WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

STRATEGIC INVENTORIES AND THE EXCESS VOLATILITY OF PRODUCTION

BY

JULIO J. ROTEMBERG AND GARTH SALONER*

SSM WP#1650-85 APRIL 1985

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139

STRATEGIC INVENTORIES AND THE EXCESS VOLATILITY OF PRODUCTION

BY

JULIO J. ROTEMBERG AND GARTH SALONER*

SSM WP#1650-85 APRIL 1985

*Sloan School of Management and Department of Economics respectively. We are very grateful to Robert Porter, James Poterba and Lawrence Summers for many helpful conversations and to an anonymous referee for useful comments. Financial support from the National Science Foundation (grants SES-8209266 and SES-8308782 respectively) is gratefully acknowledged.

STRATEGIC INVENTORIES AND THE EXCESS VOLATILITY

OF PRODUCTION

Julio J. Rotemberg and Garth Saloner*

April 1985

*Sloan School of Management and Department of Economics, M.I.T. respectively.
Financial support from the National Science Foundation (grants SES-8209266 and SES-8308782 respectively) is gratefully acknowledged.

I. Introduction

In this paper we analyze the role of inventories in supporting collusion. This role of inventories derives from their usefulness in punishing deviators from a collusive understanding and is quite different from the one stressed in standard neoclassical models. There they help smooth production; demand is somewhat random and movements in production are costly, either because of explicit costs of changing output or simply due to the concavity of the production function. Then, profit maximizing firms continue to produce substantial quantities even when demand is low in the hope of selling it in future high demand periods. The role of inventories is to reduce the variance of production below that of sales. Unfortunately for this view, in practice the variance of production about a trend tends to be substantially bigger than the variance of sales. This fact as well as related empirical failures of the "production smoothing" model have been documented by, among others, Blinder (1981, 1984) and West (1984).

A classical interpretation of these failures (see for example Eichenbaum (1984) and Blinder (1984)) is that there is substantial variance in productivity. In that case, profit maximizing firms tend to produce a lot, and even accumulate inventories, in periods of relatively high productivity. On the other hand, consumers desire smooth consumption flows leading to relatively small variability in sales. The only difficulty with this explanation is that there is no independent evidence for these productivity shocks so that the explanation remains untested. Moreover the lack of prominence in the discussions of managers and journalists of productivity explanations for inventory movements casts doubt on the existence of such independent evidence.

There is an additional and related puzzle concerning the behavior of inventories. This is that inventories tend to be high in exactly the same periods that sales are high. Table 1 reports coefficients of sales in regressions explaining the level of inventories by a time trend, a square time trend, and the level of sales in a variety of two digit manufacturing industries. These coefficients are often positive and significantly so.[1,2]

Our model of strategic inventories is consistent with the facts and thus may be helpful in explaining them. Since it views inventories namely as deterrents to deviations from implicit collusion, it is closely related to the capcity model of Brock and Scheinkman (1981).

In Brock and Scheinkman (1981), as here, there is an oligopoly which is trying to sustain collusion. As in Radner (1980), Friedman(1971), Green and Porter (1984), Porter(1984) and Rotemberg and Saloner (1985), collusion is sustained by the threat to revert to competitive behavior if any firm deviates. In Brock and Scheinkman (1981) each firm has a capacity constraint. Then, if the firms increase capacity they can use this increased capacity when punishing deviations. Thus, the penalty faced by an individual deviating firm is increased. Therefore, individual firms may be less tempted to deviate and the oligopoly may be able to sustain outcomes with higher profits.

---

[1]Note that the statistical significance of this relationship does not depend on whether sales or inventories are chosen as the dependent variable in the regression.
[2]The coefficients for durable goods industries must be interpreted with caution because inventories here refer only to inventories held by manufacturing firms and a large fraction of total durable inventories are held by retailers. Unfortunately the retailing figures for inventories are not directly comparable to those of manufacturers.

Inventories share with capacity the property that they not only enable firms to respond with higher sales to deviations by any firm, but also make it *ex post* rational to do so. Thus, if firms have more inventories on hand, they punish deviations more severely. However there are three differences between capacity and inventories as disciplining devices.

