

How Many Firms Should Be Leaders? Beneficial Concentration Revisited

Hiroaki Ino*

School of Economics, Kwansai Gakuin University

and

Toshihiro Matsumura†

Institute of Social Science, University of Tokyo

Abstract

We investigate the relationship between the Herfindahl-Hirschman Index (HHI) and welfare. Daughety (1990) finds that in the model wherein m leaders and $N - m$ followers compete under linear cost and demand, the Stackelberg model yields larger welfare and HHI than the Cournot model (beneficial concentration). We find that this always occurs under general cost and demand when m is sufficiently large, but does not always occur when m is small. However, if we consider the free entry of followers, beneficial concentration always occurs regardless of m . In particular, the more persistent the leadership, the more likely it is to be beneficial.

JEL classification numbers: L13, L40

Key words: HHI, beneficial concentration, leadership, free entry market

*Corresponding author. Address: 1-155 Uegahara Ichiban-cho, Nishinomiya, Hyogo 662-8501, Japan. E-mail: hiroakiino@04.alumni.u-tokyo.ac.jp.

†Address: 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. E-mail: matsumur@iss.u-tokyo.ac.jp

1 Introduction

Market concentration, measured by the Herfindahl-Hirschman Index (HHI), has played a rather important role in the context of antitrust legislations and economic regulations and it is increasingly gaining importance.¹ For many years, HHI has also received considerable attention from researchers in a wide range of fields, including both economists and law scholars.

Market concentration and economic welfare are affected by both the number of firms and the asymmetries among the firms. The effect of the number of firms is usually straightforward. Typically, in symmetric equilibria, an increase in the number of firms improves economic welfare and reduces the HHI.² In contrast, the effect of the asymmetries among the firms is complex. Daughety (1990) investigates this problem by considering asymmetric roles among the firms. He formulates a model where m Stackelberg leaders and $N - m$ Stackelberg followers compete in a homogeneous good market and discusses the relationship between m and economic welfare. Using a linear demand function and constant marginal cost, he finds an inverse U-shaped relationship between m and economic welfare. He demonstrates that the Stackelberg model ($m \in (0, N)$) yields larger welfare and HHI as compared to the Cournot model ($m = 0$ or $m = N$). The heterogeneity of the roles among the firms increases HHI and at the same time, improves welfare (beneficial concentration).

In this paper, we examine this problem closely. We extend Daughety's model in two directions. First, we consider general demand and cost functions. In particular, the generalization of the cost function is important since under strictly convex cost, asymmetric roles among the firms give rise to the production inefficiency, which does not arise under linear cost; hence, it is unclear whether beneficial concentration occurs. We find that given the total number of firms N , the introduction of a small number of followers into the Cournot model—implying that m decreases slightly from N —as in Daughety, always improves welfare and increases HHI. This result is robust because it holds

¹For example, the Fair Trade Commission in Japan finally published a merger guideline based on the HHI in 2003, although it denied using the HHI for many years.

²If we consider economies of scale, a large number of firms can be harmful. See, e.g., Mankiw and Whinston (1986) and Suzumura and Kiyono (1987).

true under general cost and demand functions. The result indicates that beneficial concentration can occur even under general demand and cost functions. On the other hand, we find that the introduction of a small number of leaders into the Cournot model—implying that m increases slightly from 0—also increases HHI but does *not* always improve welfare. Such an introduction can yield smaller welfare and higher HHI. Of course, if the marginal cost is constant, the introduction of a small number of leaders into the Cournot model always improves welfare and increases HHI (beneficial concentration). However, the latter result is not robust and does not hold true under increasing marginal cost. Thus, the relationship between HHI and economic welfare becomes much more complex when marginal cost is increasing.

Next, we endogenize the number of followers by considering the free entry of followers.³ We consider two types of long-run models that are different with respect to the persistence of leadership. In the first model (model 1), the production commitment by the leaders does not hold longer than the entry decision of followers; in the second model (model 2), it holds longer. In contrast to the case where the number of firms is exogenous, in both models 1 and 2, we find that the existence of leaders always improves welfare regardless of the number of leaders m . In other words, beneficial concentration caused by the heterogeneous roles among the firms always occurs in free-entry markets. This result indicates that HHI might be less appropriate for the measurement of welfare in free-entry markets as compared to markets with high entry barriers. It is also shown that model 2 yields larger welfare than model 1 for any m . This indicates that the more persistent the leadership, the more likely it is to be beneficial for welfare.

Finally, we discuss the properties of optimal states in a short-run model and two long-run models where m is adjusted to attain the highest welfare. We find that while the market with some leaders

³For the Stackelberg model with endogenous entry of firms, Etro (2006, 2008) also considers the model with free entry of followers. His model has a single leader and the time structure is as in our model 2. For an application of the Stackelberg model with free entry of followers to the innovation, see Etro (2004). For the limit result when the number of followers entering goes to infinity, see Ino and Kawamori (2009). For the implications of public policy in free entry markets, see Davidson and Mukherjee (2007), Brandão and Castro (2007), Mukherjee and Mukherjee (2008), and Marjit and Mukherjee (2008).

and some followers is optimal in the short-run model and long-run model 1, the market where all the followers' entries are deterred by the leaders is optimal in long-run model 2. This result provides an important implication for deregulation policies. When a regulated market is converted into a free-entry one, the non-existence of new firms' entries is often regarded as a bad signal and is associated with the leading incumbents' abuse of market power, which results in entry barriers. However, surprisingly, according to the above results, this situation can be desirable if such a leadership is sustained for very long.

The paper is organized as follows. Section 2 formulates the model where the number of firms is exogenous (short-run model). Section 3 investigates the equilibrium outcomes and presents the results in a general setting; further, Section 3 specifies the demand and cost functions and provides an insightful example. Section 4 extends the model into the two models with free entry of followers (long-run models) and compares each with the Cournot case and also compares the two models themselves. Section 5 discusses the properties of the optimal number of firms. Section 6 briefly examines the integer problem related to the number of firms. Section 7 concludes the paper and provides some insights about heterogeneity of technologies and strategic complementarity for future researches. The proofs for the Lemmas and Propositions are presented in the Appendix.

2 Model

We formulate an N -firm oligopoly ($N > 0$). Firms produce a homogeneous product with identical cost function $C(x) + f$, where $C(x) : \mathbb{R}_+ \mapsto \mathbb{R}_+$ is the production cost and $f(\geq 0)$ is the fixed entry cost. The (inverse) demand function is given by $P(X) : \mathbb{R}_+ \mapsto \mathbb{R}_+$. Each firm's payoff is its own profit. Firm i 's profit π_i is given by $P(X)x_i - C(x_i) - f$, where x_i is firm i 's output and $X \equiv \sum_{i=1}^N x_i$.

We make the following standard assumptions.

Assumption 1. $P(X)$ is twice differentiable and $P'(X) < 0$ for all X such that $P(X) > 0$.

Assumption 2. $C(x)$ is twice differentiable and $C'(x) > 0, C''(x) \geq 0$ for all $x \geq 0$.

Assumption 3. $P''(X)x + P'(X) < 0$ for all X such that $P(X) > 0$ for all $x \in (0, X)$.

Assumption 3 ensures that each firm's marginal revenue is decreasing in the rivals' outputs and that the reaction curve in the Cournot model has a negative slope, that is, the firms' strategies are strategic substitute.

The game runs as follows. In the first stage, $m(\in [0, N])$ firms (Stackelberg leaders) independently choose their outputs. In the second stage, $n (\equiv N - m)$ firms (Stackelberg followers) independently choose their outputs after observing the leaders' outputs.⁴ This model corresponds to the Cournot model when $m = 0$ or $m = N$.

We call the subgame perfect equilibrium symmetric if all leaders (followers) choose the same output level. The leaders' output is of course different from that of the followers even in a symmetric equilibrium. We make an assumption on the equilibrium outcomes.⁵

Assumption 4. The model has a unique equilibrium for all $m \in [0, N]$. The equilibrium is symmetric, and all firms produce positive outputs.

Finally, social welfare W (consumer surplus plus the profits of all the firms) is defined as

$$W = \int_0^X P(q) dq - \sum_{i=1}^N (C(x_i) + f). \quad (1)$$

HHI is defined as the sum of the square of each firm's market share:

$$HHI \equiv \sum_{i=1}^N \left(\frac{x_i}{X} \right)^2.$$

3 Short-run analysis

In this section, we analyze the relationship between m and social welfare under the short-run case where N is exogenous.

⁴Here, we ignore the integer problem of the number of firms to grasp the essential properties of the Stackelberg model with multiple leaders or free entry of followers. We discuss this integer problem in section 6.

⁵Sherali (1984) provides a sufficient condition for the existence of a unique equilibrium in a model with multiple Stackelberg leaders.

The solution $(x_L^*(m), x_F^*(m))$ of the following system of equations corresponds to the equilibrium outputs of leaders and followers.⁶ Note that (2) and (3) are the first-order conditions for leaders and followers, respectively.

$$(1 + nR'(X_L^*))P'(X^*)x_L^* + P(X^*) - C'(x_L^*) = 0, \quad (2)$$

$$P'(X^*)x_F^* + P(X^*) - C'(x_F^*) = 0, \quad (3)$$

where the equilibrium aggregate output $X^*(m) = mx_L^*(m) + nx_F^*(m)$ and the equilibrium aggregate outputs of leaders $X_L^*(m) = mx_L^*(m)$. $R(X_L)$ represents each follower's reaction to the leaders' total output X_L .⁷ $R(X_L)$ is obtained from the follower's first-order condition and is given by

$$P'(X_L + nR(X_L))R(X_L) + P(X_L + nR(X_L)) - C'(R(X_L)) = 0. \quad (4)$$

Differentiating the above yields

$$R'(X_L) = -\frac{P'(X_L + nR) + P''(X_L + nR)R}{n(P'(X_L + nR) + P''(X_L + nR)R) + (P'(X_L + nR) - c''(R))}. \quad (5)$$

From Assumptions 1–3, we obtain $nR'(X_L) \in (-1, 0)$ when $m \in [0, N]$. This implies that an increase in the leaders' output leads to a decrease in the followers' output; the total output of all firms (including the leaders and the followers), however, increases.

