ANNALS OF ECONOMICS AND FINANCE **2**, 353–377 (2001)

The Dynamics of Firms in the Presence of Adjustment Costs

Ruqu Wang*

Economics Department, Queen's University
Kingston, Ontario, Canada, K7L 3N6
E-mail: wangr@qed.econ.queensu.ca

In this paper we investigate how capacity adjustment costs affect a firm's response to demand uncertainty. We first characterize the pattern of optimal capacity adjustment for a monopolistic firm and find that capacity behaves as a stabilizer for the firm's output. For duopolistic firms the pattern is similar. However, a firm may deviate depending on the demand and capacity circumstance. We find that when there is only a small cost of adjustment, a firm has more incentive to deviate at a larger capacity. We also derive conditions under which deviation in the high-demand state (regardless of present capacity) is more profitable. The case of zero adjustment costs is also discussed.

Key Words: Capacity; Adjustment costs; Monopoly; Duopoly; Collusion.
JEL Classification Numbers: L000, E200.

1. INTRODUCTION

The optimal production plan of a firm often relates to the demand situation of the industry. When the demand is high, it tends to produce more; when the demand is low, it tends to produce less. A firm adjusts output levels by adjusting inputs. The levels of some inputs can be altered without additional costs. Raw materials, for example, can be purchased at a constant price according to a firm's need. The adjustment of the levels of other inputs, however, can be costly. For example, capital inputs, such as machines, usually require some cost of adjustment. The purchase of new machines, even at a constant price, requires additional space in the factory which costs money. Selling surplus machines, on the other hand, is energy

* I would like to thank Marilyn Banting, Alan Gunderson, Bart Lipman, Bob Rosenthal, Doug Tattrie, and Jean-Francois Wen for their very helpful comments. This research is supported by the Advisory Research Committee at Queen's University and the Social Sciences and Humanities Research Council of Canada.

1529-7373/2001

consuming, and the selling price is usually below market value if one wants to sell them quickly. A firm's production depends crucially on these inputs, which are costly to adjust. These inputs are called the *capacity* of a firm (cf. footnote 1) in this paper.

It is an indisputable fact that market demands are unpredictable. The industry's demand goes up and down, and consumers' tastes are constantly changing. How does a firm react to demand uncertainty, especially when the alternations of the levels of some inputs are costly? Obviously, a firm needs a larger capacity to produce larger outputs more efficiently, and vice versa for small outputs. But constant changes in capacity are very costly. This paper intends to provide a characterization for the firm's optimal behavior in this kind of situation.

There is a large literature on a firm's reaction in the presence of demand uncertainty, though none addresses the issue of adjustment costs. Green and Porter (1984), and Rotemberg and Saloner (1986), for example, assume that a firm can produce as much or as little as it desires at constant unit costs. The capacity issue is added in Staiger and Wolak's (1992) model, but the adjustment of capacity is costless. Similar assumptions are made in Brock and Scheinkman (1985).

It is reasonable to believe that firms behave differently when there are adjustment costs. A firm will not adjust its capacity or output as drastically as in the case of no adjustment costs. Using a model similar to Rotemberg and Saloner's but adding some capacity adjustment costs, we find that a firm's capacity acts like a shock absorber. The capacity is adjusted upward when the demand is high, but not as much as in the case of no adjustment costs, and vice versa when the demand is low. In this way, the output of a firm does not fluctuate as severely as the demand. Given this property of the adjustment process, a firm is most productive (efficient) when the same demand condition occurs in consecutive periods.

The above characteristics prevail along both the optimal production path of a monopoly and the optimal collusive production path of an oligopoly. Regarding the optimal collusive path, it is important to identify whether the path can be sustained in an equilibrium and when a firm has the most incentive to deviate from the path. Green and Porter (1984) find that firms are more likely to deviate in times of recession when the demand is low. Rotemberg and Saloner (1986), on the other hand, come to the opposite conclusion: firms are more likely to deviate during economic booms when the demand is high. All these conclusions are reached assuming a firm can increase or decrease its output without any adjustment costs. This paper re-examines the same issue but under more reasonable assumptions; that is, a firm must adjust its capacity in order to produce more efficiently under demand uncertainty and this adjustment is costly. In the analysis, we are able to derive conditions under which a larger capacity gives a firm more incentive to deviate and conditions under which a firm is most likely to deviate when the demand is high.

The main feature of our model is that we incorporate capacity adjustment costs explicitly. Capacity here refers to the facility of a firm. Adjusting it is obviously costly. In general, a large capacity is more appropriate for producing a large output, since a large capacity is also more expensive to maintain.

This paper relates a firm's capacity adjustment process to its output when market demand is stochastic. It studies a firm's long-run production strategies and obtains many conclusions, some are very different than models without capacity adjustment costs. It also provides conditions under which the results obtained by previous researchers are valid. Therefore, it can be viewed as a step closer towards the full understanding of the relationships between a firm's capacity, productivity, and market demand conditions.

The rest of the paper is organized as follows. In Section 1, we characterize the optimal production plan of a monopoly firm. The results are then used in Section 2 where the optimal collusive path of production is derived. In Section 2, we also analyze the incentives for a firm to deviate from the optimal path. A model in which the production function does not depend on capacity is considered as a special case. Section 3 contains some further remarks. Most proofs are relegated to an appendix.

2. THE CASE OF A MONOPOLY

Suppose that a firm is a monopoly in an industry. There are infinitely many periods starting with period 0. In each period, the demand for the firm's product is high with probability θ and low with probability $1-\theta$. We assume that these probabilities are independent across time. Let $S \in \{H, L\}$ denote the state of demand in a period, where H indicates the high-demand state and L the low-demand state. The inverse demand function in state S is given by $P = P^S(Q)$. That is, in order to sell Q units of output, the monopoly has to set a price equal to $P^S(Q)$. Of course, this $P = P^S(Q)$ function is decreasing in Q; the more one wants to sell, the lower the price he can charge.

Let $R^S(Q) = P^S(Q)Q$ denote the revenue for the firm in one period when the state of demand is S. Then the corresponding marginal revenue (measured as the extra revenue of selling one extra unit of product) is given by $R^S_Q(Q)$. Assume that $R^S_{QQ}(Q) < 0$; that is, the marginal revenue curve is downward sloping. This assumption is satisfied by many traditional demand functions, such as $P = a - bQ$. Its effect is to guarantee the uniqueness of an optimal solution to our problem.

The relative magnitude of the above mentioned demand functions in the *High* and *Low* demand states is captured by the following relationships:

$$P^H(Q) > P^L(Q), \quad P_Q^H(Q) \geq P_Q^L(Q), \quad \forall Q \geq 0.$$

Therefore, the price is not only higher, but also declines more slowly in the *High* demand state (compared at the same level of output). This property of the demand function is directly related to the elasticities of demand in each state: it implies that the elasticities of demand are higher in the high-demand state. As is argued in Weitzman (1982), as the number of products tends to increase in good times (the high-demand state), all products become closer substitutes and thus the elasticities of demand are also higher. In our model, these relationships guarantee that the optimal output in the high-demand state is always higher than the low-demand state.

The cost of production of the firm is a function of its output and its capacity in each period. The cost function is not linear in the firm's output. Usually, the cost of production can be divided into two categories. One is the cost of capital inputs, such as machines and tools. We call the stock of these inputs capacity.[1] The other is the cost of raw materials and labor inputs. The difference between the two is that, the adjustment of the former is costly (in addition to the cost of purchasing), while the latter is not. Imagine the addition of one more big machine. The firm has to make room for it, probably by constructing a new building or squeezing the existing workplace. This imposes explicit or implicit costs to the firm, accounting for the cost of new buildings or the loss in productivity of the existing workforce.