The first difference is that inventories can only be used to discipline for a short period of time. Nonetheless this may be important if the oligopoly can sustain only little collusion in the absence of inventories. In this sense our paper takes a different view from Friedman (1971) where the infinite reversion to noncooperative behavior that follows any deviation is always sufficient to enforce the cooperative outcome. We believe instead that actual punishments are substantially shorter and smaller. The first reason for this is provided in Green and Porter (1984). In their model the observation of rivals' actions are subject to noise. The firm must attempt to infer from its noisy observation when a deviation has taken place. Thus periods of punishment are triggered not only by an actual deviation but because the noisy observation suggests that it is likely that a deviation occurred. Thus, while an infinite pubnishment has the advantage that it minimizes the incentive to deviate, it has the disadvantage of guaranteeing only a finite period of collusion. In that setting finite punishments will often be optimal.

The second reason for finite punishments is that the identity of the firms in the industry, and their managers, change over time. This may be especially true in situations where the firm faces a sequence of product models. Thus the system only has a finite memory. Once the action that triggered the reversion has been forgotten by the system, it seems likely that the oligopoly will again attempt to reestablish implicitly collusive behavior.[3]

Thirdly, and perhaps more importantly, it is hard to believe that the firms would really resign themselves to an infinite period of punishment. In fact, what makes it plausible to imagine firms actually punishing each other in the first place is that the response to their perception of chiseling on the part of their rivals is likely to be one of anger. However, the anger dissipates over time and the individuals revert to being "sensible".[4]

The second difference between inventories and capacity as deterrents is that the former are more of a two-edged sword. If a firm has more inventories on hand it may also benefit more from deviating. In particular, it is now able to sell more if it does deviate.[5] Moreover, by deviating it avoids the inventory carrying costs.

The third difference, and inventories' main advantage, is that they are less costly to change than capacity. Thus the level of inventories can be tailored to current market conditions. In particular more inventories can be held when, as in Rotemberg and Saloner (1985), the incentive to deviate is higher because demand is temporarily high.

Specifically our model is structured as follows. We consider a duopoly in which both firms have an infinite horizon. Over time the firms offer different "models" of their product. Each model is sold for two periods and

---

[3]A number of methods for moving to an implicitly collusive scheme have been documented in the literature. "Price leadership" and "industrial statesmanship" are the most commonly mentioned.

[4]In this interpretation the "irrational" sentiment of anger serves a useful function. In induces cooperation as others fear anger's consequences. (An application of economic reasoning might suggest that the amount of anger that is triggered by bad behavior is optimal. It balances the energy "wasted" by the anger itself with its social advantages.)

[5]This is also an effect of allowing each firm to build more capacity.

only one model is available (from both firms) at any time. To simplify our paper we assume that inventories are never carried between models or, alternatively, that consumers are willing to pay only very little for older models.[6] Therefore for each model there are three decisions that must be made sequentially. These are the level of production at the beginning of the first period, the level of sales in the first period (or, equivalently, the level of inventories at the end of the first period), and the level of production in the second period. Sales in the second period are simply the sum of the two productions minus the sales in the first period. As in all dynamic programming problems we start the analysis with the last decision.

We assume that future punishments for deviating in the second period of a model are such that collaboration is ensured in this period unless there has been a breakdown of discipline in the first period. If such a breakdown has occurred firms act noncooperatively in the second period. Then, firms with more inventories, actually sell more for one of either of the following two reasons. First, inventories allow firms to sell beyond "capacity" where production beyond capacity is possible but costly. Second, with increasing marginal costs, the same level of sales leads to lower marginal costs in the second period in the presence of inventories. These lower marginal costs induce the firm to expand production further. If firms with more inventories sell more, this means that deviating firms, which usually carry no inventories, have lower profits in the second period. This is the key to the deterrent role of inventories.

In the sales stage of the first period each firm must decide whether to go along with the implicitly agreed upon level of sales or a higher,

---

[6]See Saloner (1984) for a static treatment of model-years with inventories.

currently more profitable, level of sales. Such a higher level of sales is only possible if the firms actually have produced more than the agreed upon level of sales. Thus if the duopoly has implicitly agreed to produce substantially more than what it intends to sell each firm has a lot of inventories with which cheating at the sales stage is possible. Cheating may also be desirable since it avoids the inventory carrying costs.

In the examples we consider, an additional unit of inventories both makes cheating more attractive (by increasing the amount by which sales can exceed their collusive value), and raises the punishments in the second period. Moreover both the increased benefits from cheating and the increased punishments from an additional unit of inventory fall as inventories get large. However, the latter tend to fall faster than the former. Thus there tend to be two levels of inventories at which the desire to cheat at this stage is exactly offset by the punishments to deviators. This can be seen as follows.