By substituting the equilibrium outcomes into (1), we obtain the equilibrium social welfare W^* :

$$W^*(m) = \int_0^{X^*} P(q)dq - mC(x_L^*) - nC(x_F^*) - Nf. \quad (6)$$

Further, let the equilibrium profit of a leader be given by $\pi_L^*(m) = P(X^*(m))x_L^*(m) - C(x_L^*(m))$ and that of a follower be given by $\pi_F^*(m) = P(X^*(m))x_F^*(m) - C(x_F^*(m))$.

⁶More precisely, the correspondence between $(x_L^*(m), x_F^*(m))$ and equilibrium output is as follows. When $m \in (0, N)$, $x_L^*(m)$ represents a leader's equilibrium output and $x_F^*(m)$ represents a follower's (the Stackelberg model). When $m = 0$ and $m = N$, (3) and (2) respectively correspond to the first-order condition of the Cournot model with N firms. Thus, $x_F(0) = x_L(N)$ is the equilibrium output of each firm when N firms produce simultaneously (the Cournot model). Furthermore, although (2) ((3)) has no economic implications when $m = 0$ ($m = N$), we can solve this equation and obtain $x_L(0)$ ($x_F(N)$). Although Assumption 4 does not directly guarantee that (2) ((3)) has the solution when $m = 0$ ($m = N$), we can show that $x_L(0)$ ($x_F(N)$) is well defined from Assumptions 1–4.

⁷From Assumptions 1–3, we can show that all subgames in the second stage have a unique equilibrium and that it is symmetric.

3.1 Welfare comparison

Let subscript ‘C’ denote the equilibrium outcomes when N firms produce simultaneously (the Cournot model). We present a result highlighting the difference between the Cournot outcome (where $m = 0$ or $m = N$) and Stackelberg outcome (where $m \in (0, N)$).

Lemma 1 *Suppose that Assumptions 1–4 are satisfied. Then, for all $m \in (0, N)$, (i) $x_L^*(m) > x_C^*$, (ii) $x_F^*(m) < x_C^*$ and (iii) $X^*(m) > X_C^*$.*

We present the results pertaining to the relationship between m and welfare. Proposition 1 focuses on the case of constant marginal costs.

Proposition 1 *Suppose that Assumptions 1–4 are satisfied. If $C''(x) = 0$ for all $x \geq 0$ (i.e., constant marginal cost), then $W^*(m) > W^*(0)$ for all $m \in (0, N)$.*

As seen in Lemma 1(iii), Stackelberg leaders increase the aggregate output. As such, total benefit increases and approaches the level observed in a competitive market; this results in higher welfare. Daughety (1990) has already demonstrated this result under the assumptions of linear demand and constant marginal cost. Various concentration measures, such as HHI, increase when the market structure changes from Cournot to Stackelberg. In other words, Stackelberg-type competition results in market concentration. In contrast to the traditional belief, concentration does enhance welfare. Daughety refers to this as ‘beneficial concentration’.

We discuss the robustness of Proposition 1 by considering increasing marginal cost. Proposition 1 implies that the Stackelberg model always yields greater welfare than the Cournot model regardless of m . Proposition 2 states that this result is true when m is sufficiently large, whereas it does not hold when m is sufficiently small.

Proposition 2 *Suppose that Assumptions 1–4 are satisfied. Then, (i) we can have $\frac{\partial W^*}{\partial m} \Big|_{m=0} < 0$ and (ii) we always have $\frac{\partial W^*}{\partial m} \Big|_{m=N} < 0$.*

Proposition 2(i) indicates that Proposition 1—the Stackelberg model always brings higher welfare than the Cournot model—is not robust when marginal cost is increasing. Suppose that some leaders are introduced into the Cournot model. Then, as in the case of constant marginal costs, this increases the total output. This competition acceleration effect in turn improves welfare. On the other hand, each leader’s output is strictly larger than each follower’s. Since the marginal cost is increasing, total production cost is minimized when all firms produce the same output. Thus, the introduction of a small number of leaders reduces production efficiency. Stackelberg model can bring lower welfare than the Cournot model since the latter’s welfare-reducing effect can dominate the former’s welfare-improving effect;⁸ note that the welfare-reducing effect does not exist when marginal cost is constant.

Further, while Proposition 2(i) states that the introduction of a small number of leaders into the Cournot model can reduce welfare, Proposition 2(ii) indicates that the introduction of a small number of followers always improves welfare⁹. To explain the reasoning behind this contrast, we present a supplementary result.

Lemma 2 *Suppose that Assumptions 1–4 are satisfied. Then,*

$$(i) \ x_C^* = x_L^*(N) = \lim_{m \rightarrow N} x_L^*(m) = \lim_{m \rightarrow N} x_F^*(m) = x_F^*(N),$$

$$\text{and } (ii) \ x_L^*(0) = \lim_{m \rightarrow 0} x_L^*(m) > \lim_{m \rightarrow 0} x_F^*(m) = x_F^*(0) = x_C^*.$$

The outputs of the Stackelberg leaders and followers converge to the Cournot output when m is close to N , whereas the leaders’ output does not converge to the Cournot output (while the followers’ output does) when m is close to 0.

When $m = N$, there are no followers. Thus, the leaders do not exhibit strategic behavior against other firms. The introduction of a small number of followers into the Cournot model compels the

⁸Similar trade-offs of two abovementioned effects are discussed in many contexts. See, among others, Lahiri and Ono (1988, 1998) and Matsumura (1998, 2003).

⁹Note that ‘introducing a small number of followers’ implies that m decrease slightly from N . Needless to say, $\partial W^*/\partial m|_{m=N}$ represents how W^* moves when m increase slightly from N . Thus, the introduction of a small number of followers enhances welfare if the sign of $\partial W^*/\partial m|_{m=N}$ is negative.

leaders to exhibit strategic behavior. However, the number of leaders is large, whereas that of followers is small. Thus, the strategic effect is small, and each leader's output increases slightly. This is why $x_L^*(m)$ tends to x_C^* when m tends to N . The same principle is not applicable when m tends to 0. The introduction of a small number of leaders into the Cournot model compels the leaders to exhibit strategic behavior. Since the number of leaders is small and the number of followers is large, the strategic effect is significant on each leader's output, resulting in a non-negligible difference between $x_L^*(m)$ and x_C^* . This is why $x_L^*(m)$ does not converge to x_C^* when m is close to 0.

We will now explain the rationale behind the contrast in Proposition 2. We explain why the introduction of a small number of followers into the Cournot model always improves welfare, whereas that the introduction of a small number of leaders can both reduce and improve welfare.

The following decomposition helps to explain Proposition 2(i):

$$\left. \frac{\partial W^*}{\partial m} \right|_{m=0} = [\pi_L^*(0) - \pi_F^*(0)] + N[P(X^*(0)) - C'(x_F^*(0))] \left. \frac{\partial x_F^*}{\partial m} \right|_{m=0}. \quad (7)$$

The change in welfare by the introduction of small number of leaders is decomposed into two parts. First, the unit firm's marginal change into a leader directly contributes to a renewal of its own profit (the first term). Lemma 2(ii) implies that this effect is positive, that is, the leadership enhances the welfare if we focus on the firm that becomes a leader. However, the new leader steals other firms' business (the second term). In other words, the introduction indirectly results in the reduction of the other firms' (followers') outputs; this is harmful for welfare. The change in welfare depends on the relative size of these two effects.

On the other hand, Proposition 2(ii) is elucidated by the decomposition below.

$$\left. \frac{\partial W^*}{\partial m} \right|_{m=N} = [\pi_L^*(N) - \pi_F^*(N)] + N[P(X^*(N)) - C'(x_L^*(N))] \left. \frac{\partial x_L^*}{\partial m} \right|_{m=N}. \quad (8)$$

Similarly, the change in welfare by the introduction of a small number of followers is decomposed into two parts. In contrast to the previous case, Lemma 2(i), here, implies that the first effect

of the renewal (the first term) converges to zero, that is, the reduction in the profit of the firm that becomes a follower is insignificant. Therefore, the indirect effect of the introduction on the other firms (the second term) alone determines the change in welfare. In this case, this effect is business augmenting¹⁰. In other words, the change in the role a firm from a leader to a follower indirectly causes all the other firms, as leaders, to expand their outputs; this is beneficial for welfare. Therefore, welfare always increases.

Propositions 1 and 2(ii) imply the following:

Corollary 1 *Suppose that Assumptions 1–4 are satisfied. Then, there exists $m \in (0, N)$ such that $W^*(m) > W_C^*$.*

Since the introduction of a small number of followers into the Cournot model increases HHI, this corollary implies that beneficial concentration can occur under general cost and demand functions (note that HHI is minimized when $m = 0$ and $m = N$ since there is no asymmetry in these cases). However, the introduction of a small number of leaders into the Cournot model does not always result in beneficial concentration since it increases HHI but can reduce welfare. In the next subsection, we demonstrate this by using linear demand and quadratic cost functions.

3.2 Linear demand and Quadratic cost

In this subsection, we specify the demand and cost functions and further elaborate on the relation between welfare and the number of leaders. Suppose that the inverse demand function is linear, i.e., $P(X) = a - X$, and the cost function is quadratic, i.e., $C(x) = kx^2$, where $k \geq 0$. Note that if $k = 0$ ($k > 0$), we have constant (increasing) marginal cost. The results obtained solving the model under these specifications (linear demand and quadratic cost), are summarized in Table 1, where $n \equiv N - m$. We can compute equilibrium social welfare in this specific case by substituting the results of Table 1 into

$$W^*(m) = \frac{1}{2} (X^*(m))^2 + m\pi_L^*(m) + n\pi_F^*(m). \quad (9)$$

¹⁰The terms of ‘business stealing’ and ‘business augmenting’ are taken from Mankiw and Whinston (1986).

