

How Many Firms Should Be Leaders?: Beneficial Concentration Revisited

Hiroaki Ino

Graduate School of Economics, University of Tokyo

and

Toshihiro Matsumura

Institute of Social Science, University of Tokyo

December 20, 2005

Abstract

We investigate how the number of Stackelberg leaders and followers affects economic welfare. First, we discuss the model with m leaders and $N-m$ followers compete in a homogeneous good market. Daughety (1990) finds that, under constant marginal costs, the Stackelberg model ($m \in (0, N)$) yield a higher welfare than the Cournot model ($m = 0$ or $m = N$.) He also shows that beneficial concentration takes place when m is close to 0 or N . We find that beneficial concentration always takes place and an increase of HHI increase welfare when m is close N , but it is not true when m is close to 0 under increasing marginal costs. We also find that the Stackelberg model can yield a smaller welfare when m is small. We also consider the model with the free entry of followers. We find that the existence of leaders always improve welfare and beneficial concentration always takes place even when marginal costs are increasing.

JEL classification numbers: L13, L40

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

1 Introduction

Market concentration measured by the Herfindahl-Hirschman Index (HHI) has played quite an important role in the context of antitrust legislations and economic regulations in many countries. It has been received much attention from the wide range of researchers including both economists and law scholars for many years and is increasingly important.¹

Market concentration and economic welfare are affected by both the number of the firms and the asymmetries among firms. The effect of the number of firms are usually straightforward. Typically, an increase of the number of firms improves economic welfare and reduces HHI in the symmetric equilibria. On the contrary, the effect of the asymmetries is usually complicated. Daughety (1990) investigate this problems by considering asymmetric roles among firms. He formulates a model where m Stackelberg leaders and $N-m$ Stackelberg followers compete in a homogenous good market and discuss the relationship between m and economic welfare. Using a linear demand function and a constant marginal cost he finds the inverse U-shape relationship between m and economic welfare. He also finds that the relationship between HHI and economic welfare is complicated and economic welfare can increase in HHI.

In this paper we takes a close look at this problem. We extend his model for two direction. First, we consider more general demand and cost functions. We find that the existence of the small number of followers always improves welfare, like Daughety (1990). This result is robust because it holds true under general cost and demand functions. We also find that the existence of the small number of leaders always improves welfare, like Daughety (1990), if marginal cost is constant. This result is not robust because it does not hold true under increasing marginal costs. As a result, the relationship between HHI and economic welfare becomes much more complicated when marginal cost is increasing.

Next, we endogenize the numbler of the followers by considering free entries of the followers. In contrast to the case of exogenous number of the firms, we find that the existence of the leaders always improves welfare.

The remainder of this paper is organized as follows: Section 2 formulates the model with

¹For example, Fare Trade Commission in Japan finally published a merger guideline based on HHI in 2003 though it denied to use HHI for many years.

exogenous number of the firms. Section 3 investigates equilibrium outcomes and presents results. Section 4 examines the model with free entries. Section 5 concludes the paper. All proofs of Lemmas and Propositions are presented in the Appendix.

2 The Model

We set up a N firms oligopoly model. 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 $f \in \mathbb{R}_+$ is the fixed 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 following standard assumptions.

Assumption 1 $P(X)$ is twice differentiable. $P'(X) < 0, \forall X$ such that $P(X) > 0$. $\exists \bar{X}, P(\bar{X}) = 0$.²

Assumption 2 $c(x)$ is twice differentiable. $c'(x) > 0, c''(x) \geq 0, \forall x \geq 0$.

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

The game runs as follows. In the first stage $m(\leq N)$ firms (Stackelberg leaders) independently choose their outputs. In the second stage, $n (\equiv N - m)$ firms (Stackelberg followers) independently choose their outputs, after observing leaders' outputs. The model corresponds to the Cournot model when $m = 0$ or $m = N$.³

Henceforth, we call the equilibrium outcome symmetric when all leaders (followers) choose the same output level. Needless to say, the leader's output is different from the follower's even in the symmetric equilibrium. We now make an assumption on the equilibrium outcomes.

Assumption 4 The model has the unique equilibrium $\forall m \in [0, N]$ and $N > 0$, the equilibrium is symmetric and all firms produces positive outputs.

²This assumption about \bar{X} is posited just to make the functions in our analysis differentiable and is not essential for our results. Indeed, we can show all the propositions without this assumption except for the statements using derivatives.

³In Lemma 2 we show that in fact each follower's (leader's) output converges to each firm's output in the Cournot model where N firms choose their output simultaneously and independently when $m \rightarrow 0$ ($m \rightarrow N$).