When there are essentially no planned inventories, going along is obviously more desirable than cheating since firms have no goods to cheat with. Here, the main effect of an extra unit of inventory is to make cheating more desirable. As the level of inventories increases the two incentives first balance and then cheating becomes the more attractive option. However, as the level of inventories grows further, the main effect of an additional unit of inventory is to increase the punishment. This eventually leads to a high level of inventories at which the two incentives balance once again. From that point on cheating is again less desirable than going along.

These complications might well lead one to ask why, given that having inventories around often makes collusion at the sales stage more delicate, the oligopoly should build them in the first place? The answer lies in the ability of firms to cheat at the production stage. Suppose the oligopoly plans to build no inventories. Then a prospective deviator knows that its rivals have only limited ability to punish it if it produces a lot and captures a significant fraction of the first period's market. This will tempt them to deviate, particularly if future punishments are limited. Such deviations will result in the rivals being worse off than if they had built inventories to enforce cooperation. Thus, sometimes, situations with relatively low planned inventories cannot be equilibria.

An important consideration when discussing deviations at the production stage is the observability of production. Since inventories are built for their deterrent effect, firms must be capable of demonstrating that they have them. Moreover, failure to carry out this demonstration leads the rival to believe that cheating is better than cooperating. Thus, underproduction, which makes this demonstration impossible, is always detected. On the other hand, it might or might not be possible to conceal overproduction. If it is not, any slight overproduction generates the expectation that noncooperation will ensue at the sales stage of the first period. This makes overproduction unattractive even when there are few planned inventories. On the other hand, if a firm is able to successfully hide its overproduction, the other firm will still sell its planned (collusive) level of sales at the sales stage. This makes deviations at the production stage more appealing and eliminates candidates for equilibria with few inventories. Therefore, observability can improve the equilibrium from the point of view of the oligopoly.

So far we have only argued that there exist equilibria in which the oligopoly can sustain superior outcomes by planning to build inventories. We also show that under some circumstances the level of inventories the oligopoly must build increases when demand increases.

What light can this model shed on the puzzling empirical regularities with which we began? First, the fact that inventories rise when demand (or sales) rise rationalizes the correlations reported in Table 1. Second, consider the often quoted fact that the variance of production exceeds that of sales. Suppose that demand is the same in both periods of a model. Then, sales in the two periods will either be identical or very similar. Thus the fact that inventories are built up in the first period and run down in the second period means that the variance of production exceeds that of sales.

The paper proceeds as follows. In the next section we develop the notation and discuss some modelling issues. Then, in later sections we consider the equilibria that emerge from two different cost structures for the firms. In the first, which is the focus of Section III, there are constant marginal costs but there is a capacity constraint. Producing above this constraint is possible but costly. In Section IV we assume instead that marginal costs increase linearly with output. This case is more realistic than the previous one but this realism is achieved at a price. We are able to derive fewer closed form expressions and must rely instead on numerical simulations. Conclusions and possible extensions are presented in Section V.

II. The Formulation

This section considers specific assumptions which we make in the rest of the paper. These are of two types. In the first subsection we present our demand and cost functions together with the necessary notation. In the second subsection we overview various alternative ways (including the one we eventually focus on) in which the duopoly can collude in the presence of inventories.

a) Notation

We assume that the demand for the industry's output is linear in every period. In particular letting P be the price in a period and S the industry level of sales, the demand in the first period of a model is:

$$P = a - bS.$$

In the second period we replace a by $\bar{a}$ in this expression to allow demand to vary between periods. This distinction is only exploited in Section IV.

Production is carried out at the beginning of a period. At the beginning of the second period firm i produces X_i. As long as prices for old models are low, firm i will choose to sell $X_i + \tilde{I}_i$ in the second period where $\tilde{I}_i$ are the inventories carried by firm i between periods.

The duopoly tries to sustain a collusive outcome by letting each firm carry over some inventories. The level of inventories firm i is intended to carry over is I_i. It is potentially different from $\tilde{I}_i$ since firm i might cheat and sell more than the intended amount. Thus, if firm i goes along at the production stage with the duopoly's plan it will produce I_i plus the intended level of sales. At first we assume that this intended level of sales is given by the static profit maximizing level, S^m, although we later relax this constraint and allow the duopoly to choose S_i optimally.