Let φ_t be the firm's capacity and Q_t be its output in period t. The cost of production is given by $C(Q_t, \varphi_t)$. Assume that $C_Q > 0$, $C_{Q\varphi} < 0$, $C_\varphi(0, \varphi) > 0$, $C_{QQ} > 0$, $C_{\varphi\varphi} > 0$, and $C_{QQ}C_{\varphi\varphi} - C_{Q\varphi}^2 \geq 0$; i.e., $C(Q, \varphi)$ is increasing in Q and (weakly) convex in (Q, φ). The assumption $C_{Q\varphi} < 0$ implies that a larger capacity is more suitable for producing a larger output, since it reduces the marginal cost of production (C_Q). This is one of the major assumptions we make in this paper. It characterizes the common property shared by most production (if not all): more machines and tools make producing extra units output easier, faster, and cheaper. This can be illustrated in the following two examples.

[1]Here, the use of the term 'capacity' is not extremely precise. As we shall see later, it does not measure the maximal output the firm can produce. Rather, it measures the amount of equipment or facility in the firm. But we have not been able to find a better term for it.

Example 1 Consider the Cobb-Douglas production function $Q = L^{\alpha}\varphi^{\beta}$, where L is the labor input, φ is the capital input (i.e., capacity, such as machines), $\alpha > 0$, $\beta > 0$, and $\alpha + \beta < 1$. Let w be the wage rate for labor and r be the interest rate for capital. Then the total cost of producing Q units of output is given by $C(Q, \varphi) = r\varphi + wQ^{\frac{1}{\alpha}}\varphi^{-\frac{\beta}{\alpha}}$. We can easily verify that all the above assumed properties on $C(Q, \varphi)$ are satisfied. In this example, an increase in capacity increases the marginal productivity of labor; that is, $C_{Q\varphi}(Q, \varphi) > 0$. Given the unit cost of capacity, it is more cost effective to have a large capacity when a higher level of output is being produced.

Example 2 Let φ be the number of plants a firm owns, and $G(Q)$ be the (identical) cost function for each plant, with $G(0) > 0$, $G'(Q) > 0$, $G''(Q) > 0$. The optimal way of producing a total of Q units of output is to split the production equally among all the plants, since the production has decreasing returns to scale ($G''(Q) > 0$). Thus, the total cost of production is given by

$$C(Q, \varphi) = \varphi G\left(\frac{Q}{\varphi}\right) + c\varphi,$$

where c is the maintenance cost per plant. Generalize the meaning of φ such that it can be any positive number. Simple calculations show that

$$C_Q = G'(\frac{Q}{\varphi}) > 0, \quad C_{Q\varphi} = -\frac{Q}{\varphi^2}G''(\frac{Q}{\varphi}) < 0, \quad C_{\varphi}(0, \varphi) = G(0) + c > 0,$$

$$C_{QQ} = \frac{1}{\varphi}G''(\frac{Q}{\varphi}) > 0, \quad C_{\varphi\varphi} = \frac{Q^2}{\varphi^3}G''(\frac{Q}{\varphi}) > 0, \quad \text{and} \quad C_{QQ}C_{\varphi\varphi} - C_{Q\varphi}^2 = 0.$$

In this example, more plants are suitable for producing a higher level of output, as it reduces the unit cost of production. Of course, because of the cost of plant maintenance, setting up too many plants would not be optimal.

From time to time, the firm may want to adjust the level of capacity. The cost of adjusting from φ_{t-1} to φ_t is denoted by $H(\varphi_t - \varphi_{t-1})$. (The cost of purchasing extra capacity or selling excessive capacity is already in the cost function $C(Q, \varphi)$. For capacity φ_t in period t, a firm has to pay an amount equivalent to the interest of borrowing φ_t in that period.) The adjustment cost is assumed to be independent of the level of output, mainly because it represents the cost of physically adding or deleting the capacity in question. This cost is incurred regardless of the production plan.

Assume that $H(\cdot)$ is twice continuously differentiable, $H(0) = 0$, $H'(0) = 0$, $H''(\varphi) > 0$, $\forall\varphi$. Here, $H(\varphi)$ is assumed to be convex to ensure that the firm's optimization problem has a unique interior solution. There are several reasons this assumption can be justified. Eisner and Strotz (1963) give two. The first is that, as the firm increases its demand for investment goods in a single period, pressure will be put on the supply of investment goods. The second argument is that there are increasing costs associated with integrating new equipment into a going concern: reorganizing production lines, training workers, etc. Even though this assumption is sometimes criticized,[2] it is widely used in modern dynamic economic analysis. To simplify the analysis, we further assume that $H(-\varphi) = H(\varphi)$. (Selling extra capacity means paying rent on empty buildings and/or receiving unfavorable prices.) As an example, $H(\varphi) = \varphi^2$ satisfies the above requirements.

The timing of the monopolist's dynamic maximization problem is as follows. At the beginning of each period, the state of that period's demand is revealed. The firm then simultaneously chooses the capacity φ, the output Q, and the price P for that period.

Let δ be the discount factor for the firm. In period t, let $V(\varphi_{t-1})$ be its maximized discounted expected profit given that the previous capacity is φ_{t-1} (before the state of demand is revealed in period t); let $V^S(\varphi_{t-1})$ be the corresponding profit when that state of demand is revealed and equal to S. Note that $V(\cdot)$ and $V^S(\cdot)$ are not indexed by t since they are time-independent; in our infinite horizon setting they depend on the capacity of the last period only. Furthermore, in what follows, we suppress the subscript t whenever there is no ambiguity. By definition,

$$V(\varphi_{t-1}) = \theta V^H(\varphi_{t-1}) + (1-\theta)V^L(\varphi_{t-1}).$$

The Bellman equation for the firm's profit maximization can be expressed as follows (rf. Ross (1983, p.74), Theorem 1.1, the optimality equation):

$$V^S(\varphi_{t-1}) = \max_{Q_t,\varphi_t} \left\{ R^S(Q_t) - C(Q_t,\varphi_t) - H(\varphi_t - \varphi_{t-1}) + \delta V(\varphi_t) \right\}. \quad (1)$$

Denote $\Pi^S(Q,\varphi) = R^S(Q) - C(Q,\varphi)$ as the current period profit (excluding the adjustment costs) and $\Gamma^S(Q_t,\varphi_t,\varphi_{t-1}) = \Pi^S(Q_t,\varphi_t) - H(\varphi_t - \varphi_{t-1})$ as the current period net profit. (1) states the principle of optimality: the optimal value of the firm starting from this period can be obtained by optimizing profit it can earn this period plus its optimal value starting from next period.

Even though the existence of a solution to a general Bellman equation is not guaranteed, (1) has some nice properties. For example, $\Gamma^S(Q_t,\varphi_t,\varphi_{t-1})$ is strictly concave in φ_{t-1} and is jointly strictly concave in (Q_t,φ_t). This is because $H(\cdot)$ is strictly convex, $C(Q,\varphi)$ is weakly concave in (Q_t,φ_t), and $R^S(Q)$ is strictly concave in Q. Given these properties of Γ^S, we can

[2] See Das (1991) for example.

conclude that there is a unique solution to (1) and the $V^S(\varphi)$ and $V(\varphi)$ functions are strictly concave.[3]

Differentiating the right-hand side of (1) with respect to Q_t and φ_t, we have the first-order conditions for the maximization problem (1):

$$\Pi_Q^S(Q_t, \varphi_t) = 0, \tag{2}$$

and

$$\Pi_\varphi^S(Q_t, \varphi_t) - H'(\varphi_t - \varphi_{t-1}) + \delta V'(\varphi_t) = 0. \tag{3}$$

From (2) and (3), we can obtain the optimal output function $Q_t = Q^S(\varphi_t)$ and the optimal capacity adjustment function $\varphi_t = \varphi^S(\varphi_{t-1})$. As we shall prove in Lemma 1, they are both increasing functions. This is quite intuitive. A larger capacity makes a larger scale of production cheaper. Therefore, the firm tends to produce more. Meanwhile, capacities are costly to adjust. Therefore, a larger capacity in the last period cultivates a larger capacity in the current period under the same demand condition.

LEMMA 1. *(i)* $\dfrac{dQ^S(\varphi_t)}{d\varphi_t} > 0$; *(ii)* $\dfrac{d\varphi^S(\varphi_{t-1})}{d\varphi_{t-1}} \geq 0$.