3 Welfare

In this section, we analyze the relationship between m and social welfare (consumer surplus plus profits of all firms). The social welfare W is defined as

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

Let $x_L^*(m)$ denote a leader's equilibrium output and $x_F^*(m)$ denote a follower's. These are obtained as the solution of following system of equations (equation (2) is a leader's first order condition and equation (3) is a follower's).

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

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

where $X^*(m) = mx_L^*(m) + nx_F^*(m)$ and $X_L^*(m) = mx_L^*(m)$.⁴ $R(X_L)$ represents a follower's reaction to leaders' action, that is, a follower's output given total output of leaders X_L . Note $R(X_L)$ is obtained from the follower's first order condition,

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

Differentiating this 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 have $nR'(X_L) \in (-1, 0)$. This implies that an increase in the output of the leader decreases the total output of the followers, and that it increases the total output of all firms (including the leader and the followers).

The equilibrium aggregate output X^* is given by

$$X^*(m) = mx_L^* + nx_F^*. \quad (6)$$

The equilibrium social welfare W^* is given by

$$W^*(m) = \int_0^{X^*} P(X) dX - mc(x_L^*) - nc(x_F^*) - (m + n)f \quad (7)$$

⁴Note when $m = 0$ and $m = N$, (3) and (2) respectively corresponds to the first order condition of 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).

We present a result highlighting the difference between the Cournot outcome (where $m = 0$ or $m = N$) and the Stackelberg outcome (where $m \in (0, N)$).

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

We now present results on 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 $c''(x) = 0$ (i.e., linear demand), (i) $\frac{\partial W^*}{\partial m} \Big|_{m=0} > 0$, (ii) $\frac{\partial W^*}{\partial m} \Big|_{m=N} < 0$ and (iii) $W^*(m) > W^*(0) \forall m \in (0, N)$.*

Stackelberg leaders push up the aggregate output as seen in Lemma 1-(iii). This brings total benefit in the industry closer to that in competitive market – resulting in the higher welfare. Daughety (1990) have already shown this result under the assumptions of linear demand and constant marginal costs. Various concentration measures, such as HHI, rise when the market structure changes from Cournot-type to Stackelberg-type. In other words, Stackelberg-type competition causes concentration to the leaders in a industry. Despite of the traditional belief, such kind of concentration enhances welfare. Daughety call that “beneficial concentration”.

We now discuss the robustness of Proposition 1 by considering increasing marginal costs. The following Proposition 2 states that Proposition 1(ii) is robust, whereas Proposition 1(i) and (iii) are not.

Proposition 2 *Suppose that Assumptions 1-4 are satisfied. Then, (i) $\frac{\partial W^*}{\partial m} \Big|_{m=N}$ can be negative when $c'' > 0$ and (ii) $\frac{\partial W^*}{\partial m} \Big|_{m=N}$ is always negative.*

Proposition 2(ii) implies the following result.

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

Proposition 2(ii) indicates that introducing a small number of followers into the Cournot model always improves welfare. Since introducing a small number of followers into the Cournot model increases HHI, this result implies that beneficial concentration can occur under general cost and

demand functions. However, Proposition 1(i) and (iii) are not robust. In the next section we show it using examples of quadratic costs.

We now explain the intuition behind the asymmetry. We explain why Proposition 1 (ii) is robust while Proposition 1 (i) is not under increasing marginal cost. Or equivalently, we explain why introducing a small number of followers into the Cournot model always improves welfare while introducing a small number of the leaders can reduce welfare.

First, we present a supplementary result, which help us to understand the intuition. Let x_C^* denote the equilibrium output of each firm when N firms produce simultaneously (the Cournot model) .

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

$$(i) \ x_C^* = x_L^*(N) = \lim_{m \rightarrow N-0} x_L^*(m) = \lim_{m \rightarrow N-0} x_F^*(m) \equiv x_F^*(N),$$

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

Both Stackelberg leader's output and follower's output converge to the Cournot output when m is close to N , whereas the leader's output does not converge to the Cournot output (the follower's output converges to the Cournot one).

Next, we explain the intuition behind the asymmetry in Proposition 2. Suppose that a small number of the leader is introduced. The leader's output is strictly larger than the follower's. Since the marginal cost is increasing, total production costs are minimized when all firm produce the same output. Thus, introducing a smaller number of leaders reduces the production efficiency. On the other hand, introducing a smaller number of leaders increases total output and increases consumer surplus. This competition acceleration effect improves welfare. The former welfare-reducing effect can dominate the latter welfare-improving effect.⁵ This is why Proposition 2(i) holds. Note that the former effect does not exist when marginal cost is constant.