When firms cheat at the production stage in the first period, they consider producing an amount different from the intended sales plus I_i. We refer to this as Q_i.

We consider two cost functions for the members of the duopoly. The first is somewhat artificial but generates analytically tractable equilibria. Here, we assume that marginal cost is always equal to zero. However, there is a capacity constraint. This can be circumvented by the payment of a fixed cost, f. Thus

$$c(Q) = 0 \qquad Q < \bar{Q}$$
$$c(Q) = f \qquad Q > \bar{Q} \qquad (1)$$

where c(Q) is the cost of producing Q. This captures in a simple way the increase in costs that follows from increasing production.[7] The second cost function we consider is more standard but leads to equilibria which can only be characterized numerically. In particular we let:

$$c(Q) = cQ^2/2. \qquad (2)$$

Thus marginal cost is increasing throughout. Firms are assumed to have a discount factor of δ and the cost of carrying I units of inventory is g+dI per period. Thus, inventory carrying costs have a fixed component g and a constant marginal cost d. The fixed component only plays a role in Section IV and, without loss of generality, we set g=0 in Section III.

b) Modelling Issues

We now turn to some of the issues which arise when modeling this type of cooperation in the presence of inventories. To some extent these issues reflect choices available to the duopoly. Some of these choices lead to outcomes superior to those of others. However, these choices also have different degrees of plausibility.

First, the duopoly has to decide on its desired level of sales. A natural candidate is the usual (zero inventory) joint profit maximizing level. However, in order to support that level, the firms will have to carry inventories. But once one includes the possibility of carrying inventory it

[7]We can think of f as being the fixed costs associated with employing an additional plant or of rearranging production so as to increase effective capacity.

is no longer optimal to support the zero inventory collusive level of sales. This is the case for two reasons: First, since production and, therefore, marginal costs, are bigger, a larger marginal revenue and thus lower sales may be desirable. Second, a larger level of sales, by itself, promotes collusion and thus may permit smaller expenditures on inventories. We will consider both the optimal level of sales in the presence of inventories and the zero inventory joint profit maximizing level. Although the latter, which is the exclusive focus of Section III, isn't optimal it permits analytic solutions to the duopoly's problem whereas the former does not. Moreover, when we consider both in Section IV, they yield nearly identical levels of sales.

The oligopoly also has to choose whether it will penalize only deviations from the agreed level of sales or also deviations from the agreed level of production. There are three possibile formulations to consider.

In the first formulation the duopoly penalizes any upward or downward deviation from either the agreed level of production or the agreed level of sales. Of course this assumes that the amount produced is common knowledge. Since this imposes a lot of discipline it also leads to the most desirable outcomes for the duopoly. In equilibrium the level of inventories is such that neither firm finds it profitable to overproduce in order to sell a lot in the first period.

The firm is deterred from so deviating by the inventory enhanced punishment its rival will mete out in the second period. This equilibrium has so few inventories however, that, at the sales stage neither firm would deviate from the agreed level of sales if the other held slightly fewer inventories. The strategies specify these "excessive" inventories only to deter cheating at the production stage. Thus, each firm has an incentive to reduce its inventories, since once the sales stage is reached they are not all necessary to deter the other firm from cheating. What prevents a firm from doing this is the threat that underproduction will be penalized. Even though this threat is credible this seems implausible. It is like penalizing a firm for behaving too cooperatively.[8]

In the second possible formulation, the duopoly penalizes only deviations from some predetermined level of sales. This seems plausible since sales are the only thing the oligopoly ultimately cares about. Thus firms are free to produce as much as they wish. However, if they produce in excess of the implicitly agreed upon level (and are incapable of hiding the extra production), their rival, anticipating large sales, might also sell off its planned inventory.

Finally, one could consider a formulation in which the level of sales that the duopoly attempts to sustain is contingent on the actual quantities that the duopolists produce. The implicit agreement thus specifies sales as a function of production. This makes the duopoly very flexible, particularly with respect to any random productivity shocks. Naturally this formulation leads to less collusion as each firm knows that if it produces slightly more, the implicit agreement will accommodate it selling slightly more.

Our analysis focuses on the second formulation which appears somewhat more plausible.