Proof. See Appendix. ▌

The following lemma states that given the same level of capacity, the firm will produce more in the high-demand state than in the low-demand state. This seems natural, as a higher demand means higher price for the firm, which in turn means production is more profitable, inducing the firm to produce more.

LEMMA 2. $\forall \varphi \geq 0$, $Q^H(\varphi) > Q^L(\varphi)$.

Proof. See Appendix. ▌

Let φ^* be the φ that maximizes $V^H(\varphi)$. This φ^* is the ideal capacity for the firm when it faces the high-demand state. Similarly, we can denote φ_* as the ideal capacity for the firm when it faces the low-demand state. Of course, it maximizes $V^L(\varphi)$. Given that last period's capacity is φ^*, the firm is certainly not going to adjust it when the demand is indeed high. (Otherwise, it is not the ideal capacity.) Similar arguments can be made

[3]See, for example, Stokey and Lucas (1989, p.263), Theorem 9.6. The problem of non-existence of an optimal solution in Example 1.1 of Ross (1983, p.74) does not arise here because of the discount factor.

for φ_*. These assessments can be confirmed by applying the Envelope Theorem to (1):

$$\frac{dV^H(\varphi^*)}{d\varphi} = H'(\varphi^H(\varphi^*) - \varphi^*) = 0$$

and

$$\frac{dV^L(\varphi_*)}{d\varphi} = H'(\varphi^L(\varphi_*) - \varphi_*) = 0.$$

From $H'(0) = 0$ and $H''(\cdot) > 0$, we can conclude that $\varphi^H(\varphi^*) = \varphi^*$ and $\varphi^L(\varphi_*) = \varphi_*$.

Since jumping to these ideal capacities is costly, the firm gradually adjusts its capacities towards these ideal ones (i.e., (ii) in Lemma 5). Eventually, when high (or low) demands occur consecutively, with very small probability, of course, the capacity approaches φ^* (or φ_*) in the limit (i.e., (iii) in Lemma 5). It is not surprising to see that the observed capacity is usually between φ^* and φ_*. Given this property of slow adjustment, the firm responds to a one-unit change in the past capacity by less than a one-unit change in its current capacity (i.e., (i) in Lemma 5). Obviously, these descriptions make sense only if φ^* is greater than φ_*, which is confirmed in the first part of Lemma 3.

We now turn to issues concerning the firm's output. Usually, how much the firm actually produces in each demand state will depend on the last period's capacity and we therefore cannot conclude that a firm always produces more in the high-demand state. (Imagine the comparison between the firm's outputs when it faces a high demand but its existing capacity is very small and when it faces a low demand but its existing capacity is very large.) Nevertheless, when the firm is equipped with its ideal capacities, it is quite obvious that it will produce more in the high demand state than in the low demand state. This is stated in the second part of Lemma 3.

Lemmas 3 and 5 below summarize the properties of φ^* and φ_* described above. Lemma 4 has no direct economic interpretation; it states that the curvature of the firm's value function is bounded, in a specific way, by the curvature of the firm's profit function. The reason it is stated as a lemma is because it is needed in the proof of Lemma 5.

LEMMA 3. *$\varphi^* > \varphi_*$;* $\quad Q^H(\varphi^*) > Q^L(\varphi_*)$.

Proof. See Appendix. ∎

LEMMA 4.

$$\delta V''(\varphi) < \frac{\Pi^2_{Q\varphi}(Q^S(\varphi),\varphi)}{\Pi_{QQ}(Q^S(\varphi),\varphi)} - \Pi_{\varphi\varphi}(Q^S(\varphi),\varphi), \quad \forall \varphi \in [\varphi_*, \varphi^*].$$

Proof. It results directly from (19) in the proof of Lemma 3. ▮

LEMMA 5.
(i) $\frac{d\varphi^S(\varphi)}{d\varphi} < 1$, $\forall \varphi \in [\varphi_*, \varphi^*]$;
(ii) $\varphi^H(\varphi) \in (\varphi, \varphi^*]$, $\varphi^L(\varphi) \in [\varphi_*, \varphi)$, $\forall \varphi \in [\varphi_*, \varphi^*]$;
(iii) $\varphi^H(\varphi^H(\varphi^H(\cdots\varphi))) \to \varphi^*$, $\varphi^L(\varphi^L(\varphi^L(\cdots\varphi))) \to \varphi_*$, $\forall \varphi \in [\varphi_*, \varphi^*]$.

Proof. See Appendix. ▮

Given Lemmas 1 to 5, we are now ready to describe the optimal path of production of the monopoly firm, which is the main objective of this section. In the first period, the firm will choose capacity φ_0 that is equal to either φ^* or φ_* depending on the initial state of demand (say, S), produce $Q = Q^S(\varphi_0)$, and charge a price that clears the market. In the subsequent periods, the firm chooses the current capacity φ_t and output Q_t according to $\varphi_t = \varphi^S(\varphi_{t-1})$ and $Q_t = Q^S(\varphi_t)$ depending on the current demand situation S. In period t, φ_t must be between φ_* and φ^*; if the high-demand (or low-demand) state occurs consecutively, φ_t approaches φ^* (or φ_*) monotonically and asymtotically.

Because capacity is costly to alter, the adjustment of capacity is not as quick as in the absence of adjustment costs. Therefore, we may suspect that there is a positive correlation between the capacity in any two consecutive periods, even though the demands in those periods are independent. Because of this, the outputs in these periods could also be positively correlated. To claim these relationships, we need to establish one more lemma.

LEMMA 6. *Suppose that $g_1(x)$ and $g_2(x)$ are both positive and strictly increasing functions, and the random variable x has a non-degenerate c.d.f. F. Then*

$$\int_{-\infty}^{\infty} g_1(x)g_2(x)dF(x) > \int_{-\infty}^{\infty} g_1(x)dF(x) \cdot \int_{-\infty}^{\infty} g_2(x)dF(x).$$

Proof. See Appendix. ▮

Lemma 6 relates two random variables, generated by $g_1(x)$ and $g_2(x)$ respectively. Because they are both positive and increasing functions, when x is large, both $g_1(x)$ and $g_2(x)$ become large. When x is small, both $g_1(x)$ and $g_2(x)$ become small. Therefore, there is positive correlation between the two. In expectation terms, the inequality in Lemma 6 can be rewritten as $E(g_1(x)g_2(x)) > E(g_1(x)) \cdot E(g_2(x))$, or equivalently, $E\{[g_1(x) - E(g_1(x))][g_2(x) - E(g_2(x))]\} > 0$. This implies that $g_1(x)$ and $g_2(x)$ are statistically positively correlated.

Looking ahead from period 0, the firm's optimal capacities and outputs in periods $t-1$ and t, φ_{t-1}, φ_t, Q_{t-1}, and Q_t are all random variables. Since $\varphi_t = \varphi^S(\varphi_{t-1})$ (where S is the demand state in period t) is an increasing function (from Lemma 1), we can conclude from Lemma 6 that φ_t and φ_{t-1} are two positively correlated variables. Furthermore, since $Q_{t-1} = Q^{S'}(\varphi_{t-1})$ (where S' is the demand state in period $t-1$) and $Q_t = Q^S(\varphi_t) = Q^S(\varphi^S(\varphi_{t-1}))$ are both positive and strictly increasing functions of φ_{t-1}, from Lemma 6, Q_{t-1} and Q_t are also positively correlated. Hence, we have the following theorem; its proof is directly from Lemma 6.

Theorem 1. *$E(\varphi_{t-1} \cdot \varphi_t) > E(\varphi_{t-1}) \cdot E(\varphi_t)$; $E(Q_{t-1} \cdot Q_t) > E(Q_{t-1}) \cdot E(Q_t)$. That is, φ_{t-1} and φ_t are positively correlated, and Q_{t-1} and Q_t are also positively correlated.*

Even though demands are independent across time, the capacities and the outputs are interdependent because of the presence of adjustment costs in the capacity.[4] This implies that capacities and outputs do not fluctuate as severely as the industry demand. In some sense, capacities in this model serve as a stabilizer for the industry output. The results obtained in this paper should also apply to the situation where the costs of production depend solely on output and where there are costs for adjusting outputs. It will become clearer in the next section that the optimal collusive path of oligopolistic competitors also shares the same properties.