On the contrary, production inefficiency caused by introducing a smaller number of followers is insignificant. As Lemma 2(i) indicates, both the leader's output and follower's output converges to the Cournot one when m is close to N . Thus, production inefficiency is insignificant and it is dominated by the welfare-improving effect of competition acceleration. This is why

⁵Similar trade off of two effects are discussed in many contexts. See, among others, Lahiri and Ono (1988, 1998) and Matsumura (1998, 2003).

Proposition 1(ii) is robust.

4 Linear Demand and Quadratic Cost Functions

In this section we specify demand and cost functions. Suppose the inverse demand function is linear, $P(X) = a - X$, and the cost function is quadratic, $c(x) = kx^2$, where $a > 0$ and $k \geq 0$. Note if $k = 0$ ($k > 0$), we have a constant (increasing) marginal cost.⁶

Under these specialization, we obtain $R'(X_L) = -1/(2k + n + 1)$, where n is the number of followers and $n \equiv N - m$. Therefore, the first order conditions (2) and (3) turn out to be

$$a - (2k + m + 1)x_L^* - nx_F^* + \frac{nx_L^*}{2k + n + 1} = 0 \quad (8)$$

$$a - (2k + n + 1)x_F^* - mx_L^* = 0. \quad (9)$$

By solving these, we obtain

$$x_L^*(m) = \frac{a(2k + 1)}{4k^2 + 2(m + n + 2)k + m + 1}, \quad (10)$$

$$x_F^*(m) = \frac{a(4k^2 + 2(n + 2)k + 1)}{(2k + n + 1)(4k^2 + 2(m + n + 2)k + m + 1)}. \quad (11)$$

By substituting (10)(11) into (7), we obtain the social welfare in equilibrium.

$$W^* = \frac{m^2(2k + 1)^2 + 2m(2k + 1)(2k^2 + (2n + 3)k + 1)}{2(4k^2 + 2(m + n + 2)k + m + 1)^2} a^2 + \frac{n(2k + n + 2)(4k^2 + 2(n + 2)k + 1)^2}{2(2k + n + 1)^2(4k^2 + 2(m + n + 2)k + m + 1)^2} a^2 - (m + n)f. \quad (12)$$

We then show that leadership can reduce welfare if and only if the cost function is strictly concave.

Proposition 3 *There exists $N (> 0)$ such that $\frac{\partial W^*}{\partial m} \Big|_{m=0} < 0$ if and only if $k > 0$.*

This proposition implies the following result.

Corollary 2 *Suppose that $k > 0$. Then, there exists $m \in (0, N)$ and $N > 0$ such that $W^*(m) < W^*(0)$.*

⁶In our specification, the constant part of marginal cost is 0. However, our analysis is robust even if constant part of marginal cost $\gamma > 0$, i.e. $c(x) = \gamma x + kx^2$.

We present a numerical example describing such a situation.

Example 1 Suppose $a = 100$, $k = 0.2$, $f = 20$, $N = 10$. Using (12), we can simulate how the ratio of leaders in a industry effects on social welfare. Figure 1 depicts a change of welfare when m changes in $[0, 10]$ given $N = m + n = 10$. While, as predicted in Proposition 3-(i), the welfare goes down when the ratio of leaders is small, as predicted in Proposition 2-(i), the welfare goes up when the ratio of leaders is large comparing the Cournot case. We also depicts how HHI changes in the same situation. HHI is defined as the sum of (each firm's market share)² (multiplied by 1000) In our model it is given by

$$HHI = 1000 \left(m \left(\frac{x_L^*(m)}{X^*(m)} \right)^2 + n \left(\frac{x_F^*(m)}{X^*(m)} \right)^2 \right). \quad (13)$$

Figure 1 illustrates this example. From Figure 1 we find that HHI provides a good insight for welfare's change except for around $m = N$ in this example. Decrease (increase) in HHI contributes to increase (decrease) in welfare. In the constant marginal cost's case that Daughety (1990) investigate, the range where HHI is increasing but the welfare is increasing is wider than this example. Therefore, the case where HHI works well seems exception. If we consider the increasing marginal cost, however, increase in concentration is more likely to be harmful. As seen in the above example and Proposition 3-(i), this harmfulness is more likely true especially when there is a few leaders, which is the situation the antimonopoly policy is considered more seriously.

We must note strict convexity of cost function does not immediately cause the welfare-deteriorating activities by a few leaders. The following two examples clarify this point.