III. The Model with Zero Marginal Costs

In this section we analyze the cost function given in (1) above. We assume that for all possible values of a, $S^m < \bar{Q} < S^c$, where $S^m = \frac{a}{4b}$ is one half of the joint profit-maximizing level of sales, and $S^c = \frac{a}{3b}$ is the firm's Cournot output. Since the firms are capable of producing their joint profit-

[8] To use an analogy to which we shall return, it is like starting a war to penalize the other country for producing too few weapons.

maximizing levels of output in every period, there would be no role for inventories if the firms were able to enforce cooperation.

The sine qua non of our model is that it be impossible for the duopoly to sustain the single-period joint profit-maximizing output in the first period without holding inventories. For this to be so, it must be the case that one of the firms would find it profitable to deviate by overproducing. We thus require that:

$$(a - b\bar{Q} - bS^m)\,\bar{Q} + \delta\,(a - 2b\bar{Q})\,\bar{Q} - K > (1+\delta)(a - 2bS^m)S^m, \text{ or}$$
$$(a - bS^b(S^m) - bS^m)\,S^b(S^m) - f + \delta(a-2b\bar{Q})\bar{Q} - K > (1+\delta)(a-2bS^m)S^m, \tag{3}$$

where $S^b(S^m) \equiv (a-bS^m)/2b$ is a firm's best-response to the sales of S^m by the other. The first equation applies when the deviating firm produces to its capacity when it deviates, while the second applies when the firm exceeds its capacity when it deviates. Both assume that neither firm exceeds its capacity in the noncooperative game that ensues in the second period. (A sufficient condition for this to be optimal is provided below).

The strategic role of inventories is readily seen by examining the consequences of a deviation by one firm at the sales stage in the first period. In that event a noncooperative game ensues in the second period. We now study that noncooperative game.

a) The Noncooperative Game in the Second Period

Without inventories, in the absence of the capacity constraint, each firm would produce and sell S^c. However to do that here the firm must incur the cost f. For a given output, X_j, of its rival, firm i is only willing to incur that cost if:

$$(a - bX_j - bX_i^b(X_j))\,X_i^b\,(X_j) - f > (a - bX_j - b\bar{Q})\,\bar{Q},$$

where $X_i^b(X_j) \equiv (a - bX_j)/2b$. If f is "large enough" firm i prefers instead to produce $\bar{Q}$; that is when period 2 is reached, firm i doesn't find it optimal to exceed its capacity. (In game-theoretic terms, it may not be sequentially rational for firm i to exceed its capacity constraint). This correspondingly limits the effective punishment in the second period since firm j's second-period profits if it deviates in the first period will be no lower than $(a-bX_j-b\bar{Q})X_j$. If this punishment is "too low", the firms may be unable to sustain high profits in the first period because the incentive to unilaterally deviate is too high. Suppose, instead, that firm i has inventories of $\tilde{I}_i$ when it reaches the second period. If, in addition, it produces to capacity in the second period, second period profits for firm j are $(a-bX_j-b(\bar{Q} + \tilde{I}_i))\, X_j$. Since this expression is strictly decreasing in $\tilde{I}_i$, the punishment firm j faces if it deviates is increasing in its rival's inventories (providing it is optimal for firm i to produce and sell $\bar{Q} + \tilde{I}_i$ if a deviation occurs).

We turn now to characterizing equilibria in which inventories serve a disciplining role. In these equilibria the duopoly plans a level of production for each firm for the first period. Some of this production is intended to be sold in the first period. The rest is earmarked to be carried as inventory to provide sufficient punishment capability to prevent either of the firms from deviating at the sales stage.

In this section we assume that the targeted sales level in the first period is the standard zero-inventory joint profit-maximizing level, S^m. Thus, the problem for the duopoly reduces to finding a pair of planned inventories for the firms, $\{I_1^*, I_2^*\}$, such that neither firm wants to deviate either at the production stage or at the sales stage.

b) The Sales Stage in the First Period

Consider the sales stage and suppose that the firms have produced $\{S^m + I_1, S^m + I_2\}$ so that I_i is the intended level of inventories. If neither firm deviates, each firm earns its share of the joint profit maximizing level:

$$\frac{a^2}{8b}(1+\delta) - d\ I_i. \tag{4}$$

Assuming for the moment that a firm that cheats at the sales stage sells its entire output, the deviating firm, (say firm 1) in the first period earns:

$$(a-b(S^m+I_1) - bS^m)(S^m+I_1). \tag{5}$$

In the second period the noncooperative game ensues.

