How Many Firms Should Be Leaders? Beneficial Concentration Revisited

Hiroaki Ino* and Toshihiro Matsumura†
Institute of Social Science, the University of Tokyo

January 24, 2008

Abstract

We investigate the relationship between the Herfindahl-Hirschman Index (HHI) and welfare. First, we discuss the model wherein m leaders and $N-m$ followers compete. Daughety (1990) finds that under linear demand and constant marginal costs, the Stackelberg model yields larger welfare and HHI than the Cournot model. Thus, he demonstrates that beneficial concentration occurs. We find that it always occurs when m is sufficiently large under general cost and demand functions but does not always occur when m is small. Next, considering free entries of followers, we find that beneficial concentration always occurs regardless of m.

JEL classification numbers: L13, L40

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

*JSPS Research Fellow. The corresponding author. Address: Institute of Social Science, the University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. E-mail: hiroakiino@04.alumni.u-tokyo.ac.jp.
†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, it 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 asymmetries among firms. The effect of the number of firms is usually straightforward. Typically, an increase in the number of firms improves economic welfare and reduces the HHI in symmetric equilibria.[2] In contrast, the effect of asymmetries among firms is complex. Daughety (1990) investigates this problem by considering asymmetric roles among firms. He formulates a model where m Stackelberg leaders and $N - m$ Stackelberg followers compete in a homogeneous good market, and he discusses the relationship between m and economic welfare. Using a linear demand function and a 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 than the Cournot model ($m = 0$ or $m = N$) does. The heterogeneity of roles among firms increases the HHI, and at the same time it improves welfare (beneficial concentration).

In this paper, we take a close look at this problem. We extend Daughety's model in two directions. First, we consider general demand and cost functions. In particular, generalization of the cost function is important since under the strictly convex costs, asymmetric roles among firms yield the production inefficiency which does not occur under the linear costs and thus it is ambiguous 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 slightly

[1] For 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.

[2] If 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).

changes from N to a smaller number—always improves welfare and increases the HHI, similar to the model of Daughety. This result is robust because it holds true under general cost and demand functions. It 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 slightly changes from 0 to a larger number— also increases the HHI but does *not* always improves welfare. It 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 the HHI (beneficial concentration). However, this latter result is not robust. It does not hold true under increasing marginal costs. 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 free entries of followers.[3] In contrast to the case of an exogenous number of firms, we find that the existence of leaders always improves welfare in the case of an endogenous number of firms. In other words, beneficial concentration caused by heterogeneous roles among firms always occurs in free-entry markets. This result indicates that the HHI might be less appropriate for the measurement of welfare in free-entry markets than at the markets with high entry barriers. It also indicates that leadership is more likely to be beneficial for welfare in the long run.

The paper is organized as follows. Section 2 formulates the model with an exogenous number of firms. Section 3 investigates the equilibrium outcomes and presents results. Section 4 specifies demand and cost functions and provides an insightful example. Section 5 examines the model with free entries. Section 6 concludes the paper. All proofs of Lemmas and Propositions are presented in the Appendix.

[3] If we consider free entries of leaders, no follower enters the market, resulting in the Cournot outcome at the free-entry equilibrium.

2 The Model

We formulate an N-firm oligopoly model ($N > 0$). Firms produce a homogeneous product with the 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 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 $P(X) > 0$ and 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.

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. This model corresponds to the Cournot model when $m = 0$ or $m = N$.[4]

Henceforth, we consider the equilibrium outcome symmetric if all leaders (followers) choose the same output level. The leaders' output is of course different from that of the followers even in the symmetric equilibrium. We make an assumption based 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.

[4] In Lemma 2, we show that 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 $m \to 0$ ($m \to N$).

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

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

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

The solution of the following system of equations, i.e., $(x_L^*(m), x_F^*(m))$, 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.

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

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

where $X^*(m) = mx_L^*(m) + nx_F^*(m)$ and $X_L^*(m) = mx_L^*(m)$. $R(X_L)$ represents a follower's reaction to the leaders' total outputs X_L.[7] $R(X_L)$ is obtained from the follower's first-order condition,

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

Differentiating it yields

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

[6] More 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.

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

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 and that it increases the total output of all firms (including the leaders and the followers).