$x_L^*(m)$	$a(2k+1)/\beta$
$x_F^*(m)$	$a(4k^2 + 2k(2+n) + 1)/\alpha\beta$
$X^*(m)$	$a(4k^2(m+n) + 2k(2+n)(m+n) + (mn + m + n))/\alpha\beta$
$\pi_L^*(m)$	$a^2(2k+1)^2(2k^2 + k(3+n) + 1)/\alpha\beta^2$
$\pi_F^*(m)$	$a^2(k+1)(4k^2 + 2k(2+n) + 1)^2/\alpha^2\beta^2$

Table 1: Results under the linear demand and quadratic cost functions: $\alpha = (2k + n + 1)$, $\beta = (4k^2 + 2k(2 + m + n) + (1 + m))$ and $n = N - m$.

We demonstrate that leadership can reduce welfare under increasing marginal costs if the number of firms is sufficiently large.

Proposition 3 *Assume linear demand and quadratic cost. If $k > 0$, there exists $N' > 0$ such that*

$$\left. \frac{\partial W^*}{\partial m} \right|_{m=0} \begin{cases} \leq 0 \\ \geq 0 \end{cases} \text{ if and only if } N \begin{cases} \geq \\ \leq \end{cases} N'.$$

Propositions 1 and 3 imply the following:

Corollary 2 *Assume linear demand and quadratic cost. There exists $N > 0$ and $m \in (0, N)$ such that $W^*(m) < W_C^*$ if and only if $k > 0$.*

As discussed above, the introduction of leaders into the Cournot model always increases HHI. This corollary implies that the introduction of leaders into the Cournot model can reduce welfare, which is the traditional belief on market concentration. We present a numerical example for such a situation.

Numerical Example

Suppose that $a = 100, k = 0.2, f = 100$ and $N = 10$. Using (9), we simulate how the ratio of leaders in the industry affects social welfare. Table 2 and Figure 1 depict a change in welfare when m changes in $[0, 10]$ given $N = m + n = 10$. Compared to the Cournot case, as predicted in Proposition 3, the welfare is smaller when the ratio of leaders is small, while as predicted in Proposition 2(ii), the welfare is larger when the ratio of leaders is large. From Table 2, we can find that the welfare is maximum at $m = 8$.

m	1	2	3	4	5	6	7	8	9
$W^*(m)/W^*(0) - 1$	-0.41%	-0.45%	-0.33%	-0.13%	0.08%	0.28%	0.46%	0.58%	0.55%

Table 2: Welfare increment by leadership in short run

Figure 1 also depicts how HHI reacts to the changes in the ratio of leaders. From this figure, we find that in our example, HHI provides a good insight into the change in welfare, except for when m tends to N . A decrease (increase) in HHI increases (decreases) welfare. In the case of constant marginal cost that Daughety (1990) investigates, the range in which both HHI and welfare increase is much wider. Under increasing marginal cost, however, an increase in concentration is more likely to be harmful, and this in turn is more likely to hold true when the number of leaders is small. Thus, it is reasonable that antitrust departments focus on markets with a small number of leaders.

Insert Figure 1 here.

4 Free entry of followers

In this section, we consider the free entry of followers. We provide long-run cases where the existence of leaders always enhances social welfare even if the cost function is strictly convex. In other words, beneficial concentration always occurs when we introduce leaders into the Cournot model, even when marginal cost is increasing. We construct two models. In both models, the number of leaders m is given exogenously¹¹ while the number of followers n is determined endogenously.

The two models are distinguished by their time structures. In model 1, potential entrants (followers) decide whether to enter the market in the first stage. Then, all firms observe m and n . In the second stage, the m leaders choose their outputs. In the third stage, the n followers (n is

¹¹If we consider the free entry of leaders, that is, if the number of leaders is endogenous and is determined by the zero-profit condition, the followers obtain negative profits and thus, no follower enters the market. In other words, the equilibrium becomes the long-run Cournot equilibrium.

determined endogenously in the first stage) choose their outputs after observing the leaders' output. In model 2, the m leaders choose their outputs in the first stage. In the second stage followers decide whether to enter the market. In the third stage, followers choose their outputs after observing n and the leaders' outputs.

We provide some cases rationalizing each model.¹² The key difference in these two models is the time at which the followers enter, that is, the relative flexibility of the followers' entry choices to leaders' production choices. Consider model 1. Suppose that m firms have a special ability to produce faster than the other firms.¹³ In other words, m firms can commit their production by inventory control and become leaders. Usually, inventory control does not increase the duration of production commitment compared to the entry decision. Thus, in this case, the leaders' production follows the followers' entry. Consider model 2. Suppose that m firms are incumbents and can make a certain capacity before the other firms. In other words, m firms can commit to their output level by capacity investments and become leaders. Usually, capacity investments can result in a production commitment before the entries of followers. Thus, in this case, the new entrants enter the market after the leaders have satisfied their production commitment. As a result, the new entrants become followers.

Henceforth, we assume the following to ensure that our long-run analysis is meaningful.

Assumption 5. Each long-run model has a unique equilibrium. The equilibrium number of followers is positive.

Note that this assumption implies that $f > 0$. Thus, the average cost $(C(x) + f)/x$ follows a U-shaped pattern if the output minimizing average cost exists and is monotonically decreasing if the

¹²Other than the cases mentioned below, we can consider a wide variety of situations suitable for our long-run models. For instance, production alliance, R&D facility, long-term contracts with customers, or sophisticated sales networks would give firms asymmetric rolls in an industry. The validity duration of commitments yielded in these ways depends on their frameworks and the environment of the industry. Thus, which model 1 or 2 is suitable is determined on a case-by-case basis.

¹³Some readers may think that such a special skill not only results in the emergence of the leadership but also often enhances technology. However, our long-run analysis indicates that even if the effect on technology is neglected, beneficial concentration occurs.

output minimizing average cost does not exist.¹⁴

In the following two subsections, we demonstrate that the two models have similar implications on beneficial concentration.

4.1 Model 1: Followers' entry before Stackelberg competition

In this subsection, we investigate model 1. Let $n^1(m)$ be the equilibrium number of followers, which is obtained from the following zero-profit condition for followers:

$$P(X^1(m))x_F^1(m) - C(x_F^1(m)) - f = 0, \quad (10)$$

where $X^1(m) = mx_L^1(m) + n^1(m)x_F^1(m)$, and $x_L^1(m)$ and $x_F^1(m)$ represent the long-run equilibrium output of a leader and a follower, respectively. $N^1(m)$ represents the total number of firms and is given by $N^1(m) = m + n^1(m)$. Since the subgames following the followers' entry decisions are the same as in the short-run model analyzed in Section 3, these equilibrium outcomes are obtained by conditions (2), (3) and (10). Using the expressions in the previous section, we have $x_L^1(m) = x_L^*(m)$ and $x_F^1(m) = x_F^*(m)$ for $N = N^1(m)$. The case where $m = 0$ corresponds to the long-run Cournot model (Cournot model with free entry). As in the previous sections, subscript C denotes the equilibrium outcome in the Cournot model.

We provide the relationship between the equilibrium outcomes of the Cournot and Stackelberg models.

Lemma 3 *Suppose that Assumptions 1–5 are satisfied. Then, $\forall m > 0$, (i) $x_F^1(m) = x_C^1$, (ii) $X^1(m) = X_C^1$, (iii) $x_L^1(m) > x_C^1$ and (iv) $N^1(m) < N_C^1$.*

Lemma 3(i) implies that each follower's output does not depend on m . Since the equilibrium price is equal to the follower's average cost (zero-profit condition), the equilibrium price remains

¹⁴If the average cost is decreasing (including the case of constant marginal cost with some fixed cost), Assumption 5 is never satisfied in model 2 (while it can be satisfied in model 1). This is because the aggregate output of the leaders undercuts the Cournot price in model 2, and no follower can enter in equilibrium. However, even in this case, if the given number of leaders is not wastefully large, beneficial concentration occurs since the optimal case with decreasing average cost is monopoly.

unchanged unless the follower's output changes (this implies Lemma 3(ii)).

Lemma 3(i) is a key result. Regardless of m , the average cost curve of each follower must be tangent to the "residual demand curve" of each follower at the long-run equilibrium. As is easily guessed, the change from the Stackelberg model ($m > 0$) to the Cournot model ($m = 0$) reduces the leaders' output for given n . This leads to an upward shift in each follower's residual demand curve; however, in the long-run, it induces new entries, leading to a downward shift in each follower's residual demand curve. Eventually, the upward shift of the residual demand curve is negated by the new entries, resulting in an unchanged curve. This is why the two models yield the same equilibrium output for each follower at the long-run equilibrium (see Figure 2).

Insert Figure 2 here.

Let $W^1(m)$ denote the long-run equilibrium welfare for given m . Note that $W_C^1 = W^1(0)$ is the equilibrium welfare in the Cournot-type model. Proposition 4 indicates a striking contrast to the short-run case.

Proposition 4 *Suppose that Assumptions 1–5 are satisfied. Then $W^1(m) > W_C^1 \forall m > 0$.*

The introduction of a leader yields asymmetry among firms and increases HHI. In addition, it reduces the total number of firms (Lemma 3(iv)), again leading to an increase in HHI. However, it improves welfare (Proposition 4). In other words, beneficial concentration always occurs.¹⁵ Notice that as seen in Lemma 3 (ii), since consumer surpluses are the same in both Cournot and Stackelberg equilibria, beneficial concentration occurs owing to the extra profit of the leaders. Thus, in this context, it is also inappropriate to uncritically use the extra profit of particular leading firms as a dangerous sign indicating monopoly power.

¹⁵Long- and short-run analysis often have different implications in oligopoly models. See, among others, Lahiri and Ono (1995) and Matsumura and Kanda (2005).