3. THE CASE OF A DUOPOLY

In previous section, there is only one firm in the industry. When there is more than one firm competing with another, is capacity still used as a stabilizer to absorb the shocks in demand? Or is it used as an amplifier to exaggerate the shocks? When is a firm most tempted to deviate from the collusive strategy? These and other questions will be investigated in this section.

Suppose that there are two firms competing in the industry. (The case of N firms can be considered similarly. See Footnote 5.) These two firms have

[4]Scheinkman and Weiss (1986) show similarly that the equilibrium path of outputs in their model exhibits higher order serial correlation than the exogenous uncertainty.

identical production technology, which is the same as in the last section. As in the case of a single firm, these firms observe the state of demand (S) in the current period before they make their decisions on φ, Q and P simultaneously. All φ, Q, and P of both firms in the past become common knowledge. We shall look for a symmetric capacity and production path that maximizes their joint profits. Given that their initial capacities are the same, the firms will select, along the path, the same capacity, produce the same amount of output, and set the same price. Since the cost function $C(Q, \varphi)$ is convex, splitting their output in any period minimizes the total cost of production; that is, $C(2Q, \varphi) > 2C(Q, \varphi)$. Let Q_t denote the production of one firm in period t. Then $2Q_t$ represents the joint output in the collusive path. Replacing $P^S(Q_t)$ by $P^S(2Q_t)$ in (1), since now there are two firms each producing Q_t. We define the solution to the following equations as their optimal collusive path of production:[5]

$$V(\varphi_{t-1}) = \theta V^H(\varphi_{t-1}) + (1-\theta)V^L(\varphi_{t-1}) \tag{4}$$

and

$$V^S(\varphi_{t-1}) = \max_{Q,\varphi} \left\{ P^S(2Q)Q - C(Q,\varphi) - H(\varphi - \varphi_{t-1}) + \delta V(\varphi) \right\}. \tag{5}$$

(5) is similar to (1) in the case of monopoly. Therefore, we can similarly confirm the existence and the uniqueness of the solution to the above dynamic optimization problem. It is easy to see that the properties obtained in the previous section also apply to this optimal collusive path. Therefore, capacity in the case of duopoly also serves as a stabilizer for outputs under demand shocks.

Because the firms are competing with each other, a collusive plan may not be viable. At any time, if it wishes to, a firm may undercut the other firm's price and take over the entire market. (Consumers choose to purchase from the firm with the lower price. If the prices are the same, they choose each firm with equal probability.) Of course, this firm's take-over is temporary. The other firm may retaliate the next period and charge a lower price. This price war could last for a very long period of time. There are two issues that we intend to investigate. The first issue is under what circumstance the two firms can cooperate all the time. Note that this cooperation is tacit. Overt agreements on collusion are not allowed by law and therefore not enforceable. Still, firms may not deviate and undercut their rival's price due to the loss of profits during the likely retaliation. The second issue is how capacity is related to the possibility of deviation. Is deviation more

[5]Replace $P^S(2Q)$ by $P^S(NQ)$ in (5) when there are N firms competing in the industry.

likely to happen when a firm's capacity is large, or when a firm's capacity is small. If a firm has a large capacity and the demand is high, it is very tempting for this firm to undercut the other firm's price and take over the entire market. But then the other firm's large capacity makes the price war that follows more nasty and the loss of future revenue more severe. Further investigation is needed on this issue.

We first argue that the collusive path is viable and can be sustained in a subgame-perfect equilibrium when discounting is small (i.e., when δ is large and close to 1). A subgame-perfect equilibrium requires the prescribed strategies to be in Nash equilibrium in every proper subgame. Abreu (1986, 1988) has shown that in a repeated (super)game all subgame-perfect equilibrium payoff paths can be obtained by very simple strategies. It will be more difficult to obtain the subgame-perfect strategies in our game, since the game is not repeated identically in each period (due to the capacity adjustment). Nevertheless, following the lines of Fudenberg and Maskin (1986) and Bernheim and Whinston (1987), in principle, we are able to construct a subgame-perfect equilibrium such that no firm wants to deviate from the optimal collusive path.

We shall see that any attainable levels of profit can be sustained in a subgame-perfect equilibrium, as long as δ is sufficiently large and the level of profit for a firm exceeds its minimax payoff in the one-period game, which is equal to zero here. Suppose that a firm deviates from the proposed path, the other firm can set a large capacity and produce a large amount of output for several periods so that the deviating firm's gain from the deviation is negated by this punishment. After that, the two firms return to the proposed collusive path. In case that the punishing firm deviates from punishment, it will be punished by its rival using a similar strategy. Therefore, to sustain the optimal collusive outcome described, we can punish any deviator by an equilibrium yielding zero profit to the deviating firm. As long as δ is sufficiently close to 1, the collusive path described by (4) and (5) is indeed an equilibrium path in our dynamic game.

In what follows, we shall investigate the incentive for a firm to deviate from the collusive path in each state of the demand. Our goal is to find out whether a firm has more incentive to deviate in the high-demand state or the low-demand state and when the capacity is high or low. To do so, we set up an artificial value function that includes a possible deviation in either state. Let $\lambda \in [0, 1]$ index the demand situation, with 0 representing low demand and 1 representing high demand. Let $k \in [1, 2]$ index the number of firms supplying the market. If they collude, both firms supply the market, and $k = 2$. However, if one firm deviates, only that firm supplies the market (in that period), and $k = 1$. λ and k are allowed to change continuously for technical convenience. Define

$$\begin{aligned} & \tilde{V}(\varphi_{t-1}, k, \lambda) \\ = \ & \max_{Q,\varphi} \{[\lambda P^H(kQ) + (1-\lambda) P^L(kQ)]Q - C(Q, \varphi) \\ - \ & H(\varphi - \varphi_{t-1}) + \delta(k-1)V(\varphi)\} \end{aligned} \tag{6}$$

where $k \in [1, 2]$, $\lambda \in [0, 1]$. Then

$$\tilde{V}(\varphi_{t-1}, 2, 0) = V^L(\varphi_{t-1}),$$

and

$$\tilde{V}(\varphi_{t-1}, 2, 1) = V^H(\varphi_{t-1}).$$

Assume that the optimal price a firm chooses when deviating is always lower than the price in the collusive path. This is reasonable because a deviating firm, anticipating a larger demand for its product during the deviation, would always choose a larger capacity. This makes producing more output cheaper. Thus the optimal price the deviator would charge is very likely to be lower than the collusive price. [6] Let $\tilde{V}(\varphi_{t-1}, 1, 0)$ represent the profit of deviation from the optimal collusive path in the low-demand state, and $\tilde{V}(\varphi_{t-1}, 1, 1)$ represent that profit in the high-demand state. Thus, $\tilde{V}(\varphi_{t-1}, 1, 0) - \tilde{V}(\varphi_{t-1}, 2, 0)$ is the gain from deviation in the low-demand state, and $\tilde{V}(\varphi_{t-1}, 1, 1) - \tilde{V}(\varphi_{t-1}, 2, 1)$ is the gain from deviation in the high-demand state.

Let φ^* and φ_* again be the φ_{t-1} as that maximize $V^H(\varphi_{t-1})$ and $V^L(\varphi_{t-1})$ respectively (c.f. (4) and (5)). As in the previous section a firm will always choose a φ that is between φ^* and φ_* along the collusive path. We further assume that $P^S(0) > C_Q(0, \varphi_*)$. This stronger assumption on $P^S(Q)$ is needed to guarantee that a firm always produces a positive amount of output in any period on the optimal collusive path.