Example 2 Suppose $a = 100$, $f = 20$, $m = 1$, $N = 10$. Using (12), we can simulate how convexity of cost function effects on social welfare. Figure 2 depicts the difference between Stackelberg-type model's welfare ($m \in (0, N)$) and Cournot-type model's welfare ($m = 0$ or N) when k changes in $[0, 1]$ given a single leader and $N = m + n = 10$. As predicted in Proposition 1-(ii), the $W^*(m)$ exceeds $W^*(0)$ when $k = 0$. Though we can easily find the case where $W^*(0)$ exceeds $W^*(m)$, there is a case the $W^*(m)$ exceeds $W^*(0)$ even if $k > 0$.

Example 3 Suppose $a = 100$, $f = 20$, $k = 0.01$. Figure 3 depicts the relationship between $W^{*'}(0)$ and N . As is discussed in the last paragraph in Section 3, introducing a small number of

leader has two welfare effects. One is welfare-reducing effect caused by asymmetric production level and the other is welfare-improving effect caused by accelerated competition. When N is large, the market is competitive and the price is close to the marginal cost. Thus, the welfare-improving effect caused accelerated competition is small, so the welfare-reducing effect can dominate this small welfare-improving effect.

5 Free Entries of Followers

In this section, we provide a likely long-run story in which the existence of leaders always enhance the social welfare even if the cost function is increasing. In other words, the beneficial concentration always occurs as in the short-run constant marginal cost's case.

Suppose that m firms have special ability to produce quickly and can produce other n slow firms. In other words, m firms become leaders and n firms become followers. The number of skilled firms, m is given exogenous and the number of unskilled firms are determined by the free entries.

In the first stage, followers chooses whether or not to enter the market. In the second stage m leaders and n followers face the same competition formulated in Section 2. Only the difference is n is given exogenous in the second stage but it is endogenously determined in the first stage. We assume that market is sufficiently profitable that $n > 0$ in equilibrium.

To make our long-run analysis meaningful, we add the assumption that guarantee an U-shaped average cost curve.

Assumption 5 $c''(x) > 0, \forall x > 0$ and $f > 0$.

Let $n^*(m)$ be the equilibrium number of followers. It is obtained from the following zero profit condition of followers.

$$P(X(m))x^F(m) - c(x^F(m)) - f = 0 \quad (14)$$

where $X(m) = mx^L(m) + n^*(m)x^F(m)$ and $x^L(m)$ ($x^F(m)$) represents the long-run equilibrium output of a leader (follower). $N(m)$ represents the total number of firms and is given by $N(m) = m + n^*(m)$. Using the expressions in previous section, we have the relations $x^L(m) = x_L^*(m)$ and $x^F(m) = x_F^*(m)$ given $N = N(m)$. The case when $m = 0$ corresponds to the Cournot model.

The two models' equilibrium outcomes have the following relationship.

Lemma 3 *Suppose that Assumptions 1-4 are satisfied. Then, $\forall m > 0$, (i) $x^F(m) = x^F(0)$, (ii) $X(m) = X(0)$, (iii) $x^L(m) > x^F(0)$, and (iv) $N(m) < N(0)$.*

Lemma 3(i) implies the following result.

Corollary 3 *Consumer surplus does not depend on m .*

Lemma 3(i) implies that the 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 unchanged unless the follower's output changes. This yields Corollary 3.

Let $W(m)$ denote the long-run equilibrium welfare given m . Note $W(0)$ is the equilibrium welfare in the Cournot-type model. Then the welfare comparison yields a striking contrast to the short-run case.

Proposition 4 *Suppose that Assumptions 1-4 are satisfied. Then, $\forall m > 0$, $W(m) > W(0)$*

Introducing a leader yields the asymmetry among firms and it increases HHI. In addition, introducing a leader reduces the total number of firms and it also increases HHI. However, it improves welfare. In other words, beneficial concentration always takes place in this context.⁷

We explain the intuition behind the result. Followers choose the output level which is lower than the average cost minimizing level. Introducing a leader into the Cournot model, the leader chooses the larger output which reduces the average cost and improves production efficiency. In contrast to the short-run case, introducing a leader does not reduce each follower's output. Instead, it reduces the number of entering firms, saving wasteful entry costs. This is why a leadership always improves welfare at the free-entry equilibrium.⁸

6 Concluding Remarks

In this paper we take a close look at welfare in the model with multiple Stackelberg leaders and reexamine the relationship between HHI and welfare. We find that beneficial concentration

⁷Long-run and short-run analysis often yields different implications in oligopoly models. See, among others, Lahiri and Ono (1995) and Matsumura and Kanda (2005).