We explain the reasoning behind this result. Consider that some firms in the Cournot model become leaders. Recall that in the short-run case discussed in the previous section, the introduction of leaders yields the competition acceleration (welfare-improving effect) and production inefficiency (welfare-reducing effect). The relative size of these effects determines the total impact on welfare. In the long-run case, no competition acceleration effect arises since the number of leaders does not affect total output and price (Lemma 3 (ii)). Hence, the effect that needs to be discussed pertains to production efficiency. With regard to production efficiency, intuitively speaking, leaders' aggressive behavior deters some followers from excessively entering¹⁶ and saves on the fixed cost. Why is it that this saving always dominates the production inefficiency arising from the asymmetry between leaders and followers? Lemma 3 shows that leaders' total output increases and that followers' total output decreases by the same amount. Since the increase in leaders' total output is caused by the expansion of each firm's output (Lemma 3(iii)), this production substitution approximately increases the leaders' cost by $C'(x_L)$ (leaders' *marginal* cost) for every unit of output expansion. On the other hand, the decrease in followers' total output is because the number of followers entering is reduced; each follower's output is constant (Lemmas 3(i) and 3(iv)). Thus, the production substitution reduces the followers' cost by $(C(x_F) + f)/x_F$ (followers' *average* cost) for every unit of output reduction. Since followers' average cost is equal to the price and leaders' marginal cost is lower than the price,¹⁷ each marginal unit of production substitution saves on cost. Therefore, in the long run, production substitution always improves production efficiency. Hence, the Stackelberg model always yields larger welfare than the Cournot model.

4.2 Model 2: Leadership before followers' entry

In this subsection, we formulate a model with an alternative time structure. After observing the leaders' outputs, new entrants decide whether or not to enter the market. Let superscript "2" denote the equilibrium outcomes in this model. Note that even in this model, $m = 0$ corresponds

¹⁶As for the excess entry, see Mankiw and Whinston (1986) and Suzumura and Kiyono (1987).

¹⁷Note that by Lemma 3(ii), the price is common for the Cournot and Stackelberg models.

to the long-run Cournot model. Therefore, we have $x_C^2 = x_C^1$, $X_C^2 = X_C^1$ and $N_C^2 = N_C^1$ (subscript C denotes the outcomes in the Cournot model).

Lemma 4 *Suppose that Assumptions 1–5 are satisfied. Then, $\forall m > 0$, (i) $x_F^2(m) = x_C^2$, (ii) $X^2(m) = X_C^2$, (iii) $x_L^2(m) > x_C^2$ and (iv) $N^2(m) < N_C^2$.*

Lemma 4 is similar to Lemma 3, indicating that the results derived in model 1 also holds true in model 2.

Fundamentally, models 2 and 1 differ only in terms of the followers' reactions to the leaders' actions. In model 2, when the followers react to the leaders' output levels, the former may not only change their output levels but also take decisions pertaining to their entry/exit in the market. Therefore, at any given level of the leaders' output, each follower's output and thus, the total output are always equal to those in the Cournot model due to the same reason as in Lemma 3. Thus, the equilibrium price is constant regardless of the leaders' actions. In other words, the leaders become price takers at the Cournot price. As a result, a leader's output in this model exceeds that in the Cournot model.

Proposition 5 *Suppose that Assumptions 1–5 are satisfied. Then, $W^2(m) > W_C^2 \forall m > 0$.*

The introduction of leaders increases HHI; this is as in the previous subsection. Therefore, Proposition 5 states that beneficial concentration always occurs in model 2 as well as in model 1. The proof and intuition for Proposition 5 are similar to that for Proposition 4; therefore, these have been excluded. Etro (2008) also shows that when $m = 1$, Stackelberg competition in model 2 yields higher welfare than Cournot competition with free entry.¹⁸ Thus, our result in Proposition 5 is an extension of his result for the model with multiple leaders.

¹⁸See Proposition 3 in his paper.

4.3 Comparison of models 1 and 2

As seen in the previous two subsections, when the Stackelberg case is compared to the Cournot case, models 1 and 2 have a similar effect on beneficial concentration. However, these two models are distinct if we compare their equilibrium outcomes.

Lemma 5 *Suppose that Assumptions 1–5 are satisfied. Then, $\forall m > 0$, (i) $x_F^1(m) = x_F^2(m)$, (ii) $X^1(m) = X^2(m)$, (iii) $x_L^1(m) < x_L^2(m)$, and (iv) $N^1(m) > N^2(m)$.*

Lemma 5 implies that each follower's output and total output are the same in the two models. However, each leader's output levels and the number of followers are different for the two models.

In model 2, the leaders behave as price takers at the Cournot price. On the other hand, in model 1, the leaders' marginal cost is still lower than the Cournot price. Since the Cournot price is the same in both models, each leader's equilibrium output is larger in model 2 than in model 1, and as a result, the number of followers is smaller in model 2 than in model 1. Since in model 2, leaders choose their outputs taking followers' entry into consideration, they become more aggressive and as a result, more effectively deter followers from entering. Since the source of beneficial concentration is the savings in followers' entry cost, it is straight-forward that model 2 yields higher welfare than model 1.

Proposition 6 *Suppose that Assumptions 1–5 are satisfied. Then, $W^2(m) > W^1(m) \forall m > 0$.*

Proposition 6 states that the beneficial impact of long-run leadership on welfare is greater in model 2 than in model 1. This implies that leadership is more desirable if the leadership persistently adopts a more inflexible production position to followers' entry choices.

Numerical Example

To illustrate the properties of the two long-run models in detail, we provide numerical results. Suppose linear demand $P(X) = a - X$ and quadratic cost $C(x) = kx^2$, where $k \geq 0$. Further,

$a = 100, k = 0.2, f = 100$ (these parameters are the same as in the numerical example in Subsection 3.2). Figures 3, 4 and 5, respectively, depict how the long-run equilibrium output of a leader, number of followers and welfare change in m . From Lemma 5 and Proposition 6, we have $x_L^1(m) < x_L^2(m)$, $n^1(m) > n^2(m)$ and $W^2(m) > W^1(m)$ as far as followers enter. Table 3 numerically expresses the welfare change. We find that $m = 4$ yields maximum welfare in model 1. On the other hand, in model 2, we can see that welfare is increasing in m . This observation provides us with a conjecture that there will be some differences between models 1 and 2 with regard to the reaction of welfare to the changes in the number of leaders. We will discuss this in the next section.

m	1	2	3	4	5	6	7	8	9
$W^1(m)/W^1(0) - 1$	2.28%	4.14%	5.39%	5.92%	5.79%	5.11%	4.02%	2.62%	0.98%
$W^2(m)/W^2(0) - 1$	2.74%	5.48%							

Table 3: Welfare and the number of leaders in the long run

Insert Figures 3–5 here.

5 Optimal number of leaders

In this section, we discuss the properties of the optimal number of leaders in the short-run model and the two long-run models.

Proposition 7 *Suppose that Assumptions 1-4 are satisfied. (i) In the short run, the number of leaders that yields the maximum $W^*(m)$ is in $(0, N)$. Suppose, in addition, that Assumption 5 is satisfied. (ii) In the first long-run model (model 1), the number of leaders that yields the maximum $W^1(m)$ is in $(0, N_C^1)$. (iii) In the second long-run model (model 2), $W^2(m)$ is increasing in the number of leaders.*

Both (i) and (ii) state that the market with both leaders and followers is desirable. However, with regard to (i), as seen in Figure 1, we can state that the equilibrium welfare when the number of leaders is small can be less than that in the Cournot case (rotated S-shaped relationship between m and $W^*(m)$ in Figure 1). On the other hand, with regard to (ii), as seen in Figure 5, the equilibrium welfare for any given number of leaders is greater than that in the Cournot case (inverse U-shaped relationship between m and $W^1(m)$ in Figure 5). Therefore, while the optimal number of leaders in model 1 is typically not that large (somewhere around $N_C^1/2$), the optimal number of leaders in the short-run model tends to be large (biased towards N).¹⁹

There is a sharp contrast in (ii) and (iii). While the optimal number of leaders in model 1 is such that some followers are allowed to enter, the number of leaders in model 2 is, as long as some follower enters (Assumption 5), more desired if it gets larger. This implies that the optimal number of leaders in model 2 is such that no follower is allowed to enter. Indeed, if we relax Assumption 5 and include the case where no follower is allowed to enter, we obtain the optimal number of leaders for model 2.²⁰

If we consider the constant marginal cost case, the second contrast is remarkable. As noted in Etro (2006), in Model 2 with constant marginal cost, single leader blocks all the followers from entering. Thus, this leader can earn huge profits by undercutting the long-run Cournot price. Welfare is enhanced by this amount; the savings made on followers' fixed cost is substantially large. On the other hand, in Model 1 with the constant marginal cost, some followers enter the market as long as $m < N_C^1$. Therefore, $m \geq N_C^1$ must hold if the leaders are to block all the followers from

¹⁹Recall that in the simulations, the optimal number of leaders in $m = 8$ in the short-run model and $m = 4$ in model 1.

²⁰ Consider the case where m is so large that $x_F^2(m) = 0$ holds. When $m \geq N_C^2$, the m -firms' Cournot price is lower than or equal to $P(X_C^2)$. Thus, we find that the equilibrium of model 2 is the Cournot equilibrium with m firms. When $m \leq N_C^2$, the m -firms' Cournot price is higher than or equal to $P(X_C^2)$. Thus, the situation where each leader produces X_C^2/m is one of the equilibria of model 2 (the equilibrium price is $P(X_C^2)$, and $x_F^2(m) = 0$). In this case, the dotted line in Figure 5 (Figure 3) denotes the change in welfare (each leader's output) in m . If m is larger than the point at which the dotted line tangent to the thin line, no followers enter. If $m = N_C^2$ which is the point at which $X_C^2/m = x_C^2$ holds, the situation corresponds to the m -firm Cournot equilibrium. Therefore, by Proposition 7(iii), the optimal number of firms must be between these two points. Observe that in the figure, if we connect the thin line and the dotted line, the inverse U-shaped relationship between m and $W^2(m)$, which is greater than $W^1(m)$, emerges.

entering. Hence, there are no savings on fixed cost.

6 Integer problem

Finally, we provide some qualifications for our results on beneficial concentration with respect to an integer problem. Although our main results in Corollary 1, Proposition 4 and Proposition 5 suggest a tendency towards beneficial concentration, we have, thus far, not considered the integer constraint for the number of firms. Corollary 1 depends on the fact that we can slightly decrease the number of leaders, while Propositions 4 and 5 depend on the fact that the number of followers is a real number. Thus, it is important to consider the integer constraint.