To characterize the equilibrium conditions and to compare the deviation incentives, we need the following lemma:

Lemma 7. *There exists an $\epsilon > 0$ such that if $H'(\varphi^* - \varphi_*) < \epsilon$, then*

$$\frac{\partial^2 \tilde{V}(\varphi_{t-1}, k, \lambda)}{\partial \varphi_{t-1} \partial k} < 0, \tag{7}$$

[6] In the unusual case that some of the prices are higher than those in the collusive path, (A.14) in the proof of Lemma 7 is no longer valid and the deviating firm can charge no higher than the prices in the collusive path. In this case, the firm's deviating profit is lower than what is calculated. We hope that this will not alter the qualitative results in the rest of the paper.

$\forall \varphi_{t-1} \in [\varphi_*, \varphi^*]$, $k \in [1,2]$, $\lambda \in [0,1]$.

Proof. See Appendix. ▌

Lemma 7 characterizes the condition under which the incentive to deviate is an increasing function of capacity, noting that $-\frac{\partial \tilde{V}(\varphi_{t-1},k,\lambda)}{\partial k}$ represents the gradual change from two firms to one firm. The condition $H'(\varphi^* - \varphi_*) < \epsilon$ does not necessarily imply that $H(\cdot)$ has to be small. In the case where $H(\cdot)$ is large, φ^* and φ_* are very close to each other, since adjusting capacity is quite costly. In this case, $H'(\varphi^* - \varphi_*) < \epsilon$ is easy to satisfy.

Integrating (7) from $k = 1$ to $k = 2$, we have

$$\frac{\partial}{\partial \varphi_{t-1}} \left[\tilde{V}(\varphi_{t-1}, 2, \lambda) - \tilde{V}(\varphi_{t-1}, 1, \lambda) \right] < 0,$$

or equivalently,

$$\frac{\partial}{\partial \varphi_{t-1}} \left[\tilde{V}(\varphi_{t-1}, 1, \lambda) - \tilde{V}(\varphi_{t-1}, 2, \lambda) \right] > 0. \tag{8}$$

This implies the following theorem; it characterizes one necessary and sufficient condition under which the optimal path is an equilibrium path:

Theorem 2. *Let ϵ be given by Lemma 7 and let $H'(\varphi^* - \varphi_*) < \epsilon$. Then the optimal collusion path can be sustained in a subgame-perfect equilibrium if and only if*

$$\tilde{V}(\varphi^*, 1, 0) = V^L(\varphi^*) \quad \text{and} \quad \tilde{V}(\varphi^*, 1, 1) = V^H(\varphi^*). \tag{9}$$

Proof. From (8), the gain from deviation is increasing in φ_{t-1} . Therefore φ^* maximizes both $\tilde{V}(\varphi_{t-1}, 1, 0) - V^L(\varphi_{t-1})$ and $\tilde{V}(\varphi_{t-1}, 1, 1)$ $= V^H(\varphi_{t-1})$. If, and only if, the conditions in (9) hold, neither of the deviations is profitable. ▌

This theorem states that the optimal path is an equilibrium path if and only if a firm does not want to deviate at the largest capacity (φ^*). If these firms cooperate when their capacities are the largest, they have more incentive to cooperate when their capacities are smaller. This is because a deviation increases the current-period demand for the deviating firm. Since the adjustment cost function is convex, a larger capacity in the previous period reduces the cost of adjustment due to the increase in output to meet the demand. At the same time, since the deviating firm is punished from the next period on, it earns zero total profit in the future. Therefore, the deviating firm needs not take into account the fluctuation in demand starting from the next period. As long as the adjustment cost involved is small the consideration of adjustment costs has only a small effect which is then dominated by the direct effect of the increasing capacity. Thus, the most profitable time to deviate is when $\varphi_{t-1} = \varphi^*$, which is, therefore, the most likely time for the collusive agreement to break down.

Condition (9) relates the incentive to deviate to a firm's capacity. How the demand situations relate to the deviation incentives is still unknown. Does a firm have more incentive to deviate when the demand is high, or when the demand is low?

Applying the Envelope Theorem to (6), we have for the optimal Q,

$$\frac{\partial \tilde{V}}{\partial \lambda} = \left[P^H(kQ) - P^L(kQ)\right] Q.$$

So

$$\begin{aligned}\frac{\partial^2 \tilde{V}}{\partial k \partial \lambda} &= \left[P^H(kQ) - P^L(kQ)\right] \frac{dQ}{dk} \\ &+ Q\left[P_Q^H(kQ) - P_Q^L(kQ)\right](Q + k\frac{dQ}{dk}).\end{aligned} \tag{10}$$

If $P_Q^H(kQ) - P_Q^L(kQ)$ is sufficiently uniformly small, the price difference between the high demand and the low demand states does not vary too much across different output levels. This implies that an increase in output will be rewarded consistently. Since the increase in output (when a firm deviates) is higher when the total demand is high, given any level of capacity, it is always more tempting to deviate when the demand is high. Mathematically, when $P_Q^H(kQ) - P_Q^L(kQ)$ is sufficiently uniformly small, the second term on the right-hand side of (10) is dominated by the first term. We have the following theorem:

THEOREM 3. *Let ϵ be given by Lemma 7 and let $H'(\varphi^* - \varphi_*) < \epsilon$. If $P_Q^H - P_Q^L$ is sufficiently uniformly small, then the most profitable deviation is when the demand is high.*

Proof. If $H'(\varphi^* - \varphi_*) < \epsilon$ is satisfied, $\frac{d\varphi}{dk} < 0$. From (26) in the proof of Lemma 7, we conclude that $\frac{dQ}{dk} < 0$. When $P_Q^H - P_Q^L$ is sufficiently uniformly small, $\frac{\partial^2 \tilde{V}}{\partial k \partial \lambda} < 0$. Integrating this inequality both sides from $k = 1$ to $k = 2$ and from $\lambda = 0$ to $\lambda = 1$, we have

$$\tilde{V}(\varphi, 2, 1) - \tilde{V}(\varphi, 2, 0) < \tilde{V}(\varphi, 1, 1) - \tilde{V}(\varphi, 1, 0),$$

or equivalently,

$$\tilde{V}(\varphi, 1, 1) - V^H(\varphi) > \tilde{V}(\varphi, 1, 0) - V^L(\varphi).$$

That is, deviation at high demand is most profitable. ▮

Theorems 2 and 3 characterize sufficient conditions under which a firm has more incentive to deviate at higher level of capacity or at high-demand state. As a special case, suppose that a firm has no capacity constraint and the cost of production depends solely on how much a firm produces. In this case, the capacity variable φ does not show up in the cost function, i.e., $C(Q, \varphi)$ reduces to $C(Q)$, and $H(\cdot) \equiv 0$. In this case, simpler conditions could be obtained. We have the following corollary of Theorem 3:

COROLLARY 1. *Suppose that $P^H_Q(Q) > P^L_Q(Q)$, $\forall Q$. Then a firm has more incentive to deviate in the high-demand state if $C''(Q)$ is sufficiently uniformly small. A firm has more incentive to deviate in the low-demand state if $C''(Q)$ is sufficiently uniformly large.*

Proof. See Appendix. ▮

When C'' is uniformly small, the marginal cost of production $C'(Q)$ does increase too dramatically when the firm increases its output. Therefore, deviating when the demand is high is more profitable, as the demand stolen from the other firm is larger. One special case is when firms have constant marginal costs. In that case, a similar result has been obtained by Rotemberg and Saloner (1986).

When C'' is sufficiently large, the marginal cost of production $C'(Q)$ increases dramatically, and an equal increase in the output will cost the firm more when the output is higher. Since it is too costly to increase output when the output level is already high, a firm prefers to deviate when the demand is low.

4. CONCLUDING REMARKS

In this paper, we presented and analyzed a model in which a firm's cost of production depends on both its output and capacity. This model is different from most conventional models in that the adjustment of capacity is costly. Because of this adjustment cost, a firm's capacity acts like a stabilizer for its output, which does not fluctuate as violently as the market demand. This feature of our model is unique; previous models do not possess such a property.