⁸For the discussion on wasteful entry costs, see among others, Mankiw and Whinston (1986) and Suzumura and Kiyono (1987).

always takes place when a smaller number of follower is introduced. On the contrary, introducing a smaller number of leaders can reduce welfare. It always increase HHI and the change in welfare can go both ways. We also find that the latter result does not hold if we consider free entries of the followers and beneficial concentration always takes place when we introduce a leader into the Cournot model.

In this paper we assume that all firms have identical cost function and only the difference of roles among firms yield the asymmetry among the firms. If we introduce the cost differences among the firms, the analysis becomes much complicated. A leadership be the firm with a lower (higher) marginal cost more likely improves welfare. Then, whether the firm with lower cost or the firm with higher cost more likely is quite important. Extending our analysis to this direction remains for future research.⁹

⁹For the discussion on endogenous role among the firms with heterogeneous costs, see Ono (1978,1982). For the discussion of endogenous role of N-firm oligopoly, see among others, Matsumura (1999). For the discussion on the role of firms based on experimental studies, see Huck et al. (2001).

APPENDIX

Proof of Lemma 1 : By the first order condition (3) with $m = 0$,

$$P'(X^*(0))x_F^*(0) + P(X^*(0)) - c'(x_F^*(0)) = 0 \quad (15)$$

where $X^*(0) = Nx_F^*(0)$. Thus, $R(mx_F^*(0))$ satisfies (4) when $R(mx_F^*(0)) = x_F^*(0)$. Then, the uniqueness of Cournot equilibrium yields $R(mx_F^*(0)) = x_F^*(0)$.

(i) Suppose contrary $x_L^*(m) \leq x_F^*(0)$. Since $nR' > -1$, $X^*(m) = mx_L^*(m) + nR(mx_L^*(m)) \leq mx_F^*(0) + nR(mx_F^*(0)) = X^*(0)$. We compare the LHS of (15) with that of (2).

$$\begin{aligned} & P'(X^*(0))x_F^*(0) + P(X^*(0)) - c'(x_F^*(0)) \\ & \leq P'(X^*(m))x_F^*(0) + P(X^*(m)) - c'(x_F^*(0)) \quad (\because X^*(m) \leq X^*(0), \text{ Assumption 3}) \\ & \leq P'(X^*(m))x_L^*(m) + P(X^*(m)) - c'(x_F^*(0)) \quad (\because x_L^*(m) \leq x_F^*(0), P' < 0) \\ & \leq P'(X^*(m))x_L^*(m) + P(X^*(m)) - c'(x_L^*(m)) \quad (\because x_L^*(m) \leq x_F^*(0), c'' > 0) \\ & < (1 + R')P'(X^*(m))x_L^*(m) + P(X^*(m)) - c'(x_L^*(m)) \quad (\because -1 < R' < 0, P' < 0) \end{aligned}$$

Therefore, if (15) is satisfied, LHS of (2) must be strictly positive – a contradiction.

(ii)(iii) We have already proven $x_L^*(m) > x_F^*(0)$ in (i). Since $R' < 0$, $x_F^*(m) = R(mx_L^*(m)) < R(mx_F^*(0)) = x_F^*(0)$. Since $nR' > -1$, $X^*(m) = mx_L^*(m) + nR(mx_L^*(m)) > mx_F^*(0) + nR(mx_F^*(0)) = X^*(0)$. **Q.E.D.**

Proof of Proposition 1 : Statement (i)(ii) is the corollary of statement (iii). Therefore, we provide the proof of (iii). We denote the constant marginal cost by $\gamma \geq 0$. Then, $\forall N$, $\forall m \in (0, N)$,

$$W^*(m) - W^*(0) = \int_{X^*(0)}^{X^*(m)} P(X)dX - \gamma(X^*(m) - X^*(0)) > 0.$$

The last inequality comes from lemma 1-(iii) and the fact the price exceeds the marginal cost at $X^*(m)$. **Q.E.D.**

Proof of Proposition 2 : We show statement (i) in section 4. Therefore, we provide the proof of (ii) here. Note, by Lemma 2(i) and the Implicit Function Theorem, x_F^* and x_L^* is differential at $m = N$. Differentiability at $m = N$ guarantee the following decomposition.

$$\frac{\partial W^*}{\partial m} \Big|_{m=N} = [Px_L^*(N) - c(x_L^*(N))] - [Px_F^*(N) - c(x_F^*(N))] + N[P - c(x_L^*(N))] \frac{\partial x_L^*}{\partial m} \Big|_{m=N}$$

The fact $x_L^*(N) = x_F^*(N)$ from Lemma 2(i) cancels out the first and the second term of RHS. Lemma 1(i) implies the last term is strictly negative. Thus, the LHS is also strictly negative.

Q.E.D.