The equilibrium aggregate output X^* is given by

$$X^*(m) = mx_L^* + nx_F^*. \tag{6}$$

The equilibrium social welfare W^* is given by

$$W^*(m) = \int_0^{X^*} P(q)dq - mC(x_L^*) - nC(x_F^*) - Nf. \tag{7}$$

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 the 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 results regarding 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 $m \in (0,N)$.*

Stackelberg leaders increase the aggregate output as seen in Lemma 1(iii). Thus, total benefit in the industry becomes closer to that in the competitive market, resulting 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,[8] increase when the market structure changes from the Cournot to Stackelberg type. In other words, Stackelberg-type competition results

[8] HHI is defined as the sum of the square of each firm's market share. In our model, it is given by $HHI = m\left(x_L^*(m)/X^*(m)\right)^2 + n\left(x_F^*(m)/X^*(m)\right)^2$.

in market concentration. Despite traditional belief, concentration enhances welfare. Daughety refers to it as "beneficial concentration."

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

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

Proposition 1 and Proposition 2(ii) imply the following result.

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

Proposition 2(ii) indicates that the introduction of a small number of followers into the Cournot model always improves welfare. Since the introduction of 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 (note that HHI is minimized when $m = 0$ and $m = N$). However, the introduction of a small number of leaders into the Cournot model does not always improve welfare. In the next section, we demonstrate this by using the example of quadratic costs.

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

Before explaining the reasoning behind Proposition 2, we present a supplementary result that helps understand the reasoning.

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

(i) $x^*_C = x^*_L(N) = \lim_{m \to N} x^*_L(m) = \lim_{m \to N} x^*_F(m) = x^*_F(N)$,

and (ii) $x^*_L(0) = \lim_{m \to 0} x^*_L(m) > \lim_{m \to 0} x^*_F(m) = x^*_F(0) = x^*_C$.

The outputs of Stackelberg leaders and followers 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 output) when m is close to 0.

When $m = N$, no follower exists. Thus, the leaders do not exhibit strategic behavior against other firms. The introduction of a small number of followers into the Cournot model makes the leaders assume strategic behavior. However, the number of leaders is large, whereas that of followers is small. Thus, the strategic effect is small, and each leader increases output slightly. This is why $x_L^*(m)$ is close to x_C^* when m is close to N. The same principle cannot apply to the case where m is close to 0. The introduction of a small number of leaders into the Cournot model makes the leaders exhibit strategic behavior. Since the number of leaders is small and the number of followers is large, the strategic effect affects significantly each leader's output, resulting in a nonnegligible 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 explain the rationale behind the asymmetry in Proposition 2. Suppose that a small number of leaders are 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 firms produce the same output. Thus, the introduction of a small number of leaders reduces the production efficiency. On the other hand, the introduction of a small number of leaders increases the total output. This competition acceleration effect in turn improves welfare. The former welfare-reducing effect can dominate the latter welfare-improving effect.[9] This is why Proposition 2(i) holds true. Note that the former effect does not exist when the marginal cost is constant.

In contrast, production inefficiency caused by the introduction of a small number of followers is insignificant. As Lemma 2(i) indicates, both the leaders' and the followers' outputs converge to the Cournot output 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 Proposition

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

$x_L^*(m)$	$(a-c)(2k+b)/\beta$
$x_F^*(m)$	$(a-c)(4k^2+2bk(2+n)+b^2)/\alpha\beta$
$X^*(m)$	$(a-c)(4k^2(m+n)+2bk(2+n)(m+n)+b^2(mn+m+n))/\alpha\beta$
$\pi_L^*(m)$	$(a-c)^2(2k+b)^2(2k^2+bk(3+n)+b^2)/\alpha\beta^2-f$
$\pi_F^*(m)$	$(a-c)^2(k+b)(4k^2+2bk(2+n)+b^2)^2/\alpha^2\beta^2-f$

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

1(ii) is robust.

4 Linear Demand and Quadratic Cost Functions

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

Under these specifications, we obtain $R'(X_L)=-b/(2k+nb+b)$, where $n\equiv N-m$. Therefore, the first-order conditions (2) and (3) are

$$(a-c)-(2k+mb+b)x_L^*-nbx_F^*+\frac{nb^2x_L^*}{2k+nb+b}=0, \tag{8}$$

$$(a-c)-(2k+nb+b)x_F^*-mbx_L^*=0. \tag{9}$$

Solving these equations yields the results that have been summarized in Table 1. $\pi_L^*(m)$ ($\pi_F^*(m)$) represents the equilibrium profit of a leader (follower). We can compute equilibrium social welfare in this specific case by substituting the results in Table 1 into

$$W^*(m)=\frac{b}{2}\left(X^*(m)\right)^2+m\pi_L^*(m)+n\pi_F^*(m). \tag{10}$$

We demonstrate that leadership can reduce welfare if and only if the cost function is strictly convex.

