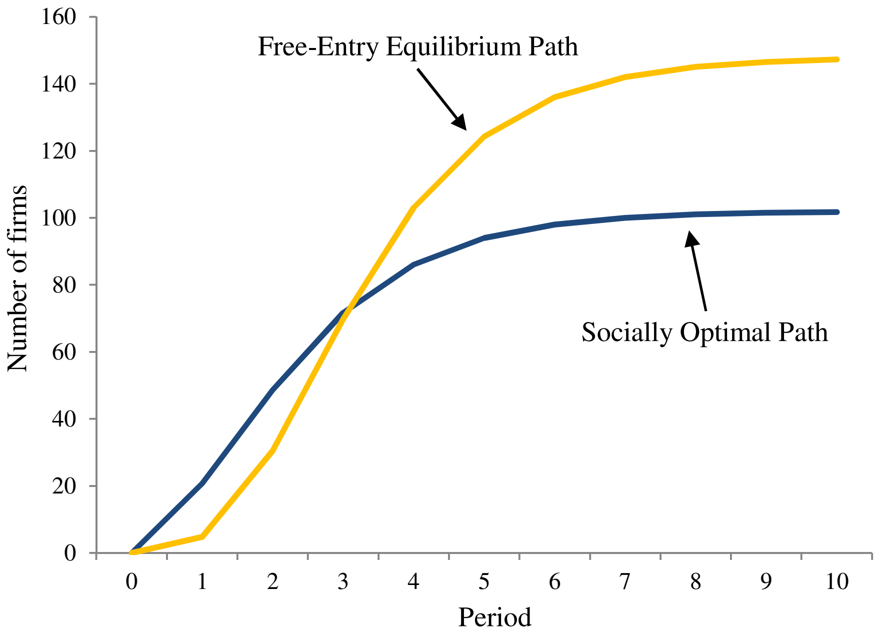

Figure 3: Optimal subsidy/tax schedule for a numerical example

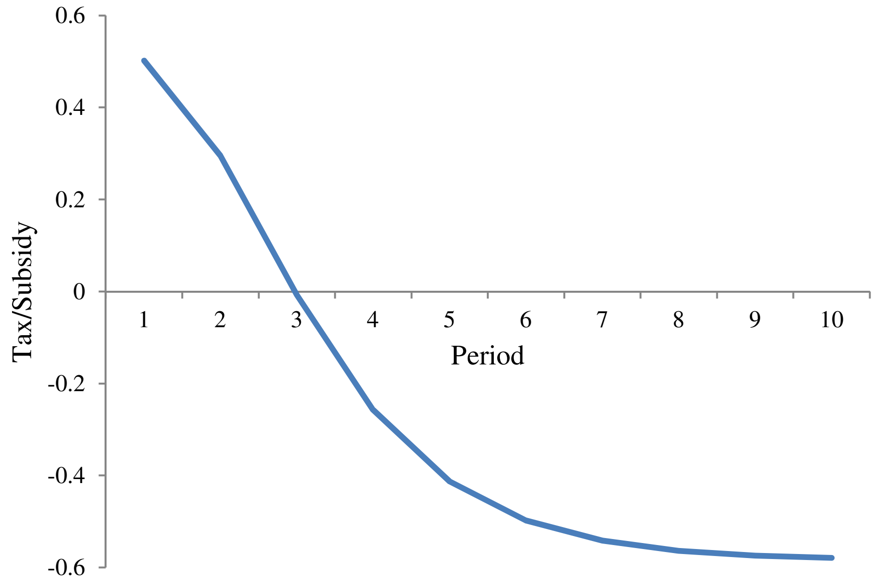

Table 1: Number of firms and optimal subsidy/tax schedule for a numerical example

Period	Free-entry Path	Optimal Path	Subsidy/Tax[a]
0	0	0	–
1	4.77	20.68	0.501
2	30.49	48.58	0.295
3	69.77	71.59	-0.007
4	103.10	86.02	-0.257
5	124.29	93.94	-0.412
6	136.01	98.00	-0.498
7	142.05	100.02	-0.541
8	145.07	101.02	-0.563
9	146.55	101.50	-0.574
10	147.27	101.73	-0.579

Note: For $a = 8$, $b = 0.5$, $c = 5$, $f = 18$, $\beta = 0.97$, and $\sigma(Q_{t-1}) = 6\sqrt{Q_{t-1}}$.

[a] Positive (negative) figures indicate the level of subsidy (tax).