Proof of Lemma 2 : We regard the LHS of (2) as a function $F(m, x_L^*(m), x_F^*(m))$ and the LHS of (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 $\forall m \in (0, N)$, $F(m, x_L^*(m), x_F^*(m)) = 0$, $\lim_{m \rightarrow +0} F(m, x_L^*(m), x_F^*(m)) = 0$. 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 system of equations (2)(3) when $m = 0$. Thus, by definition of $x_F^*(0)$, $\lim_{m \rightarrow +0} x_F^*(m) = x_F^*(0)$. By taking $m \rightarrow N - 0$, similarly we obtain $F(N, \lim_{m \rightarrow N-0} (x_L^*(m), x_F^*(m))) = 0$ and $G(N, \lim_{m \rightarrow N-0} (x_L^*(m), x_F^*(m))) = 0$. Therefore, $\lim_{m \rightarrow N-0} (x_L^*(m), x_F^*(m))$ is a solution of system of equations (2)(3) when $m = N$. Thus, by definition of $x_L^*(N)$, $\lim_{m \rightarrow N-0} x_L^*(m) = x_L^*(N)$.

Take $m = 0$. Then, by (3), $x_F^*(0)$ comes from $P'(Nx_F^*)x_F^* + P(Nx_F^*) = c'(x_F^*)$. Then, given $x_F^*(0)$, $x_L^*(0)$ is determined according to (2), i.e. $(1 + NR'(0))P'(Nx_F^*)x_L^* + P(Nx_F^*) = c'(x_L^*)$. By $(1 + NR'(0))P'(Nx_F^*) < 0$ and $c'' \geq 0$, the LHS of this equation is strictly decreasing in x_L^* and the RHS is increasing. Since $0 < (1 + NR'(0)) < 1$, $x_L^*(0) > x_F^*(0)$. Next, take $m = N$. Note when $m = N$, nR' vanish from (2). Thus, $x_F^*(N)$ comes from $P'(Nx_L^*)x_L^* + P(Nx_L^*) = c'(x_L^*)$. Then, given $x_L^*(N)$, $x_F^*(N)$ is determined according to (3), i.e. $P'(Nx_L^*)x_F^* + P(Nx_L^*) = c'(x_F^*)$. By $P'(Nx_L^*) < 0$ and $c'' \geq 0$, the LHS of this equation is strictly decreasing in x_F^* and the RHS is increasing. $x_F^*(N)$ satisfies (3) only when $x_L^*(N) = x_F^*(N)$. **Q.E.D.**

Proof of Proposition 3 : The contrapositive of Proposition 1-(i) yields $k > 0$ if $\exists N$, $\frac{\partial W^*}{\partial m} \Big|_{m=0} < 0$. Next, we prove the reverse, i.e. if $k > 0$, $\exists N$, $\frac{\partial W^*}{\partial m} \Big|_{m=0} < 0$. By differentiating (12) given $N = m + n$ and evaluating at $(m, n) = (0, N)$, we obtain

$$\begin{aligned} & \frac{\partial W^*}{\partial m} \Big|_{m=0} \\ &= \frac{a^2 N (32k^5 + (8k^4 + k)(10 + 3N) + (3k^3 + 2k^2)(20 + 9N) + 1 - 2k^2(N^2 + N^3) - kN^2)}{(1 + 2k + N)^3 (1 + 4k + 2k(2 + N))^3}. \end{aligned}$$

Since the denominator is positive, the RHS of above equation is strictly negative if and only if

$$(24k^4 + 36k^3 + 18k^2 + 3k)N + 32k^5 + 80k^4 + 80k^3 + 40k^2 + 10k + 1 < 2k^2N^3 + (2k^2 + k)N^2.$$

The LHS is linear and the RHS is cubic with respect to N . Since k is constant and strictly positive, the RHS must exceed the LHS when we take enough large N . **Q.E.D.**

Proof of Lemma 3 : Superscript C denote the equilibrium outcome in the Cournot model. Let $X(m) = X^S$ for $m > 0$. (i) Let x^* denote the output minimizing average cost of the firm $(c(x) + f)/x$. First we show that $x^F \leq x^*$. Suppose contrary that $x^F > x^*$. Suppose that one follower (firm i) deviates from the equilibrium strategy and chooses $x_i = x^*$. Since the other firms' output is constant, the deviation reduces the total output, resulting in the rise of the price (Note that in the Stackelberg-type model the leaders have chosen their output before observing the followers' outputs). By (14), firm i's profit is zero before the deviation. The deviation raises the price and reduces firm i's average cost, so firm i obtains strictly positive profits after the deviation – a contradiction. For the same reason we have $x^C \leq x^*$.