If f is sufficiently high, firm 1 (which sells its entire stock-on-hand in the first period) sells $\bar{Q}$ in the second period. Firm 2 sells $\min\{\bar{Q}+I_2,\ X_2^b(\bar{Q})\}$ i.e. firm 2 sells the lesser of its best-response to $\bar{Q}$ and $\bar{Q}+I_2$ (which is the maximum it can sell in the second period given that it has brought in inventories of I_2). We assume for the most part that $\bar{Q}+I_2 < X_2^b(\bar{Q})$.

Then firm 1's discounted second-period profits are:

$$\delta(a-b(\bar{Q} + I_2) - b\bar{Q})\bar{Q}. \tag{6}$$

Combining equations (4) - (6), firm 1 chooses not to deviate if

$$\Delta(I_1, I_2) \equiv (a-b(S^m + I_1) - bS^m)(S^m + I_1) + \delta(a - b(\bar{Q} + I_2)-b\bar{Q})\bar{Q} - K$$
$$-\frac{a^2}{8b}(1+\delta) + dI_1 < 0. \tag{7}$$

The function $\Delta(I_1, I_2)$ is the difference between the profits from deviating and those from participating in the collusive understanding. In other words, it is the net benefit to firm 1 from deviating if the firms are holding inventories of I_1 and I_2.

Equation (7) highlights another effect of inventories, namely that, although high inventories may deter the firm's rival from deviating, they increase the firms own incentive to deviate. This occurs for two reasons: first if the firm deviates it does not have to incur the costs of carrying the inventory, dI_i and second, at least for moderate levels of I_i, $(a-b(S^m+ I_i)-bS^m)(S^m+I_i)$ is increasing in I_i i.e. the firm makes larger first-period profits when it does deviate.

Substituting $S^m = \frac{a}{4b}$ in (7) and rearranging yields the result that I_2 just deters firm 1 from deviating when it has inventories of I_1 if[9]

$$I_2 = \frac{1}{\delta b\bar{Q}} \{I_1(-bI_1+ \frac{a}{4} + d) + M(a)\}, \tag{8}$$

where $M(a) \equiv \frac{-\delta a^2}{8b} + \delta a\bar{Q} - 2b\delta\bar{Q}^2 - K < 0,$ (9)

since $\bar{Q} > Q^m$.

The symmetric Nash equilibria, if they exist, have $I_1 = I_2$. Of particular interest are those symmetric equilibria in which the inventories are such that deviations are just deterred. Setting $I_1 = I_2 = I^*$ in (8) and solving gives:

$$I^* = \left[(\tfrac{a}{4} + d - \delta b\bar{Q}) \pm \sqrt{(\tfrac{a}{4} + d - \delta b\bar{Q})^2 + 4bM(a)}\right]/2b. \tag{10}$$

Thus these symmetric equilibria come in pairs. We denote the one with the lower (higher) level of inventories by $I^{*\ell}$ (I^{*u}).

---

[9]This analysis implicitly assumes that I_2 is not so large that it implies optimal output of less than $\bar{Q}$ in the second period in the noncooperative game there.

The functions $I_2(I_1)$ and $I_1(I_2)$ are illustrated in Figure 1. $I^{*\ell}$ and I^{*u} are given by the intersections of these curves. The points $\{I_i', I_2'\}$ for which $I_1 > I_1(I_2')$ and $I_2' > I_2(I_1')$ are also equilibria at the sales stage. These points have the feature that each firm has enough inventories to deter the other from deviating. But notice that at all these points with the exception of $I^{*\ell}$, I^{*u} and the origin, at least one of the firms is holding more inventories than it needs to deter the other from cheating at the sales stage. No firm would be willing to produce such excessive inventories unless it believed its rival would punish it for underproducing.[10] If we rule out such implausible beliefs the only candidates for equilibria that remain are the origin, $I^{*\ell}$ and I^{*u}. However if equation (3) holds, which is the premise of this paper, there is no equilibrium with zero inventories. That leaves I^{*u} and $I^{*\ell}$.

The intuition behind the existence of these two equilibria is as follows. Suppose that both firms change their level of inventories together and consider firm 1's incentive to deviate. This is given by $\frac{\partial \Delta}{\partial I} = (\frac{a}{4} + d - \delta b Q) - 2bI$. This is positive if $I < \frac{1}{2b}\{\frac{a}{4} + d - \delta b Q\}$. Thus the change in the incentive to deviate as the firms' level of inventories changes depends on the level of inventories from which the change is made. When inventories are low and both firms increase their inventories slightly, the deterrent effect of the rival's extra inventories is outweighted by the increase in one's own temptation to deviate. Thus $I_2(I_1)$ increases more rapidly than the 45° line at low values of I_1. However as I_1 increases, the relative strength of the two effects is reversed and $I_2(I_1)$ bends back

[10] This is equivalent to waging war against a country for having too small an arms build-up.

towards the 45° line.