Proposition 3 *If $k>0$, there exists N' (>0) such that $\frac{\partial W^*}{\partial m}\big|_{m=0}\lesseqgtr 0$ if and only if $N\gtreqless N'$.*

Propositions 1 and 3 imply the following result.

Corollary 2 *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 the HHI. This corollary implies that the introduction of leaders into the Cournot model can reduce welfare in accord with the traditional belief on market concentration. We present a numerical example for such a situation.

Example

Suppose that $a = 110$, $b = 1$, $c = 10$, $k = 0.2$, $f = 20$, and $N = 10$. Using (10), we can simulate how the ratio of leaders in the industry affects social welfare. Figure 1 depicts 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 decreases when the ratio of leaders is small, while as predicted in Proposition 2(ii), the welfare increases when the ratio of leaders is large. Figure 1 depicts how the HHI changes in the same situation.

From Figure 1, we find that the HHI provides a good insight into the change in welfare in this example, except the range around $m = N$. A decrease (an increase) in the HHI increases (decreases) welfare. In the case of constant marginal costs that Daughety (1990) investigates, the range in which HHI increases but welfare also increases is substantially wider than that in this example. 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 in small. Thus, it is reasonable that antimonopoly departments focus on markets with a small number of leaders.

5 Free Entries of Followers

In this section, we consider free entries of followers.[10] We provide long-run cases in which 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 exogenously.[11]

We first explain the time structure of model 1. In the first stage, potential entrants (followers) decide whether to enter the market. Then, all firms observe m and n. In the second stage m leaders choose their output. In the third stage n followers (n is determined endogenously in the first stage) choose their output after observing the leaders' output.

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

We provide some cases rationalizing each model.[12]

Consider model 1. Suppose that m firms have a special ability to produce faster than the other firms. 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 story, the leaders' production follows the followers' entry. The

[10] For the Stackelberg model with endogenous entry of firms, Etro (2006) also considers the model with free entries of followers. His model has a single leader and the time structure is model 2 explained below.

[11] If the number of leaders is endogenous and is determined by the zero-profit condition, the followers obtain negative profits (thus, no follower enters the market). In other words, the equilibrium outcome becomes Cournot type.

[12] Other 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.

number of skilled firms m is given exogenously[13] and the number of unskilled firms is determined by the free entries.

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 story, the new entrants enter the market after the leaders have satisfied their production commitment. As a result, the new entrants become followers.

Henceforth, we include an assumption that makes our long-run analysis meaningful.

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

Note that this assumption implies $f > 0$. Thus, the average cost $(C(x) + f)/x$ is U shaped if the output minimizing average cost exists, while it monotonically decreases if the 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.

5.1 Model 1: Entries before Stackelberg competition

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

$$P(X^{**}(m))x_F^{**}(m) - C(x_F^{**}(m)) - f = 0, \tag{11}$$

[13] Some 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.

[14] If the average cost is decreasing, Assumption 5 is not satisfied in model 2. This is because the aggregate output of the leaders undercuts the Cournot price, and no followers can enter in the equilibrium. However, even in this case, if the given number of leaders is sufficiently small, beneficial concentration occurs since the optimal case with the decreasing average cost is monopoly.

where $X^{**}(m) = mx_L^{**}(m)+n^{**}(m)x_F^{**}(m)$ and $x_L^{**}(m)$ ($x_F^{**}(m)$) represent the long-run equilibrium output of a leader (follower). $N^{**}(m)$ represents the total number of firms and is given by $N^{**}(m) = m+n^{**}(m)$. Since the subgames following followers' entry decisions is same as the short-run model analyzed in Section 3, these equilibrium outcomes are obtained by the conditions (2), (3) and (11). Using the expressions given in the 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 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 of the equilibrium outcomes between the Cournot and Stackelberg models.

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

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 depends 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.

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 (e.g., the case $m > 0$) to the Cournot model ($m = 0$) reduces the leaders' output given n. This leads to an upward shift of each follower's residual demand curve, however, in the long-run, it induces new entries, leading to a downward shift of each follower's residual demand curve. Eventually, the upward shift of the residual demand curve is canceled out by 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).