However, in the short-run model, we obtain similar results even when there is an integer constraint if the model has a linear demand and quadratic cost (as in Subsection 3.2). More concretely, we can show that under the constraint $N \in \{2, 3, \dots\}$, (i) when $k > 0$, there exists $N' > 2$ such that $W^*(1) - W^*(0) \leq 0$ if and only if $N \geq N'$ and (ii) $W^*(N - 1) - W^*(N) > 0$ for all N .²¹

The integer problem is more serious in the long-run models. Suppose that the number of firms is an integer. Then, if not even only additional entry is profitable, the firms do not enter even when each follower makes a positive profit. This implies that the long-run price can exceed the followers' average cost. Thus, the integer constraint reduces the consumer surplus. If the integer problem is more severe in the Stackelberg model than in the Cournot model, integer constraint works as a counter effect to the beneficial concentration. Indeed, in both models 1 and 2, we can present counter examples wherein the beneficial concentration does not occur for some m even in the model with linear demand and quadratic cost (shown in Appendix B).

7 Concluding remarks

In this paper, we analyze welfare in a model with multiple Stackelberg leaders and reexamine the relationship between HHI and welfare. We find that beneficial concentration always occurs when

²¹The proofs are available upon request.

a small number of followers are introduced. In contrast, the introduction of a small number of leaders can both reduce and improve welfare. Such an introduction, however, always increases HHI. We also find that if we consider the free entry of followers, beneficial concentration always occurs when we introduce a leader into the Cournot model. This result indicates that HHI might be less appropriate for the measurement of welfare in free-entry markets than in markets with high entry barriers. In particular, beneficial concentration always occurs even if the leaders do not take the entry of followers into account (model 1). However, if the leadership is so persistent that the leaders can take the entry of followers into account (model 2), it is more beneficial for welfare. Indeed, in this case, the optimal number of leaders is such that no follower is allowed to enter.

In long-run models, we assume that the number of leaders is given exogenously. This setting implies that a limited number of firms have the ability to become leaders. In reality, there are various situations that are suitable for this limited leadership, for example, incumbents with bottleneck facilities as in the power industry, long-established firms in a production alliance as in the airline industry and pioneering firms successfully creating brand loyalty among their consumers as in the mobile phone market.²² In antitrust policies, limited leadership is often considered as a problem that results in entry barriers. However, our result indicates that the possibility of preventing inefficient entries enhances welfare if the leadership is retained for a long time.

In this paper, we assume that all firms have identical cost functions and that only the difference of roles among the firms yields asymmetry. However, particularly in our long-run models, the reasoning behind Lemma 3(i) is applicable even if we introduce leaders whose cost is different from that of the entrants in the Cournot model. Hence, the price and followers' profits remain unchanged after the introduction, and thus, the entire net social benefit of this introduction must serve as a reward for the leaders. Thus, we obtain the implications on welfare comparison: the welfare in the Stackelberg model ($m > 0$) is greater than that in the Cournot model ($m = 0$) if and only if

²²See Etro (2007) for other examples and applications. In chapter 6, he argues that Microsoft in the software market is an example of limited leadership.

the profits of the leaders are positive. If we introduce the cost differences among firms into the short-run model, the analysis will become more complicated. Leadership by a firm with a lower (higher) marginal cost is more likely to improve (reduce) welfare. However, it is difficult to find another condition for welfare improvement since not only the leaders but also the followers and consumers are relevant to changes in welfare. Extending our analysis in this direction remains for future research.²³

In the body of this paper, the key explanation of the welfare improvement was the aggressive behavior of the leaders. Thus, Assumption 3, which indicates that the firms' strategies are strategic substitute, seems to be the critical assumption for the beneficial concentration. However, this will only be true in the short run and not necessarily true in the long run. In the games where the firms' strategies are strategic complement, leadership makes the firms accommodative and thus, reduces the total output in the short run. This implies that the competition accelerating effect mentioned in the case of strategic substitute is replaced by the competition decelerating effect. At the same time, the production inefficiency caused by the asymmetric production between leaders and followers remains. Since leadership reduces welfare in both the aspects, welfare improvement does not occur in the short run. Let us consider the long-run time structure as in model 1 and the firms' strategies as strategic complement. Then, the free entry in Cournot competition results in insufficient entries.²⁴ If some firms assume leadership, their accommodative behavior will further increase the number of entries and this will alleviate the insufficiency of the entries and improve welfare. However, more formal and general analysis will be needed in this regard. This is left for the future.²⁵

²³For the discussion with a single leader and a fixed number of followers under cost heterogeneity, see Levin (1988). For the discussion on endogenous roles among firms with heterogeneous costs, see Ono (1978,1982). For the discussion of the endogenous role of N-firm oligopoly, see Matsumura (1999) among others. For the discussion on the role of firms based on experimental studies, see Huck et al. (2001).

²⁴See footnote 7 in Mankiw and Whinston (1986).

²⁵For a time structure as in model 2, Etro (2008) provides the applications of price competition with Logit demand or Dixit-Stiglitz demand and endogenous entry, and shows that a single leader improves welfare.

APPENDIX

A Proofs

Proof of Lemma 1: In Section 3, we have already shown that an increase (decrease) of x_L decreases (increases) x_F and increases (decrease) X given m and n (see the discussion immediately after equation (5)). Thus, Lemma 1(i) implies Lemma 1(ii) (Lemma 1(iii)) and *vice versa*. Thus, we merely have to prove Lemma 1(i). We prove it by contradiction. We assume $x_L^*(m) \leq x_C^*$ and derive a contradiction. Note that $x_L^*(m) \leq x_C^*$ implies $X^*(m) \leq X_C^*$.

Substituting $m = 0$ into (3) yields

$$P'(X_C^*)x_C^* + P(X_C^*) - C'(x_C^*) = 0, \quad (11)$$

where $X_C^* = X^*(0) = Nx_C^* = Nx_F^*(0)$. We compare the LHS in (11) with that in (2). Assumption 3 ensures that $P'(X)x + P(X)$ is decreasing in X . Since $X^*(m) \leq X_C^*$, we have $P'(X_C^*)x_C^* + P(X_C^*) - C'(x_C^*) \leq P'(X^*(m))x_C^* + P(X^*(m)) - C'(x_C^*)$. Since $P' < 0$, $C'' \geq 0$, and $x_L^*(m) \leq x_C^*$, we have $P'(X^*(m))x_C^* + P(X^*(m)) - C'(x_C^*) \leq P'(X^*(m))x_L^*(m) + P(X^*(m)) - C'(x_L^*(m))$. As is shown in Section 3 (immediately after equation (5)), $-1 < (N - m)R' < 0$ for $m \in [0, N]$. Thus we have $P'(X^*(m))x_L^*(m) + P(X^*(m)) - C'(x_L^*(m)) < (1 + (N - m)R')P'(X^*(m))x_L^*(m) + P(X^*(m)) - C'(x_L^*(m))$. Therefore, if (11) is satisfied, the LHS in (2) must be positive, a contradiction. **Q.E.D.**

Proof of Proposition 1: Let c denote the (constant) marginal cost. Then we obtain

$$W^*(m) - W^*(0) = \int_{X_C^*}^{X^*(m)} P(X)dX - c(X^*(m) - X_C^*) > [P(X^*(m)) - c](X^*(m) - X_C^*) > 0.$$

The inequalities hold since the price exceeds the marginal cost and from Lemma 1(iii). **Q.E.D.**

Proof of Proposition 2: (i) Proposition 3 in Section 4 presents examples where $\frac{\partial W^*}{\partial m} \Big|_{m=0} < 0$ when $N > N'$.

(ii) We show that for any situation satisfying Assumptions 1–4, $\partial W^*(N)/\partial m < 0$. Differentiating (6) and evaluating at $m = N$, we obtain (8). Since $x_L^*(N) = x_F^*(N) = x_C^*$ (Lemma 2(i)), $\pi_L^*(N) - \pi_F^*(N)$ is canceled out. Thus, if the last term on the RHS of (8) is negative, $\partial W^*/\partial m|_{m=0}$ is also negative. Comparative statics analysis using (2) and (3) yields

$$\frac{\partial x_L^*}{\partial m} \Big|_{m=N} = \frac{P' x_C^* (P' - C''(x_C^*)) R' \Big|_{m=N}}{N(P' + P'' x_C^*) (P' - C''(x_C^*)) + (P' - C''(x_C^*))^2} < 0 \quad (12)$$

where $R' \Big|_{m=N} = -(P'(X_C^*) + P''(X_C^*) x_C^*) / (P'(X_C^*) - C''(x_C^*)) < 0$. We use $x_L^*(N) = x_F^*(N) = x_C^*$ to induce (12), which implies that the last term on the RHS is negative. **Q.E.D.**

Proof of Lemma 2: We regard the LHS in (2) as a function $F(m, x_L^*(m), x_F^*(m))$ and the LHS in (3) as a function $G(m, x_L^*(m), x_F^*(m))$. Continuity of $F(\cdot)$ yields

$$\lim_{m \rightarrow 0} F(m, x_L^*(m), x_F^*(m)) = F\left(0, \lim_{m \rightarrow 0} (x_L^*(m), x_F^*(m))\right).$$

On the other hand, since $F(m, x_L^*(m), x_F^*(m)) = 0 \forall m \in (0, N)$, $\lim_{m \rightarrow 0} F(m, x_L^*(m), x_F^*(m)) = 0$. Hence, $F(0, \lim_{m \rightarrow 0} (x_L^*(m), x_F^*(m))) = 0$. Similarly, $G(0, \lim_{m \rightarrow 0} (x_L^*(m), x_F^*(m))) = 0$. Therefore, $\lim_{m \rightarrow 0} (x_L^*(m), x_F^*(m))$ is a solution of the system of equations (2) and (3) when $m = 0$. Thus, by definitions of $x_F^*(0)$, $x_L^*(0)$, and x_C^* , we have $\lim_{m \rightarrow 0} x_F^*(m) = x_F^*(0) = x_C^*$ and $\lim_{m \rightarrow 0} x_L^*(m) = x_L^*(0)$. Similarly, taking $m \rightarrow N$, we obtain $\lim_{m \rightarrow N} x_L^*(m) = x_L^*(N) = x_C^*$ and $\lim_{m \rightarrow N} x_F^*(m) = x_F^*(N)$.