In the previous sections, we characterized the optimal path of production for a monopoly industry and for a duopoly industry. When there are more than two firms, the analysis can be directly generalized. It was shown that the optimal collusive path of production of a duopoly exhibits the same properties as the optimal path of a monopoly firm, even though a firm may want to deviate from time to time. Furthermore, we found that when the cost of adjustment is sufficiently low, the greatest incentive to deviate occurs at the largest capacity. We also derived conditions under which deviations are more profitable when the demand is high, regardless of the firms' capacities.

The analysis in this paper relates a firm's productivity to its capacity under demand uncertainty and costs of adjustment. A firm is less efficient if it has a small capacity but produces a large amount, or if it has a large capacity but produces a small amount. Our analysis shows that a firm is more efficient when the demand is more consistent. High efficiency takes place when high demand occurs consecutively, or when low demand occurs consecutively. If demand fluctuates tremendously, efficiency is harder to obtain.

The model in this paper could be used to analyze the production decisions of firms of different sizes, especially those with cost functions which are not identical. It can also be used to analyze situations where firms produce related but not identical products, and demand uncertainty affects all products in a similar way. Of course, the calculations involved could become very complex and useful conclusions may or may not be forthcoming. As these studies may never be carried out (due the complexity involved), the intuition developed in this paper could provide a useful reference for those more complicated situations.

APPENDIX: PROOFS OF LEMMAS AND THEOREMS IN THE TEXT

Proof of Lemma 1

(i) Differentiating (2) with respect to φ_t , we have

$$\Pi^S_{QQ}(Q_t,\varphi_t)\frac{dQ^S(\varphi_t)}{d\varphi_t}+\Pi^S_{Q\varphi}(Q_t,\varphi_t)=0. \tag{A.1}$$

$\Pi^S_{QQ}<0$ and $\Pi^S_{Q\varphi}=-C_{Q\varphi}>0$ imply that $\frac{dQ^S(\varphi_t)}{d\varphi_t}>0$.

(ii) Let $\varphi_t=\varphi^S(\varphi_{t-1})$ and $\hat{\varphi}_t=\varphi^S(\hat{\varphi}_{t-1})$. We have

$$\begin{aligned}&\Pi^S(Q^S(\varphi_t),\varphi_t)-H(\varphi_t-\varphi_{t-1})+\delta V(\varphi_t)\\ \geq\ &\Pi^S(Q^S(\hat{\varphi}_t),\hat{\varphi}_t)-H(\hat{\varphi}_t-\varphi_{t-1})+\delta V(\hat{\varphi}_t)\end{aligned} \tag{A.2}$$

and

$$\begin{aligned} & \Pi^S(Q^S(\hat{\varphi}_t), \hat{\varphi}_t) - H(\hat{\varphi}_t - \hat{\varphi}_{t-1}) + \delta V(\hat{\varphi}_t) \\ \geq \; & \Pi^S(Q^S(\varphi_t), \varphi_t) - H(\varphi_t - \hat{\varphi}_{t-1}) + \delta V(\varphi_t). \end{aligned} \tag{A.3}$$

Adding (A.2) and (A.3), simplifying, and rearranging terms, we have

$$[H(\hat{\varphi}_t - \varphi_{t-1}) - H(\varphi_t - \varphi_{t-1})] - [H(\hat{\varphi}_t - \hat{\varphi}_{t-1}) - H(\varphi_t - \hat{\varphi}_{t-1})] \geq 0.$$

Since $H(\cdot)$ is twice continuously differentiable, using Taylor's Theorem twice, we have

$$H''(\tilde{\varphi}_t - \tilde{\varphi}_{t-1})(\hat{\varphi}_t - \varphi_{t-1})(\hat{\varphi}_t - \varphi_{t-1}) \geq 0, \tag{A.4}$$

where $\tilde{\varphi}_{t-1}$ is between $\hat{\varphi}_t$ and φ_t, and $\tilde{\varphi}_{t-1}$ is between $\hat{\varphi}_{t-1}$ and φ_{t-1}. $H'' > 0$ and (A.4) imply that if $\hat{\varphi}_{t-1} > \varphi_{t-1}$, then $\hat{\varphi}_t \geq \varphi_t$. That is, $\varphi^S(\varphi_{t-1})$ is monotone increasing. ∎

Proof of Lemma 2 Let

$$P(Q;\alpha) \equiv \alpha P^H(Q) + (1-\alpha)P^L(Q), \;\; \alpha \in [0,1], \tag{A.5}$$

and

$$\Pi(Q,\varphi;\alpha) \equiv P(Q;\alpha)Q - C(Q,\varphi). \tag{A.6}$$

So $P(Q;0) = P^L(Q)$, and $P(Q,1) = P^H(Q)$. Maximizing Π with respect to Q, we have

$$\Pi_Q(Q,\varphi;\alpha) = 0, \tag{A.7}$$

and

$$\Pi_{QQ}(Q,\varphi;\alpha) \leq 0.$$

Differentiating (A.7) with respect to α, we have

$$\Pi_{QQ}\frac{dQ}{d\alpha} + \Pi_{Q\alpha} = 0.$$

Since

$$\Pi_{Q\alpha} = [P^H_Q(Q) - P^L_Q(Q)]Q + [P^H(Q) - P^L(Q)] > 0,$$

we conclude that $\frac{dQ}{d\alpha} > 0$. Therefore, Q is strictly increasing in α and $Q^H(\varphi) > Q^L(\varphi)$. ▮

Proof of Lemma 3 Let $P(Q;\alpha)$ and $\Pi(Q,\varphi;\alpha)$ be as defined in (A.5) and (A.6), and let

$$W(Q,\varphi_t,\varphi_{t-1};\alpha) \equiv \Pi(Q,\varphi_t;\alpha) - H(\varphi_t - \varphi_{t-1}) + \delta V(\varphi_t).$$

Note that $V(\varphi_t)$ does not depend on α, because we vary the demand in the current period (period t) only. Maximizing $W(Q,\varphi_t,\varphi_{t-1};\alpha)$, we have the following first-order conditions:

$$W_Q = 0, \quad W_{\varphi_t} = 0, \quad W_{\varphi_{t-1}} = 0.$$

Differentiating the above equations with respect to α, we have

$$W_{QQ}\frac{dQ}{d\alpha} + W_{Q\varphi_t}\frac{d\varphi_t}{d\alpha} + W_{Q\varphi_{t-1}}\frac{d\varphi_{t-1}}{d\alpha} + W_{Q\alpha} = 0, \qquad \text{(A.8)}$$

$$W_{\varphi_t Q}\frac{dQ}{d\alpha} + W_{\varphi_t\varphi_t}\frac{d\varphi_t}{d\alpha} + W_{\varphi_t\varphi_{t-1}}\frac{d\varphi_{t-1}}{d\alpha} + W_{\varphi_t\alpha} = 0,$$

and

$$W_{\varphi_{t-1}Q}\frac{dQ}{d\alpha} + W_{\varphi_{t-1}\varphi_t}\frac{d\varphi_t}{d\alpha} + W_{\varphi_{t-1}\varphi_{t-1}}\frac{d\varphi_{t-1}}{d\alpha} + W_{\varphi_{t-1}\alpha} = 0.$$

Let

$$A = \begin{pmatrix} W_{QQ} & W_{Q\varphi_t} & W_{Q\varphi_{t-1}} \\ W_{\varphi_t Q} & W_{\varphi_t\varphi_t} & W_{\varphi_t\varphi_{t-1}} \\ W_{\varphi_{t-1}Q} & W_{\varphi_{t-1}\varphi_t} & W_{\varphi_{t-1}\varphi_{t-1}} \end{pmatrix}.$$

Then

$$\begin{pmatrix} \frac{dQ}{d\alpha} \\ \frac{d\varphi_t}{d\alpha} \\ \frac{d\varphi_{t-1}}{d\alpha} \end{pmatrix} = -A^{-1} \cdot \begin{pmatrix} W_{Q\alpha} \\ W_{\varphi_t\alpha} \\ W_{\varphi_{t-1}\alpha} \end{pmatrix}.$$