Let $W^{**}(m)$ denote the long-run equilibrium welfare given m. Note that $W_C^{**} = W^{**}(0)$ 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* $W^{**}(m) > W_C^{**}$ $\forall m > 0$.

The introduction of a leader yields asymmetry among firms and increases the HHI. In addition, it reduces the total number of firms (Lemma 3(iv)), and again leading to an increase in HHI. However, it improves welfare (Proposition 4). In other words, beneficial concentration always occurs in this context.[15]

We explain the reasoning behind this result. Consider that one firm (firm 1) in the Cournot model becomes the leader. Then, in both short-run and long-run cases, firm 1's output increases, while the other firms' outputs decrease. In other words, production substitution from the followers to the leader occurs.

In the short-run case, each of firm 2, firm 3,..., and firm N reduces its output. Production substitution approximately increases the leader's cost $C'(x_1)|\Delta x_1|$ and reduces the followers' costs $(N-1)C'(x_2)|\Delta x_2|$, where Δx_i is the difference in firm i's output between the Stackelberg and Cournot models. When marginal cost is increasing and the leader's output level is higher than the follower's, production substitution impairs production efficiency. This is why the introduction of a leader can reduce welfare in the short-run case discussed in the previous section.

In the long-run case, production substitution from the followers to the leader occurs but in a completely different manner. No follower entering the market changes its output. Instead, the number of followers entering is reduced. Production substitution approximately increases the leader's cost $C'(x_1)|\Delta x_1|$ and reduces the follower's cost $|\Delta n(C(x_2) + f)|$, where Δn is the difference in the number of followers entering between the Stackelberg and Cournot models. In the short-run case, per

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

output reduction of costs is $C'(x_2)$ (follower's marginal cost). In the long run, it is $(C(x_2)+f)/x_F$ (follower's average cost). Since firm i's average cost is equal to the price and the leader's marginal cost is lower than the price, production substitution in the long run improves production efficiency. This is why the Stackelberg model always yields larger welfare than the Cournot model.

5.2 Model 2: Leadership by incumbents before the entry of followers

In this subsection, we formulate a model with an alternative time structure. After observing the leaders' outputs, new entrants decide whether to enter the market. Let superscript "***" denote the equilibrium outcomes in this model. Note the case where $m = 0$ corresponds to the long-run Cournot model in this model as well. Therefore, we have the relations $x_C^{***} = x_C^{**}$, $X_C^{***} = X_C^{**}$, and $N_C^{***} = N_C^{**}$ (subscript C denotes the outcomes in the Cournot model).

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

Lemma 4 states that Lemma 3, which is derived in model 1, still holds true in model 2. Lemmas 4 and 3 imply that models 1 and 2 yield the same output level for each follower and the total output level, although the two models yield different output levels for each leader and the number of followers.

The only aspect in which model 2 differs from model 1 is in terms of the followers' reactions to the leaders' actions. In model 2, when the followers react to the leaders' output levels, the former can change not only their output levels but also their decisions of entry and exit. Therefore, at any given level of the leaders' output, the follower's output and thus the total output always equals that in the Cournot model due to the same reason that mentioned 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,* $W^{***}(m) > W_C^{***}$ $\forall m > 0$.

Proposition 5 states that beneficial concentration always occurs in model 2 as well as in model 1. The proofs of Lemma 4 and Proposition 5 are similar to those of Lemma 3 and Proposition 4, respectively, and therefore, they can be omitted.

6 Concluding Remarks

In this paper, we analyze welfare in a model with multiple Stackelberg leaders and reexamine the relationship between the HHI and welfare. We find that beneficial concentration always occurs when a small number of followers are introduced. In contrast, the introduction of a small number of leaders can reduce welfare. It always increases the HHI, although welfare can either increase or decrease. We also find that the latter result does not hold if we consider free entries of followers and that beneficial concentration always occurs when we introduce a leader into the Cournot model. In particular, beneficial concentration occurs even if the leaders do not take entry of followers into account (model 1). This result indicates that the HHI might be less appropriate for the measurement of welfare in free-entry markets than in markets with high entry barriers.