We now prove $x^F = x^C$ by contradiction. Suppose that $x^F < x^C$. By (14), $P(X^S) = c(x^F)/x^F$ and $P(X^C) = c(x^C)/x^C$. Since $x^F < x^C \leq x^*$, the U-shaped average cost curve implies $P(X^S) > P(X^C)$. Since $P' < 0$, we must have $X^S < X^C$. We compare LHS of (15) with that of (3).

$$\begin{aligned} P'(X^C)x^C + P(X^C) - c'(x^C) &< P'(X^S)x^C + P(X^S) - c'(x^C) \quad (\because X^S < X^C, \text{ Assumption 3}) \\ &< P'(X^S)x^F + P(X^S) - c'(x^C) \quad (\because x^F < x^C \text{ and } p' < 0) \\ &< P'(X^S)x^F + p(X^S) - c'(x^F) \quad (\because x^F < x^C \text{ and } c'' > 0) \end{aligned}$$

Thus, if (15) is satisfied, LHS of (3) must be strictly positive – a contradiction. Similarly, the supposition $x^F > x^C$ leads a contradiction.

(ii) By (14), we have $P(X^S) = c(x^F)/x^F$ and $P(X^C) = c(x^C)/x^C$. In Lemma 3-(i), we have shown that $x^F = x^C$. Thus, the U-shaped average cost curve implies $P(X^S) = P(X^C)$. $X^S = X^C$ is derived from $p' < 0$.

(iii) Suppose contrary $x^L(m) \leq x^C$. We compare the LHS of (15) with that of (2).

$$\begin{aligned} P'(X^C)x^C + P(X^C) - c'(x^C) &\leq P'(X^S)x^L + P(X^S) - c'(x^C) \quad (\because \text{Lemma 3(ii)}, x^L \leq x^C, P' < 0) \\ &\leq P'(X^S)x^L + P(X^S) - c'(x^L) \quad (\because x^L \leq x^C, c'' > 0) \\ &< (1 + R')P'(X^S)x^L + P(X^S) - c'(x^L) \quad (\because -1 < R' < 0, P' < 0) \end{aligned}$$

Therefore, if (15) is satisfied, LHS of (2) must be positive – a contradiction.

(iv) It is derived from Lemma 3(i)-(iii).

Q.E.D.

Proof of Proposition 4 : Lemma 3(ii) implies that consumer surplus in the Stackelberg-type model ($m > 0$) is equal to that in the Cournot-type model ($m = 0$). Profits of all followers and Cournot firms are zero. Thus, producer surplus is zero in the Cournot-type model. In the Stackelberg-type model, profits of all followers are zero. The leaders can choose the same output as $x^F = x^C$. Thus, by definition, a leader's profit is higher than a Cournot firm. From Lemma 3(iii), producer surplus is strict positive in the Stackelberg-type model.

Q.E.D.

References

- Daughety, A. F. (1990). “Beneficial Concentration,” *American Economic Review*, **80**, 1231–1237.
- Huck, S., Müller, W., and Normann, H. T. (2001). “Stackelberg Beats Cournot: On Collusion and Efficiency in Experimental Markets.” *Economic Journal*, **111**, 749–765.
- Lahiri, S. and Ono, Y. (1988). “Helping Minor Firms Reduces Welfare.” *Economic Journal*, **98**, 1199–1202.
- Lahiri, S. and Ono, Y. (1995). “The Role of Free Entry in an Oligopolistic Heckscher-Ohlin Model.” *International Economic Review* **36**, 609–624.
- Lahiri, S. and Ono, Y. (1998). “Foreign Direct Investment, Local Content Requirement, and Profit Taxation.” *Economic Journal*, **108**, 444–457.
- Mankiw, N. G. and Whinston, M. D. (1986). “Free Entry and Social Inefficiency.” *Rand Journal of Economics*, **17**: 48–58.
- Matsumura, T. (1998). “Partial Privatization in Mixed Duopoly.” *Journal of Public Economics*, **70**, 473–483.
- Matsumura, T. (1999): “Quantity-setting Oligopoly with Endogenous Sequencing,” *International Journal of Industrial Organization*, **17**, 289–296.
- Matsumura, T. (2003) “Consumer-Benefiting Exclusive Territories.” *Canadian Journal of Economics*, **36(4)**, 1007–1025.
- Matsumura, T. and Kanda, O. (2005). “Mixed Oligopoly at Free Entry Markets.” *Journal of Economics*, **84(1)**, 27–48.
- Ono, Y. (1978). “The Equilibrium of Duopoly in a Market of Homogeneous Goods,” *Economica*, **45**, 287–295.
- Ono, Y. (1982). “Price Leadership: A Theoretical Analysis,” *Economica*, **49**, 11–20.
- Suzumura, K. and Kiyono, K. (1987). “Entry Barriers and Economic Welfare.” *Review of Economic Studies*, **54**, 157–167.