Simple calculations show that

$$W_{Q\alpha} = [P_Q^H(Q) - P_Q^L(Q)]Q + [P^H(Q) - P^L(Q)] > 0,\ W_{\varphi_t\alpha} = 0,\ W_{\varphi_{t-1}\alpha} = 0,$$

$$W_{QQ} = \Pi_{QQ}, \quad W_{Q\varphi_t} = \Pi_{Q\varphi_t} = -C_{Q\varphi_t} > 0, \quad W_{Q\varphi_{t-1}} = 0,$$

$$W_{\varphi_t\varphi_t} = \Pi_{\varphi_t\varphi_t} - H'' + \delta V'', \quad W_{\varphi_t\varphi_{t-1}} = H'' > 0,$$

$$W_{\varphi_{t-1}\varphi_{t-1}} = H'' < 0,$$

and

$$| A |= -\Pi_{QQ} H'' \left[\delta V'' + \Pi_{\varphi\varphi} - \frac{\Pi_{Q\varphi}^2}{\Pi_{QQ}} \right].$$

From the second-order condition for the maximization, A is negative definite. Thus, $| A |< 0$. Since $\Pi_{QQ} < 0$, and $H'' > 0$, we have

$$\delta V'' + \Pi_{\varphi\varphi} - \frac{\Pi_{Q\varphi}^2}{\Pi_{QQ}} < 0. \tag{A.9}$$

From

$$\begin{pmatrix} \frac{dQ}{d\alpha} \\ \frac{d\varphi_t}{d\alpha} \\ \frac{d\varphi_{t-1}}{d\alpha} \end{pmatrix} = -\frac{1}{| A |} \begin{pmatrix} H''(C_{\varphi\varphi} - \delta V'') & * & * \\ C_{Q\varphi_t} H'' & * & * \\ C_{Q\varphi_t} H'' & * & * \end{pmatrix} \begin{pmatrix} W_{Q\alpha} \\ 0 \\ 0 \end{pmatrix}, \tag{A.10}$$

we have

$$\frac{d\varphi_t}{d\alpha} = \frac{d\varphi_{t-1}}{d\alpha} = \frac{C_{Q\varphi_t} H'' W_{Q\alpha}}{| A |} > 0.$$

Hence, φ_t and φ_{t-1} are increasing functions of α. Since φ^* is the optimal φ_t corresponding to $\alpha = 1$, and φ_* is the optimal φ_t corresponding to $\alpha = 0$, we conclude that $\varphi^* > \varphi_*$.

Since $W_{QQ} < 0$, $W_{Q\varphi_t} > 0$, $\frac{d\varphi_t}{d\alpha} > 0$, $W_{Q\varphi_{t-1}} = 0$, and $W_{Q\alpha} > 0$, from (A.8) we conclude that

$$\frac{dQ}{d\alpha} > 0. \tag{A.11}$$

As $Q^H(\varphi^*)$ and $Q^L(\varphi_*)$ are the optimal Q's corresponding to $\alpha = 1$ and $\alpha = 0$ respectively, $Q^H(\varphi^*) > Q^L(\varphi_*)$. ▌

Proof of Lemma 5

(i) Differentiating (3) with respect to φ_{t-1} and solving for $\frac{d\varphi_t}{d\varphi_{t-1}}$, we have

$$\frac{d\varphi_t}{d\varphi_{t-1}} = \frac{H''}{H'' - \left[\frac{d\Pi^S_\varphi}{d\varphi_t} + \delta V''\right]}.$$

From (A.1) and Lemma 4, we have

$$\frac{d\Pi^S_\varphi}{d\varphi_t} + \delta V'' = \Pi_{\varphi Q}\frac{dQ}{d\varphi_t} + \Pi_{\varphi\varphi} + \delta V'' = \Pi_{\varphi\varphi} - \frac{\Pi^2_{Q\varphi}}{\Pi_{QQ}} + \delta V'' < 0.$$

Therefore, $\frac{d\varphi_t}{d\varphi_{t-1}} < 1$.

(ii) Integrating $\frac{d\varphi_t}{d\varphi_{t-1}} < 1$, both sides from φ to φ^*, we have $\varphi^H(\varphi^*) - \varphi^H(\varphi) < \varphi^* - \varphi$. Since $\varphi^H(\varphi^*) = \varphi^*$, we have $\varphi^H(\varphi) > \varphi$. Suppose that $\varphi^H(\varphi) = \tilde{\varphi} > \varphi^*$. Then

$$\begin{aligned}
& \Pi(Q,\tilde{\varphi}) - H(\tilde{\varphi} - \varphi) + \delta V(\tilde{\varphi}) \\
< \; & \Pi(Q,\tilde{\varphi}) - [H(\tilde{\varphi} - \varphi^*) + H(\varphi^* - \varphi)] + \delta V(\tilde{\varphi}) \\
= \; & [\Pi(Q,\tilde{\varphi}) - H(\tilde{\varphi} - \varphi^*) + \delta V(\tilde{\varphi})] - H(\varphi^* - \varphi) \\
\le \; & [\Pi(Q,\varphi^*) - H(\varphi^* - \varphi^*) + \delta V(\varphi^*)] - H(\varphi^* - \varphi) \\
= \; & \Pi(Q,\varphi^*) - H(\varphi^* - \varphi) + \delta V(\varphi^*)
\end{aligned}$$

where we made use of the convexity of $H(\cdot)$ in the first inequality, φ^* maximizing $V^H(\varphi)$ in the second inequality, and $H(0) = 0$ in the last equality. Therefore, $\varphi^H(\varphi) > \varphi^*$ can never happen. Together with $\varphi^H(\varphi) = \varphi$, we have $\varphi^H(\varphi) \in (\varphi, \varphi^*]$. Similarly, we can prove that $\varphi^L(\varphi) \in [\varphi^*, \varphi)$.

(iii) From (ii) $\varphi^* \ge \varphi^H(\varphi) > \varphi$, $\forall\varphi \in [\varphi_*, \varphi^*]$. Therefore,

$$\varphi^H_n(\varphi) \equiv \overbrace{\varphi^H(\varphi^H(\varphi^H}^{n}(\cdots\varphi)))$$

is an increasing sequence but bounded by φ^*, and thus must have a limit. From (ii), $\forall\hat{\varphi} < \varphi^*$, $\varphi^H(\hat{\varphi}) > \hat{\varphi}$; thus $\hat{\varphi}$ cannot be the limit of $\varphi^H_n(\varphi)$. Hence,

$$\lim_{n\to\infty} \varphi^H_n(\varphi) = \varphi^*, \ \forall\varphi \in [\varphi_*, \varphi^*].$$

Similarly, we can prove that φ_* is the limit of $\varphi^L(\varphi^L(\varphi^L(\cdots\varphi)))$, $\forall\varphi \in [\varphi_*, \varphi^*]$. ∎

Proof of Lemma 6

Let $E_1 = \int_{-\infty}^{\infty} g_1(x)dF(x)$ be the mean of $g_1(x)$. Then $\exists x^*$, such that $g_1(x) > E_1$ if $x > x^*$, and $g_1(x) < E_1$ if $x < x^*$. Since $g_2(x)$ is positive and strictly increasing, we have

$$\int_{x^*}^{\infty} [g_1(x) - E_1] \cdot g_2(x)dF(x) > \int_{x^*}^{\infty} [g_1(x) - E_1] \cdot g_2(x^*)dF(x), \quad \text{(A.12)}$$

$$\int_{-\infty}^{x^*} [g_1(x) - E_1] \cdot g_2(x)dF(x) > \int_{-\infty}^{x^*} [g_1(x) - E_1] \cdot g_2(x^*)dF(x). \quad \text{(A.13)}$$

Adding the above (A.12) and (A.13), we have

$$\begin{aligned}\int_{-\infty}^{\infty} [g_1(x) - E_1] \cdot g_2(x)dF(x) &> \int_{-\infty}^{\infty} [g_1(x) - E_1] \cdot g_2(x^*)dF(x) \\ &= g_2(x^*) \int_{-\infty}^{\infty} [g_1(x) - E_1]dF(x) = 0.\end{aligned}$$

i.e.,

$$\int_{-\infty}^{\infty} g_1(x) \cdot g_2(x)dF(x) > \int_{-\infty}^{\infty} g_1(x)dF(x) \cdot \int_{-\infty}^{\infty} g_2(x)dF(x).$$