Note that for the robustness of beneficial concentration in long-run models, the exogeneity of m is the key setting. This setting implies that a limited number of firms have the ability to become leaders. In reality, we can find a wide variety of situations suitable for this limited leadership, for example, incumbents with bottleneck facilities as in the electricity industry, long-established firms in a production alliance as in an airline industry, and pioneering firms successfully enclosing consumers like as in the mobile phone market. Sometimes, particularly in deregulation policies, this limited leadership is considered as a problem that results in entry barriers. However, our result indicates the possibility of preventing inefficient entries if limited 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 firms yields asymmetry among them. However, in our long-run models in particular, the reasoning behind Lemma 3(i) is applicable even if we introduce leaders that have a different cost from that of entrants into 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 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 (deteriorate) welfare. However, it is difficult to find a more meaningful 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.[16]

[16] For 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).

APPENDIX

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 $x_L^*(m) \le x_C^*$ and derive a contradiction. Note that $x_L^*(m) \le x_C^*$ implies $X^*(m) \le X_C^*$.

Substituting $m = 0$ into (3) yields

$$P'(X_C^*)x_C^* + P(X_C^*) - C'(x_C^*) = 0, \tag{12}$$

where $X_C^* = X^*(0) = Nx_C^* = Nx_F^*(0)$. We compare the LHS in (12) with that in (2). Assumption 3 ensures that $P'(X)x+P(X)$ is decreasing in X. Since $X^*(m) \le X_C^*$, we have $P'(X_C^*)x_C^* + P(X_C^*) - C'(x_C^*) \le P'(X^*(m))x_C^* + P(X^*(m)) - C'(x_C^*)$. Since $P' < 0$, $C'' \ge 0$, and $x_L^*(m) \le x_C^*$, we have $P'(X^*(m))x_C^* + P(X^*(m)) - C'(x_C^*) \le P'(X^*(m))x_L^*(m) + P(X^*(m)) - C'(x_L^*(m))$. 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 (12) 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 $N < N'$ and $\frac{\partial W^*}{\partial m}\big|_{m=0} < 0$ when $N > N'$.

(ii) We show that for any situation satisfying Assumptions 1–4, $\partial W^*(N)/\partial m < 0$. Differentiating (7) and evaluating at $m = N$, we obtain

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

Since $x_L^*(N) = x_F^*(N) = x_C^*$ (Lemma 2(i)), the first and the second terms on the RHS are canceled out. Thus, if the last term on the RHS is negative, the LHS is also negative. Comparative statics analysis using (2) and (3) yields

$$\frac{\partial x_L^*}{\partial m}\Big|_{m=N} = \frac{P'x_C^*(P' - C''(x_C^*))R'|_{m=N}}{N(P' + P''x_C^*)(P' - C''(x_C^*)) + (P' - C''(x_C^*))^2} < 0 \tag{13}$$

where $R'|_{m=N} = -(P'(X_C^*) + P''(X_C^*)x_C^*)/(P'(X_C^*) - C''(x_C^*)) < 0$. We use $x_L^*(N) = x_F^*(N) = x_C^*$ to induce (13), 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

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

On the other hand, since $F(m, x_L^*(m), x_F^*(m)) = 0\ \forall m \in (0, N)$, $\lim_{m\to 0} F(m, x_L^*(m), x_F^*(m)) = 0$. Hence, $F(0, \lim_{m\to 0}(x_L^*(m), x_F^*(m))) = 0$. Similarly, $G(0, \lim_{m\to 0}(x_L^*(m), x_F^*(m))) = 0$. Therefore, $\lim_{m\to 0}(x_L^*(m), x_F^*(m))$ 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 $\lim_{m\to 0} x_F^*(m) = x_F^*(0) = x_C^*$ and $\lim_{m\to 0} x_L^*(m) = x_L^*(0)$. Similarly, taking $m \to N$, we obtain $\lim_{m\to N} x_L^*(m) = x_L^*(N) = x_C^*$ and $\lim_{m\to N} x_F^*(m) = x_F^*(N)$.

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 (12). Then, given $x_L^*(N) = x_C^*$, $x_F^*(N)$ is determined according to (3), i.e., $P'(Nx_C^*)x_F^* + P(Nx_C^*) = C'(x_F^*)$. $x_F^*(N)$ satisfies this equation if $x_F^*(N) = x_C^*$ by (12). Then, uniqueness of $x_F^*(N)$ implies that $x_F^*(N) = x_C^* = x_L^*(N)$. **Q.E.D.**

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

$$\frac{\partial W^*}{\partial m}\Big|_{m=0} = [Px_L^*(0) - C(x_L^*(0))] - [Px_F^*(0) - C(x_F^*(0))] + N[P(X^*(0)) - C'(x_F^*(0))]\frac{\partial x_F^*}{\partial m}\Big|_{m=0}. \tag{14}$$