Consider $m = 0$. Note that the proof of Lemma 1 is also applicable when $m = 0$. Therefore, $x_L^*(0) > x_C^*$. Next, consider $m = N$. Note that when $m = N$, nR' is canceled from (2). Thus, $x_L^*(N) = x_C^*$ is obtained only by (11). Then, given $x_L^*(N) = x_C^*$, $x_F^*(N)$ is determined according to (3), i.e., $P'(N x_C^*) x_F^* + P(N x_C^*) = C'(x_F^*)$. $x_F^*(N)$ satisfies this equation if $x_F^*(N) = x_C^*$ by (11). Then, uniqueness of $x_F^*(N)$ implies that $x_F^*(N) = x_C^* = x_L^*(N)$. **Q.E.D.**

Proof of Proposition 3: Differentiating (6) and evaluating at $m = 0$, we obtain (7). From the

results in table 1, we have

$$\pi_L^*(0) = P(X^*(0))x_L^*(0) - C(x_L^*(0)) = \frac{(a-c)^2(2k+b)^2(2k^2+bk(3+N)+b^2)}{(2k+Nb+b)(4k^2+2bk(2+N)+b^2)^2}, \quad (13)$$

$$\pi_F^*(0) = P(X^*(0))x_F^*(0) - C(x_F^*(0)) = \frac{(a-c)^2(k+b)}{(2k+Nb+b)^2}. \quad (14)$$

From the first-order condition (3),

$$P(X^*(0)) - C'(x_F^*(0)) = -P'(X^*(0))x_F^*(0) = bx_F^*(0) = \frac{b(a-c)}{2k+Nb+b}. \quad (15)$$

Differentiating x_F^* yields

$$\left. \frac{\partial x_F^*}{\partial m} \right|_{m=0} = \frac{-N(a-c)b^3}{(2k+Nb+b)^2(4k^2+2bk(2+N)+b^2)}. \quad (16)$$

Substituting (13)–(16) into Equation (7), we obtain

$$\left. \frac{\partial W^*}{\partial m} \right|_{m=0} = \frac{N(a-c)^2b^3A}{(2k+Nb+b)^3(4k^2+2bk(2+N)+b^2)^2}, \quad (17)$$

where A is given by

$$A = [8k^3 + 12k^2b + 6kb^2 + b^3 + (2k^2b + kb^2)N] - kb^2N^2.$$

Since the denominator is positive, the RHS of (17) is negative (positive) if and only if $A < 0$ ($A > 0$). Since $k > 0$, the first term of A , which is within the bracket, is positive, and the second term of A is nonpositive. Note that when $N = 0$, $A > 0$ because the second term is zero. The first term is linear and the second term is quadratic with respect to N . Hence, in absolute value, the second term must exceed the first term when N is greater than some positive number and *vice versa*. **Q.E.D.**

Proof of Lemma 3: We prove that the firm facing the zero-profit condition produces on the downward-sloping part of the average cost curve. If the average cost is monotonically decreasing, this is obvious. Otherwise, let \hat{x} denote the output minimizing average cost of firm $(C(x) + f)/x$. We show that $x_F^1(m) \leq \hat{x}$. Suppose the contrary: $x_F^1(m) > \hat{x}$ holds true. Consider that one follower (firm i) deviates from the equilibrium strategy and chooses $x_i = \hat{x}$. Since the other firms' output is

constant, the deviation reduces the total output, resulting in an increase in the price (note that the leaders have chosen their outputs before observing the followers' outputs). Firm i 's profit is zero before the deviation from the zero profit condition (from (10)). The deviation increases the price and reduces firm i 's average cost and therefore, firm i makes a positive profit after the deviation. This is a contradiction. The same principle can be applied to the Cournot model to obtain $x_C^1 \leq \hat{x}$.

(i) We prove $x_F^1(m) = x_C^1$ by contradiction. Suppose that $x_F^1(m) < x_C^1$. From (10) (zero profit condition), we have $P(X^1(m)) = (C(x_F^1) + f)/x_F^1$ and $P(X_C^1) = (C(x_C^1) + f)/x_C^1$. These equations imply that $P(X^1(m)) > P(X_C^1)$ since the average cost $(C(x) + f)/x$ is decreasing when $x < \hat{x}$ and $x_F^1(m) < x_C^1 \leq \hat{x}$ by the supposition. Since $P' < 0$, $X^1(m) < X_C^1$ must hold. We compare the LHS in (11) with that in (3). Assumption 3 ensures that under the condition $X^1(m) < X_C^1$, the inequality $P'(X_C^1)x_C^1 + P(X_C^1) - C'(x_C^1) < P'(X^1(m))x_C^1 + P(X^1(m)) - C'(x_C^1)$ holds. Since $P' < 0$ and $C'' \geq 0$, under the condition $x_F^1(m) < x_C^1$, we have $P'(X^1(m))x_C^1 + P(X^1(m)) - C'(x_C^1) < P'(X^1(m))x_F^1(m) + P(X^1(m)) - C'(x_F^1(m))$. Thus, if (11) is satisfied, the LHS in (3) must be positive, which is a contradiction. Similarly, $x_F^1(m) > x_C^1$ leads to a contradiction.

(ii) From (10), we have $P(X^1(m)) = (C(x_F^1) + f)/x_F^1$ and $P(X_C^1) = (C(x_C^1) + f)/x_C^1$. In Lemma 3(i), we have shown that $x_F^1(m) = x_C^1$. Thus, $P(X^1(m)) = P(X_C^1)$. $X^1(m) = X_C^1$ is derived from $P' < 0$.

(iii) Suppose the contrary, i.e., $x_L^1(m) \leq x_C^1$. We compare the LHS in (11) with that in (2). Since $P' < 0$ and $C'' \geq 0$, under the condition $x_L^1(m) \leq x_C^1$, we have $P'(X_C^1)x_C^1 + P(X_C^1) - C'(x_C^1) \leq P'(X^1(m))x_L^1(m) + P(X^1(m)) - C'(x_L^1(m))$, where we use Lemma 3(ii). Since $nR'(X_L) \in (-1, 0)$ (see the discussion immediately after equation (5)), $P'(X^1(m))x_L^1(m) + P(X^1(m)) - C'(x_L^1(m)) < (1 + n^1(m)R')P'(X^1(m))x_L^1(m) + P(X^1(m)) - C'(x_L^1(m))$. Therefore, if (11) is satisfied, the LHS in (2) must be positive, which is a contradiction.

(iv) Lemma 3(iv) is derived from Lemmas 3(i)–(iii).

Q.E.D.

Proof of Proposition 4: Lemma 3(ii) indicates that the consumer surplus in the Stackelberg model ($m > 0$) is equal to that in the Cournot model ($m = 0$). The profits of all firms in the Cournot model are zero by the zero profit condition (10). In the Stackelberg model, the profits of all followers are also zero. Since x_L^1 is larger than x_C^1 by Lemma 3(iii) and is less than the output level that equates the marginal cost with price by (2), the average cost of each leader $(C(x_L^1) + f)/x_L^1$ is less than $P(X_C^1) = (C(x_C^1) + f)/x_C^1$, which is equal to the equilibrium price in the Stackelberg model. Therefore, the profits of all leaders are positive in the Stackelberg model. **Q.E.D.**

Proof of Lemma 4: Let $m > 0$. Given the leaders' aggregate outputs, the followers' actions are determined according to the following equations (at the second stage and third stage).

$$P(X^2(m))x_F^2(m) - C(x_F^2(m)) - f = 0, \quad (18)$$

$$P'(X^2(m))x_F^2(m) + P(X^2(m)) - C'(x_F^2(m)) = 0 \quad (19)$$

Since $P(X^2(m)) = P(X_C^2)$ (from Lemma 4(ii)), a leader's output (at the first stage) is obtained by

$$P(X_C^2(m)) = C'(x_L^2(m)). \quad (20)$$

(i) Using these conditions, we can prove $x_F^2(m) = x_C^2$ in the same way as in the proof of Lemma 3(i). Note that in the proof of Lemma 3(i), (10), (3) and (11) are all the conditions to prove the fact. In this proof, (18) and (19) correspond to (10) and (3), respectively. The condition in Cournot model (11) is common in both models 1 and 2.

(ii) The proof is similar to that of Lemma 3(ii).

(iii) In the Cournot model, $P(X_C^2) > C'(x_C^2)$. Thus, by (20) and $C'' > 0$, we have $x_L^2(m) > x_C^2$.

(iv) Lemma 3 (iv) is derived from Lemmas 3(i)–(iii). **Q.E.D.**

Proof of Lemma 5: (i) From Lemma 3(i) and Lemma 4(i), it is obvious that $x_F^1(m) = x_C^1 = x_C^2 = x_F^2(m)$.

(ii) Similarly, from Lemma 3(ii) and Lemma 4(ii), it is obvious that $X^1(m) = X_C^1 = X_C^2 = X^2(m)$.

(iii) As seen in (2), $x_L^1(m)$ satisfies $P(X^1(m)) > C'(x_L^1(m))$. As seen in (20), $x_L^2(m)$ satisfies $P(X^2(m)) = C'(x_L^2(m))$. Since from (ii), we have $X^1(m) = X^2(m)$, $C'(x_L^2(m)) > C'(x_L^1(m))$ holds. Therefore, we have $x_L^2(m) > x_L^1(m)$ from $C'' > 0$.

(iv) Lemma 5(iv) is derived from Lemmas 5(i)–(iii). **Q.E.D.**

Proof of Proposition 6: Lemma 5(ii) indicates that the consumer surplus in model 1 is equal to that in model 2 for all $m > 0$. In both models 1 and 2, the profits of all the followers are zero. Observe that from (2) and (20), we get $C'(x_L^2(m)) = P(X^2(m)) = P(X^1(m)) > C'(x_L^1(m))$. Since the profit maximizing behavior is marginal cost pricing given the price $P(X^2(m)) = P(X^1(m))$, the profits of m leaders in model 2 are larger than that in model 1. **Q.E.D.**

Proof of Proposition 7: (i) This is obvious from Corollary 1.