▮

Proof of Lemma 7 Simplifying notations, let $P(Q) = \lambda P^H(Q) + (1-\lambda)P^L(Q)$. Differentiating the right-hand side of (6) with respect to Q and φ, we have the following first order conditions for maximization:

$$P(kQ) + kQP'(kQ) - C_Q(Q, \varphi) = 0, \quad \text{(A.14)}$$

and

$$-C_\varphi(Q, \varphi) - H'(\varphi - \varphi_{t-1}) + \delta(k-1)V'(\varphi) = 0. \quad \text{(A.15)}$$

Differentiating (A.14) and (A.15) with respect to k, we have

$$R''(kQ)(Q + k\frac{dQ}{dk}) - C_{QQ}\frac{dQ}{dk} - C_{Q\varphi}\frac{d\varphi}{dk} = 0, \quad \text{(A.16)}$$

and

$$-C_{Q\varphi}\frac{dQ}{dk} - C_{\varphi\varphi}\frac{d\varphi}{dk} - H''\frac{d\varphi}{dk} + \delta V' + \delta(k-1)V''\frac{d\varphi}{dk} = 0. \quad \text{(A.17)}$$

Solving for $\frac{d\varphi}{dk}$ and $\frac{dQ}{dk}$ in (A.16) and (A.17), we have

$$\frac{d\varphi}{dk} = \frac{-QR''C_{Q\varphi} - \delta V'(\varphi)[kR'' - C_{QQ}]}{[-C_{\varphi\varphi} - H'' + \delta(k-1)V''](kR'' - C_{QQ}) - C_{Q\varphi}^2}, \tag{A.18}$$

and

$$\frac{dQ}{dk} = \frac{1}{C_{Q\varphi}} \cdot \frac{[-C_{\varphi\varphi} - H'' + \delta(k-1)V''](-QR''C_{Q\varphi}) - \delta V'C_{Q\varphi}^2}{[-C_{\varphi\varphi} - H'' + \delta(k-1)V''](kR'' - C_{QQ}) - C_{Q\varphi}^2}. \tag{A.19}$$

Since $R_{QQ}^S < 0$, $R'' = \lambda R_{QQ}^H + (1-\lambda)R_{QQ}^L < 0$. If $V'' \le 0$, then because $C_{\varphi\varphi}C_{QQ} - C_{Q\varphi}^2 > 0$, the denominator of (A.18) is positive. If $V'' > 0$, $\delta(k-1)V'' \le \delta V''$ because $k \in [1,2]$. From Lemma 4,

$$\delta V''(\varphi) < \frac{\Pi_{Q\varphi}^2}{\Pi_{QQ}} - \Pi_{\varphi\varphi} = \frac{C_{Q\varphi}^2}{R'' - C_{QQ}} + C_{\varphi\varphi}.$$

Therefore,

$$-C_{\varphi\varphi} - H'' + \delta(k-1)V'' < \frac{C_{Q\varphi}^2}{R'' - C_{QQ}} - H'' < 0.$$

Since

$$\frac{C_{Q\varphi}^2}{R'' - C_{QQ}}(kR'' - C_{QQ}) - C_{Q\varphi}^2 > 0,$$

we conclude that the denominator of (A.18) is positive. Applying the Envelope Theorem to (5), we have

$$\frac{dV^S(\varphi_{t-1})}{d\varphi_{t-1}} = H'(\varphi - \varphi_{t-1}).$$

Since $H'' > 0$, φ and $\varphi_{t-1} \in [\varphi_*, \varphi^*]$, $H'(\varphi - \varphi_{t-1}) \le H'(\varphi^* - \varphi_*)$. From (4),

$$V'(\varphi_{t-1}) = \theta\frac{dV^S(\varphi_{t-1})}{d\varphi_{t-1}} + (1-\theta)\frac{dV^S(\varphi_{t-1})}{d\varphi_{t-1}} \le H'(\varphi^* - \varphi_*).$$

Let $\epsilon = \min_{Q,\varphi,k} \frac{QC_{Q\varphi}R''}{\delta[-R''k + C_{QQ}]}$. Since Q is bounded from below, $\varphi_t \in [\varphi_*, \varphi^*]$, and $k \in [0,1]$, we have $\epsilon > 0$. Thus, if $H'(\varphi^* - \varphi_*) < \epsilon$,

$$V'(\varphi_{t-1}) = H'(\varphi - \varphi_{t-1}) < H'(\varphi^* - \varphi_*) < \epsilon.$$

From (A.18), $\frac{d\varphi}{dk} < 0$. Therefore,

$$\frac{\partial^2 \tilde{V}(\varphi_{t-1}, k, \lambda)}{\partial \varphi_{t-1} \partial k} = \frac{H'(\varphi - \varphi_{t-1})}{dk} = H''(\varphi - \varphi_{t-1}) \frac{\varphi}{dk} < 0.$$

▌

Proof of Corollary Letting $C_{Q\varphi} = 0$ in (A.16) and solving for $\frac{dQ}{dk}$, we have

$$\frac{dQ}{dk} = \frac{R''Q}{C'' - R''k} < 0.$$

From (10), we have

$$\frac{\partial^2 \tilde{V}}{\partial k \partial \lambda} = \frac{R''Q(P^H - P^L) + C''Q^2(P_Q^H - P_Q^L)}{C'' - R''k}. \tag{A.20}$$

Recall that $R'' < 0$. If C'' is sufficiently uniformly small, (A.20) is always negative. By constructing a parallel analysis to the proof of Theorem 3, we can show that the firm has more incentive to deviate in the high-demand state. If C'' is sufficiently uniformly large, (A.20) is always positive, and the firm has more incentive to deviate in the low-demand state. ▌

REFERENCES

Abreu, Dilip, 1986, Extremal equilibrium of oligopolistic supergames. *Journal of Economic Theory* **39**, 191-225.

Abreu, Dilip, 1988, On the theory of infinitely repeated games with discounting. *Econometrica* **56**, 383-396.

Bernheim, B. Douglas, and Michael D. Whinston, 1987, Multimarket contact and collusive behavior. Harvard University Discussion Paper #1317.

Brock, William A. and Jose A. Scheinkman, 1985, Price setting supergames with capacity constraints. *Review of Economic Studies* **LII**, 371-382.

Das, Sanghamitra, 1991, Are convex adjustment costs necessary for investment smoothing? *Southern Economic Journal* **58**, 268-272.

Eisner, Robert and Robert Strotz, 1963, Determinants of business investment. Research study two in: *Impacts of Monetary Policy*, prepared for the commission on money and credit. Englewood Cliffs: Prentice-Hall.

Fudenberg, Drew and Eric Maskin, 1986, The folk theorem in repeated games with discounting and with incomplete information. *Econometrica* **54**, 533-554.

Green, Edward J. and Robert H. Porter, 1984, Non-cooperative collusion under imperfect price information. *Econometrica* **52**, 87-100.

Ross, Sheldon M, 1983, *Introduction to Stochastic Dynamic Programming.* San Diego: Academic Press.

Rotemberg, Julio J. and Garth Saloner, 1986, A supergame-theoretic model of price wars during booms. *American Economic Review* **76**, 390-407.

Scheinkman, Jose A. and Lawrence Weiss, 1986, Borrowing constraints and aggregate economic activity. *Econometrica* **54**, 23-45.

Staiger, Robert W. and Frank A. Wolak, 1992, Collusive pricing with capacity constraints in the presence of demand uncertainty. *RAND Journal of Economics* **23**, 203-220.

Stokey, Nancy L. and Robert E. Lucas, Jr., 1989, *Recursive Methods in Economic Dynamics*. Cambridge: Harvard University Press.

Weitzman, Martin L., 1982, Increasing returns and the foundations of unemployment theory. *Economic Journal* **92**, 787-804.