From the results in table 1, we have

$$P(X^*(0))x_L^*(0) - C(x_L^*(0)) = \frac{(a-c)^2(2k+b)^2(2k^2+bk(3+N)+b^2)}{(2k+Nb+b)(4k^2+2bk(2+N)+b^2)^2}, \tag{15}$$

$$P(X^*(0))x_F^*(0) - C(x_F^*(0)) = \frac{(a-c)^2(k+b)}{(2k+Nb+b)^2}. \tag{16}$$

From the first-order condition (3),

$$P(X^*(0)) - C'(x_F^*(0)) = -P'(X^*(0))x_F^*(0) = bx_F^*(0) = \frac{b(a-c)}{2k+Nb+b}. \tag{17}$$

Differentiating x_F^* yields

$$\frac{\partial x_F^*}{\partial m}\Big|_{m=0} = \frac{-N(a-c)b^3}{(2k+Nb+b)^2(4k^2+2bk(2+N)+b^2)}. \tag{18}$$

Substituting (15)–(18) into Equation (14), we obtain

$$\frac{\partial W^*}{\partial m}\Big|_{m=0} = \frac{N(a-c)^2b^3A}{(2k+Nb+b)^3(4k^2+2bk(2+N)+b^2)^2}, \tag{19}$$

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 (19) 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 varsa.* **Q.E.D.**

Proof of Lemma 3: (i) Let $\hat{x}$ denote the output minimizing average cost of the firm $(C(x)+f)/x$.

If the average cost is monotonically decreasing, suppose $\hat{x} = \infty$. First, we show that $x_F^{**}(m) \leq \hat{x}$. Obviously, this inequality is satisfied if $\hat{x} = \infty$. Suppose the contrary that $x_F^{**}(m) > \hat{x}$ holds true when $\hat{x} < \infty$. 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 output before observing the followers' outputs). Firm i's profit is zero before the deviation (see (11), zero profit condition). The deviation increases the price and reduces firm i's average cost, and therefore firm i obtains positive profits after the deviation, which is a contradiction. The same principle can apply to the Cournot model, and we have $x_C^{**} \leq \hat{x}$.

Then, we prove $x_F^{**}(m) = x_C^{**}$ by contradiction. Suppose that $x_F^{**}(m) < x_C^{**}$. From (11) (zero profit condition), we have $P(X^{**}(m)) = (C(x_F^{**}) + f)/x_F^{**}$ and $P(X_C^{**}) = (C(x_C^{**}) + f)/x_C^{**}$. These equations imply that $P(X^{**}(m)) > P(X_C^{**})$ since the average cost $(C(x) + f)/x$ is decreasing when $x < \hat{x}$ and $x_F^{**}(m) < x_C^{**} \leq \hat{x}$ by the supposition. Since $P' < 0$, $X^{**}(m) < X_C^{**}$ must hold. We compare the LHS in (12) with that in (3). Assumption 3 ensures that under the condition $X^{**}(m) < X_C^{**}$, the inequality $P'(X_C^{**})x_C^{**} + P(X_C^{**}) - C'(x_C^{**}) < P'(X^{**}(m))x_C^{**} + P(X^{**}(m)) - C'(x_C^{**})$ holds. Since $P' < 0$ and $C'' \geq 0$, under the condition $x_F^{**}(m) < x_C^{**}$, we have $P'(X^{**}(m))x_C^{**} + P(X^{**}(m)) - C'(x_C^{**}) < P'(X^{**}(m))x_F^{**}(m) + P(X^{**}(m)) - C'(x_F^{**}(m))$. Thus, if (12) is satisfied, the LHS in (3) must be positive, which is a contradiction. Similarly, $x_F^{**}(m) > x_C^{**}$ leads to a contradiction.

(ii) From (11), we have $P(X^{**}(m)) = (C(x_F^{**}) + f)/x_F^{**}$ and $P(X_C^{**}) = (C(x_C^{**}) + f)/x_C^{**}$. In Lemma 3(i), we have shown that $x_F^{**}(m) = x_C^{**}$. Thus, $P(X^{**}(m)) = P(X_C^{**})$. $X^{**}(m) = X_C^{**}$ is derived from $P' < 0$.