(ii) Since no followers enter the market when $m = N_C^1$, N_C^1 leaders engages in Cournot competition. Thus, $W^1(N_C^1) = W^1(0) = W_C^1$. Therefore, by Proposition 4, we have

$$\operatorname{argmax}_m W^1(m) \in (0, N_C^1).$$

(iii) From Lemma 4(ii), the leaders are price takers at the price $P(X_C^2)$ for all m as long as the followers enter the market. Thus, $x_L^2(m)$ is constant and $x_L^2(m) > x_C^2$ for all m . Therefore, the profit of each leader is constant and satisfies

$$P(X_C^2)x_L^2(m) - C(x_L^2(m)) - f > P(X_C^2)x_C^2 - C(x_C^2) - f = 0$$

for all m . Since the consumer surplus is constant and the profits of all followers are also zero for all m , $W^2(m)$ is increasing in m . **Q.E.D.**

B Examples of the integer problem

Suppose linear demand $P(X) = a - X$ and quadratic cost $C(x) = kx^2$, where $k \geq 0$. Further, we specify $a = 5, k = 3, f = 1$ and $m = 1$.

B.1 Example in model 2

If we neglect the integer problem in model 2, the following equilibrium outcomes are derived: $x_F^2(1) = \frac{1}{2}$ and $X^2(1) = \frac{3}{2}$. From $X = x_L + nx_F$, given x_L , the number of followers must satisfy

$$n(x_L) = 3 - 2x_L. \quad (21)$$

Since the leader acts like a price taker (i.e., $C'(x_L) = 7/2$), the equilibrium outcome of x_L is $x_L^2(1) = \frac{7}{12}$. Thus, from (21), we obtain $n^2(1) = 3 - \frac{7}{6}$. Hence, in the Cournot case, $N^2(0) = n^2(0) = 3$, while in the Stackelberg case with one leader, $N^2(1) = 1 + n^2(1) = 2 + 5/6$.

Now, suppose that n is always an integer. From (21), the number of followers in the Stackelberg case with one leader for a given x_L is

$$n^I(x_L) = \begin{cases} 2 & \text{if } 0 < x_L \leq \frac{1}{2} \\ 1 & \text{if } \frac{1}{2} < x_L \leq 1 \\ 0 & \text{if } 1 < x_L. \end{cases} \quad (22)$$

For fixed n , the reaction function of a follower and the aggregate output are given by

$$x_F(x_L; n) = \frac{5 - x_L}{7 + n}, \quad X(x_L; n) = x_L + nx_F(x_L; n) = \frac{5n + 7x_L}{7 + n}.$$

Thus, given x_L , the profit for the leader and a follower for a fixed n are given by

$$\pi_L(x_L; n) = \frac{7(5 - x_L)}{7 + n}x_L - 3x_L^2 - 1, \quad (23)$$

$$\pi_F(x_L, n) = \frac{4(x_L - 5)^2}{(7 + n)^2} - 1. \quad (24)$$

From (22) and (23), we obtain the leader's profit:

$$\pi_L(x_L; n^I(x_L)) = \begin{cases} x_L(35 - 34x_L)/9 - 1 & \text{if } 0 < x_L \leq \frac{1}{2} \\ x_L(35 - 31x_L)/8 - 1 & \text{if } \frac{1}{2} < x_L \leq 1 \\ x_L(35 - 28x_L)/7 - 1 & \text{if } 1 < x_L. \end{cases}$$

We depict $\pi(x_L; n^I(x_L))$ in Figure 6.

Insert Figure 6 here.

From this figure, we observe that the best strategy for the leader is $x_L = 35/62$. In this case, the number of followers entering is $n = 1$. Using this equilibrium outcome, we can calculate welfare:

$$0.5X(35/62; 1)^2 + \pi_L(35/62; 1) + \pi_F(35/62; 1) \approx 1.090.$$

On the other hand, in the Cournot case for this example, the integer problem does not arise since $n^2(0)$ is exactly 3. Thus, we have $X = 3/2$; the profit of each firm is 0 and welfare is 1.125 –which exceeds the welfare in the Stackelberg case.

B.2 Example in model 1

Suppose the same specification as above. The profit of the leader for a fixed n is same as above:

$$\pi_L(x_L; n) = \begin{cases} x_L(35 - 34x_L)/9 - 1 & \text{if } n = 2 \\ x_L(35 - 31x_L)/8 - 1 & \text{if } n = 1 \\ x_L(35 - 28x_L)/7 - 1 & \text{if } n = 0. \end{cases}$$

Thus, the leader chooses $x_L = 35/56$ if $n = 0$, $x_L = 35/62$ if $n = 1$ and $x_L = 35/68$ if $n = 2$. Substituting these results into (24), we obtain $\pi_F(35/62, 1) \approx 0.229$ if $n = 1$ and $\pi_F(35/68, 2) \approx -0.006$ if $n = 2$. Since the first entrant can earn positive profit but the second entrant cannot, $n = 1$ in the equilibrium. Therefore, the equilibrium outcome is the same as that in model 2 and the same integer problem arises.

References

- Brandão, A., Castro, S., “State-owned enterprises as indirect instruments of entry regulation,” *Journal of Economics* 92 (2007), 263-274.
- Daughety, A. F., “Beneficial concentration,” *American Economic Review* 80 (1990), 1231-1237.
- Davidson, C., Mukherjee, A., “Horizontal mergers with free entry,” *International Journal of Industrial Organization* 25 (2007), 157-172.
- Etro, F., “Innovation by leaders,” *Economic Journal* 114 (2004), 281-303.
- Etro, F., “Aggressive leaders,” *Rand Journal of Economics* 37 (2006), 146-154.
- Etro, F., *Competition, Innovation, and Antitrust: A Theory of Market Leaders and Its Policy Implications* (Springer: New York, 2007).
- Etro, F., “Stackelberg competition with endogenous entry,” *Economic Journal* 118 (2008), 1670-1697.
- Huck, S., Müller, W., Normann, H. T., “Stackelberg beats Cournot: On collusion and efficiency in experimental markets,” *Economic Journal* 111 (2001), 749-765.
- Ino, H., Kawamori, T., “Oligopoly with a large number of competitors: Asymmetric limit result,” *Economic Theory* 39 (2009), 331-352.
- Lahiri, S., Ono, Y., “Helping minor firms reduces welfare,” *Economic Journal* 98 (1988), 1199-1202.
- Lahiri, S., Ono, Y., “The role of free entry in an oligopolistic Heckscher-Ohlin model,” *International Economic Review* 36 (1995), 609-624.
- Lahiri, S., Ono, Y., “Foreign direct investment, local content requirement, and profit taxation,” *Economic Journal* 108 (1998), 444-457.
- Levin, D., “Stackelberg, Cournot and collusive monopoly: Performance and welfare comparisons,” *Economic Inquiry* 26 (1988), 317-330.
- Mankiw, N. G., Whinston, M. D., “Free entry and social inefficiency,” *Rand Journal of Economics* 17 (1986), 48-58.

- Marjit, S., Mukherjee, A., "International outsourcing and R&D: Long-Run implications for consumers," *Review of International Economics* 16 (2008), 1010-1022.
- Matsumura, T., "Partial privatization in mixed duopoly," *Journal of Public Economics* 70 (1998), 473-483.
- Matsumura, T., "Quantity-setting oligopoly with endogenous sequencing," *International Journal of Industrial Organization* 17 (1999), 289-296.
- Matsumura, T., "Consumer-benefiting exclusive territories," *Canadian Journal of Economics* 36 (2003), 1007-1025.
- Matsumura, T., Kanda, O., "Mixed oligopoly at free entry markets," *Journal of Economics* 84 (2005), 27-48.
- Mukherjee, A., Mukherjee, S., "Excess-entry theorem: The implications of licensing," *Manchester School* 76 (2008), 675-689.
- Ono, Y., "The equilibrium of duopoly in a market of homogeneous goods," *Economica* 45 (1978), 287-295.
- Ono, Y., "Price leadership: A theoretical analysis," *Economica* 49 (1982), 11-20.
- Sherali, H., "A multiple leader Stackelberg model and analysis," *Operations Research* 32 (1984), 390-404.
- Suzumura, K., Kiyono, K., "Entry barriers and economic welfare," *Review of Economic Studies* 54 (1987), 157-167.

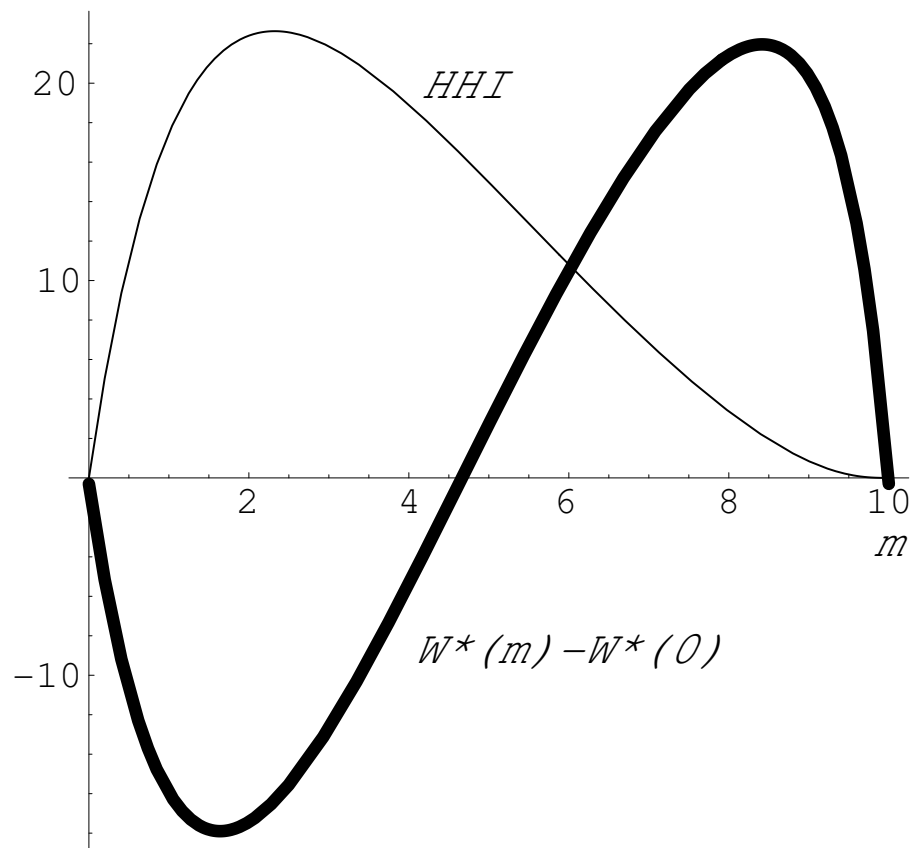


Figure 1: Welfare Comparison when the number of leaders changes for fixed $N = m + n$: The thick curve represents $W^*(m) - W^*(0)$. The thin curve represents the HHI (scaled to adjust the figure). The horizontal axis represents m and its origin is 0. The origin of the vertical axis is 0 when it measures welfare and $1/N$ when it measures HHI.