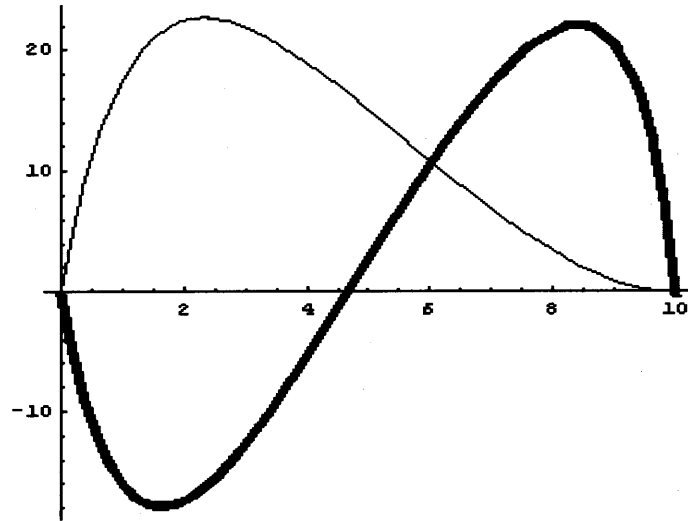


Figure 1: Welfare Comparison when the number of leaders changes for fixed $N = m + n$: The bold curve represents $W^*(m) - W^*(0)$. The other solid curve represents HHI. The horizontal axis represents m and its origin is 0. The origin of vertical axis is 0 when it measures welfare and $1000/N$ when it measures HHI.

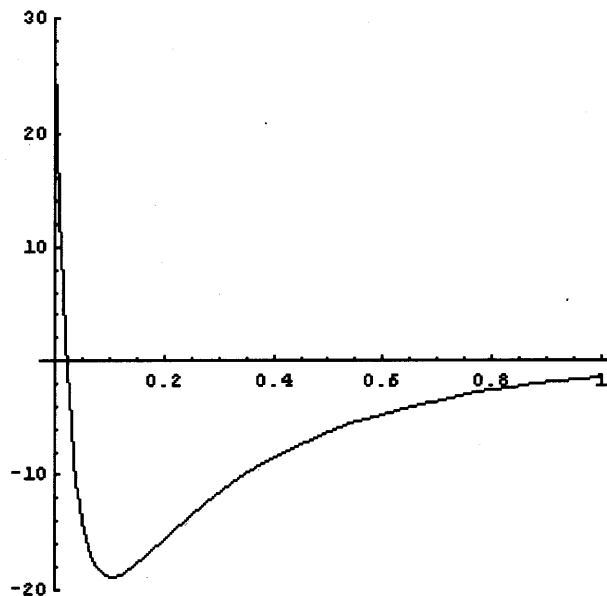


Figure 2: Welfare Comparison when the convexity of cost function changes for fixed m and N : The solid curve represents $W^*(m) - W^*(0)$. The horizontal axis represents k and its origin is 0.

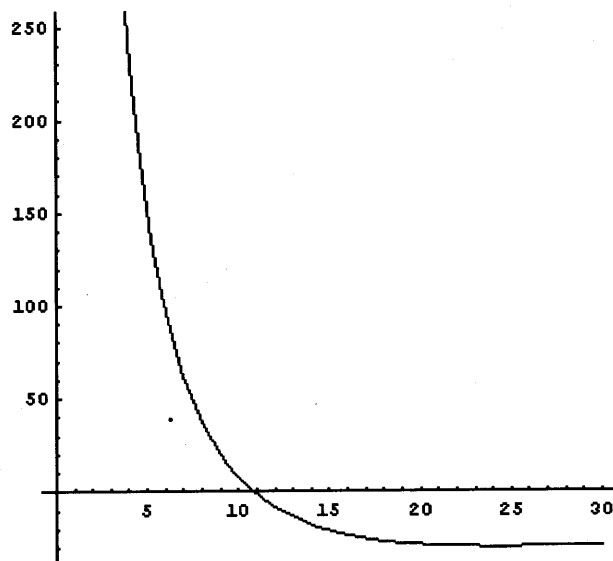


Figure 3: Suppose $a = 100$, $f = 20$, $k = 0.01$: The solid curve represents $\partial W^*(0)/\partial m$. The horizontal axis represents N .