(iii)Suppose the contrary, i.e., $x_L^{**}(m) \leq x_C^{**}$. We compare the LHS in (12) with that in (2). Since $P' < 0$ and $C'' \geq 0$, under the condition $x_L^{**}(m) \leq x_C^{**}$, we have $P'(X_C^{**})x_C^{**} + P(X_C^{**}) - C'(x_C^{**}) \leq P'(X^{**}(m))x_L^{**}(m) + P(X^{**}(m)) - C'(x_L^{**}(m))$, where we use Lemma 3(ii). Since $nR'(X_L) \in (-1, 0)$ (see the discussion immediately after equation (5)), $P'(X^{**}(m))x_L^{**}(m) + P(X^{**}(m)) - C'(x_L^{**}(m)) <$

$(1+n^{**}(m)R')P'(X^{**}(m))x_L^{**}(m)+P(X^{**}(m))-C'(x_L^{**}(m))$. Therefore, if (12) is satisfied, the LHS in (2) must be positive, which is a contradiction.

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

Proof of Proposition 4: Corollary 3 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 (11). In the Stackelberg model, the profits of all followers are also zero. Since x_L^{**} is larger than x_C^{**} by Lemma 3(iii) and is less than the output level that equals the marginal cost and price by (2), the average cost of each leader $(C(x_L^{**})+f)/x_L^{**}$ is less than $P(X_C^{**}) = (C(x_C^{**}) + f)/x_C^{**}$, 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.**

References

Daughety, A. F. (1990). "Beneficial Concentration." *American Economic Review*, **80**, 1231–1237.

Etro, F. (2006). "Aggressive Leaders." *Rand Journal of Economics*, **37**, 146–54.

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.

Levin, D. (1988). "Stackelberg, Cournot and Collusive Monopoly: Performance and Welfare Comparisons." *Economic Inquiry*, **26(2)**, 317–330.

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.

Sherali, H. (1984). "A Multiple Leader Stackelberg Model and Analysis" *Operations Research*, **32**, 390–404.

Suzumura, K. and Kiyono, K. (1987). "Entry Barriers and Economic Welfare." *Review of Economic Studies*, **54**, 157–167.