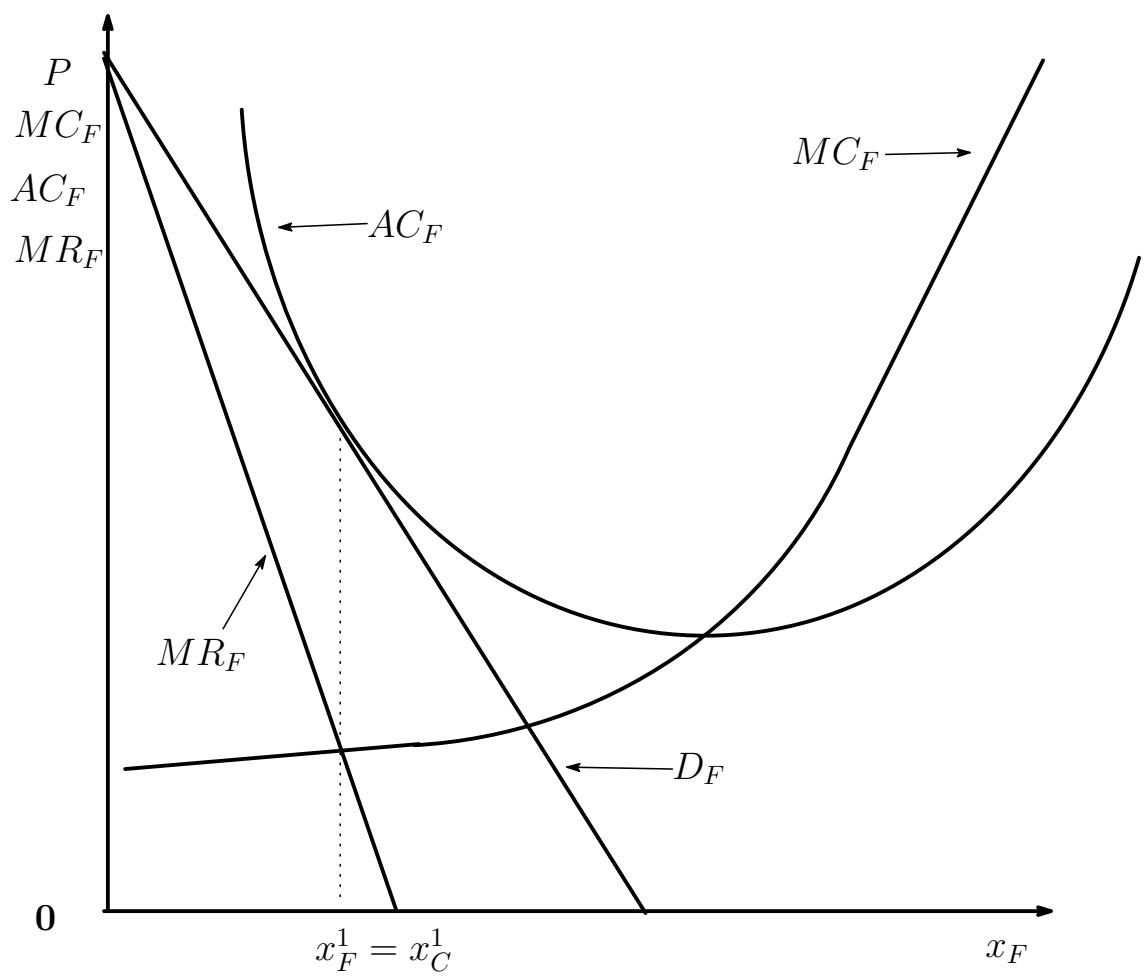


Figure 2: The equilibrium outputs in the long run: D_F : residual demand of a follower. AC_F : average cost of a follower. MR_F : marginal revenue of a follower. MC_F : marginal cost of a follower.

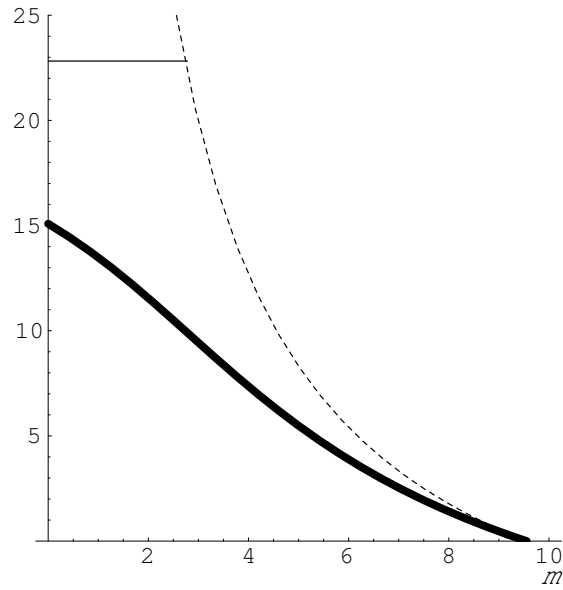


Figure 3: The equilibrium outputs in the long run: The thick curve represents $x_L^1(m) - x_F^1(0)$. The thin curve represents $x_L^2(m) - x_F^2(0)$. The dotted curve represents $X_C^2/m - x_F^2(0)$.

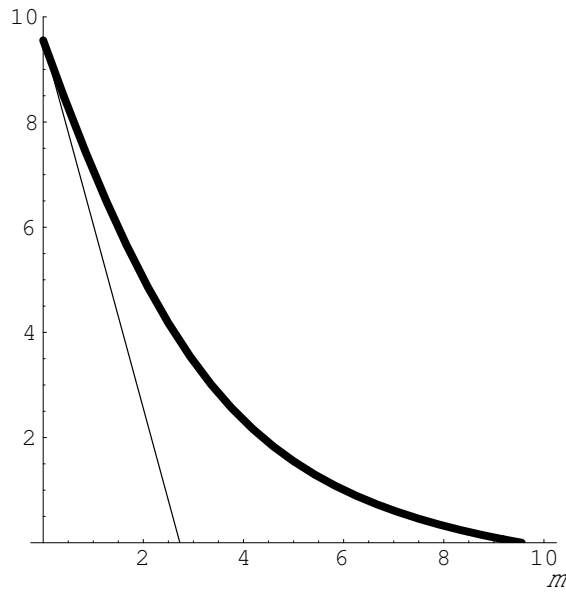


Figure 4: The equilibrium number of followers in the long run: The thick curve represents $n^1(m)$. The thin curve represents $n^2(m)$.

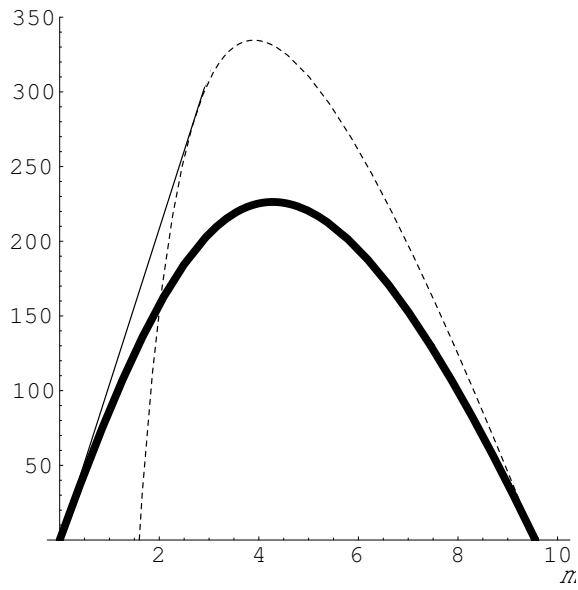


Figure 5: The equilibrium welfare in the long run: The thick curve represents $W^1(m) - W^1(0)$. The thin curve represents $W^2(m) - W^2(0)$. For the dotted curve, see footnote 20.

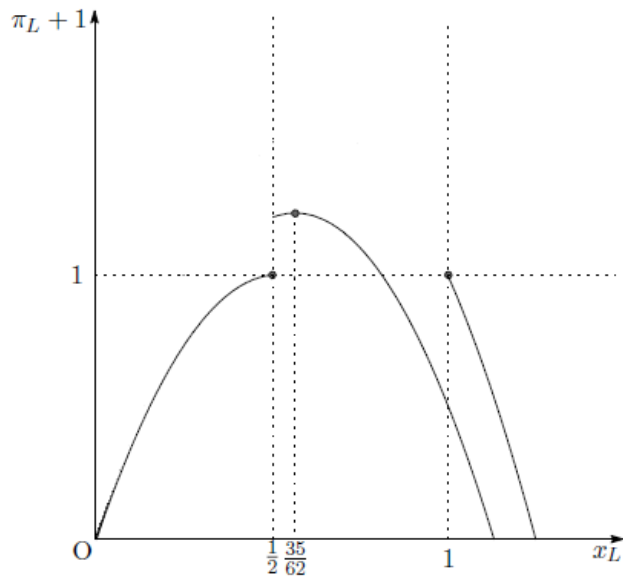


Figure 6: Profit of a leader under the integer constraint