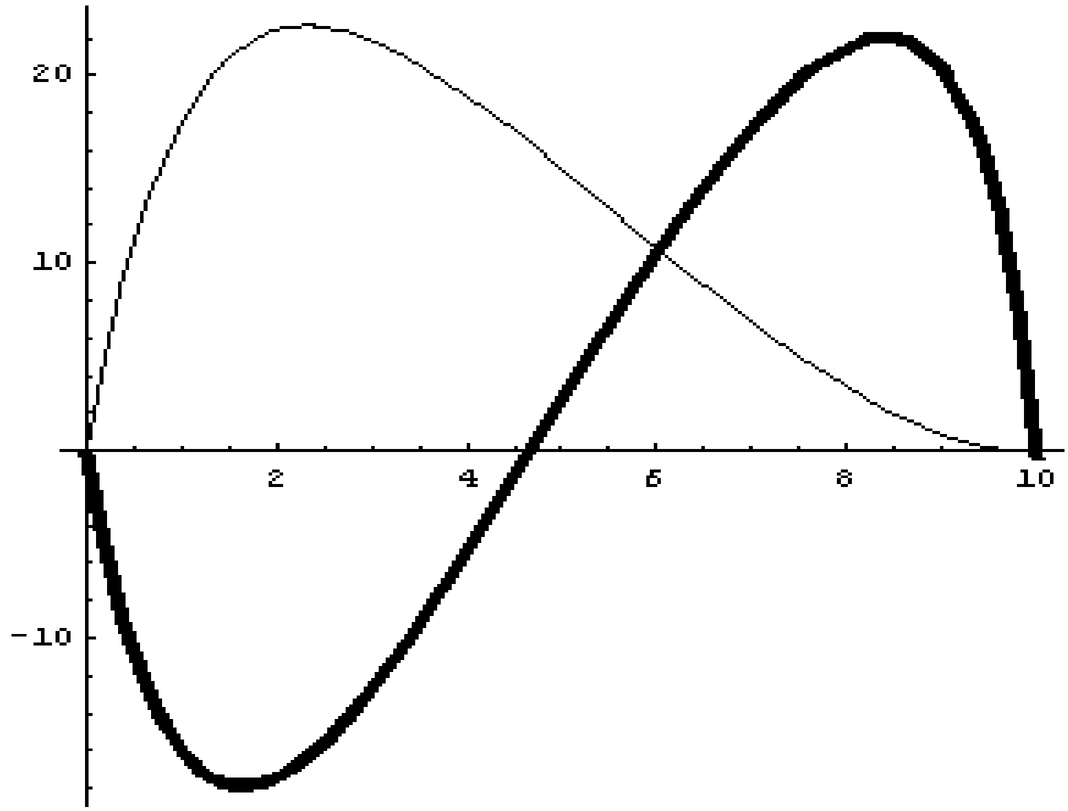
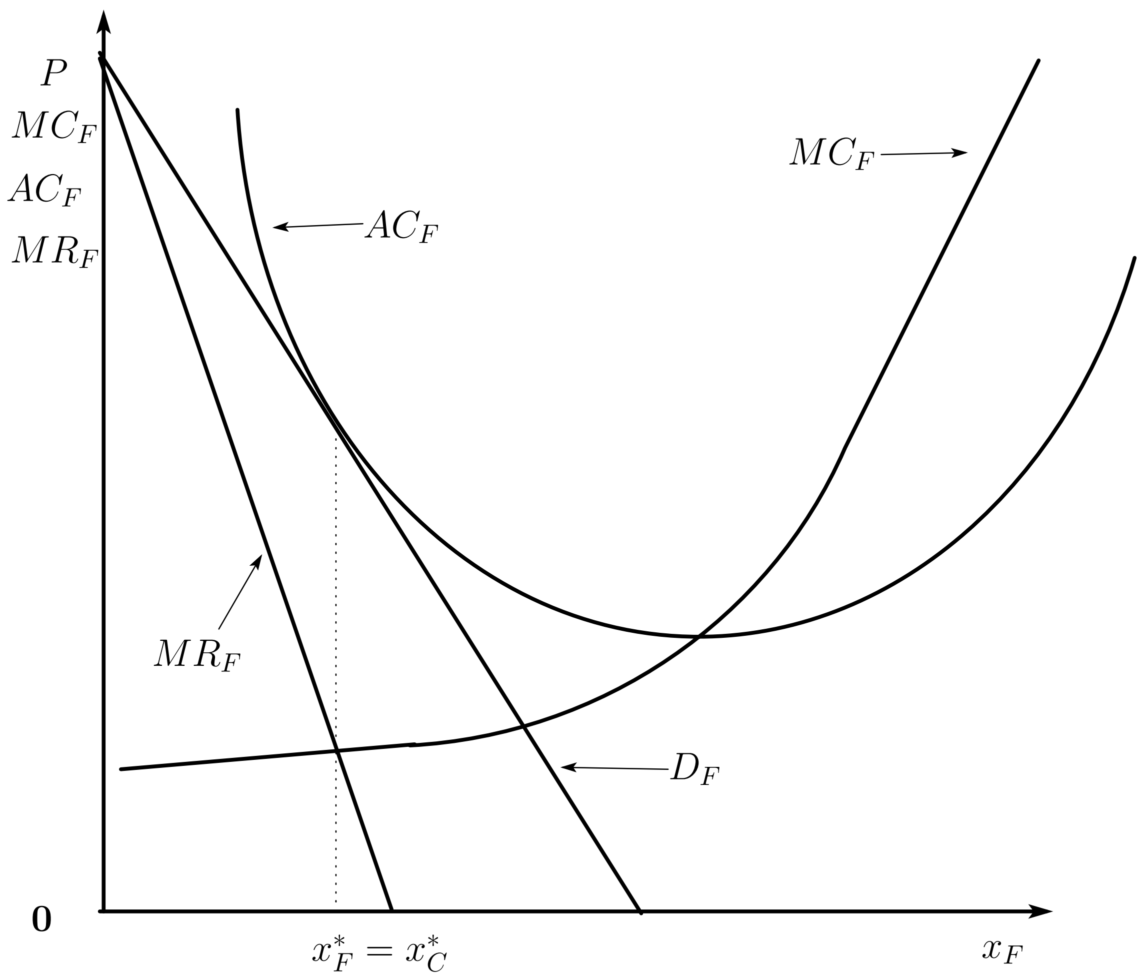


Figure 1: Welfare Comparison when the number of leaders changes for fixed $N = m + n$: The thick curve represents $W^*(m) - W^*(0)$. The thin curve represents the HHI (scaled to adjust the figure). The horizontal axis represents m and its origin is 0. The origin of the vertical axis is 0 when it measures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



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 2