Eventually $I_2(I_1)$ peaks and then begins to fall. This means that as firm 1's inventories increase beyond this point firm 2 requires even fewer inventories to deter it from cheating. This is purely an artifact of the assumption that the deviating firm sells its entire inventory when it cheats. Under this assumption a firm's own very large inventories decrease its profitability from deviating.[11] We provide a sufficient condition for a deviating firm to be willing to sell its entire production below.

At $I^{*\ell}$ firm 1 has a very small incentive to deviate both because it cannot increase sales by much (it is constrained to sell less than S^m+I^*) and because its low level of inventories means it doesn't save much in inventory carrying costs by deviating. Thus firm 2 also needs a relatively low level of inventories to prevent the deviation. At I^{*u} firm 1 has a large incentive to deviate and firm 2 requires a large level of inventories to deter it from doing so.

It remains to provide conditions for which equilibria exist i.e. for which $I_2\ (I_1)$ is as depicted in Figure 1. Firstly, $M(a) < 0$ implies that $I_2(0) < 0$ (I_2 starts below the x-axis). Second, $\frac{dI_2}{dI_1}\Big|_{I_1=0} =$ $\frac{1}{\delta b\bar{Q}}\ (\frac{a}{4} + d) > 0$ (I_2 starts with a positive slope). Third, $I_2\ (I_1)$ and the 45° line intersect if $\sqrt{(\frac{a}{4} + d - \delta b\bar{Q})^2 + 4b\ M(a)}$ has real roots i.e. if

$$(\frac{a}{4} + d - \delta b\bar{Q}\ )^2 > 4b\delta\{\frac{a^2}{8b} + K/\delta - (a-2b\bar{Q})\bar{Q}\}. \qquad (11)$$

Thus equation (11) is a sufficient condition for the existence of an

---

[11]Eventually, I_1 is so large that $I_2(I_1) = 0$. Here firm 1 has so many inventories that it would be "suicide" to use them all when it deviates even if firm 2 has none. This is analogous to a superpower using an arsenal of nuclear weapons sufficient to produce "Nuclear Winter."

equilibrium. It is more likely to be satisfied if δ and K are low and if d is high. If δ and K are sufficiently high then for any level of inventories that firm 1 has, firm 2 is able to deter it from cheating with even fewer inventories and no equilibrium with inventories exists. On the other hand if d is high, firm 1 has a large incentive to deviate in order to save on the inventory carrying costs and hence firm 2 requires a large inventory to deter it. Thus when d is high it is more likely that the $I_2(I_1)$ curve rises above the 45° line.

Of particular interest is how equilibrium levels of inventories vary with the level of demand. Since $\frac{dM(a)}{da} = \delta(\bar{Q} - \frac{a}{4b}) = \delta\,(\bar{Q} - S^m) > 0$, it is immediate from (10) that $\frac{dI^{*u}}{da} > 0$.

Now consider $I^{*\ell}$ and let $y = \sqrt{(\frac{a}{4} + d - \delta b\bar{Q})^2 + 4bM(a)}$. Then $\frac{dI^{*\ell}}{da}$ $= \frac{1}{2b}\left\{\frac{1}{4} - \frac{(\frac{a}{4} + d - \delta b\bar{Q})}{4y} - \frac{4b(\bar{Q} - \frac{a}{4b})\delta}{4y}\right\}$. Since $(\frac{a}{4} + d - \delta b\bar{Q}) > y$ and $\bar{Q} > S^m = \frac{a}{4b}$ we have that $\frac{dI^{*\ell}}{da} < 0$. At the low inventory equilibrium an increase in demand results in a decrease in equilibrium inventories.

As before, the intuition for this can be gained by examining equation (7). Evaluating $\Delta(I_1, I_2)$ at $I_1 = I_2 = I^*$ and totally differentiating gives:

$$\frac{dI^*}{da} = \frac{-\partial\Delta(I^*)}{\partial a} \Big/ \frac{\partial\Delta(I^*)}{\partial I^*}. \tag{12}$$

Now $\frac{\partial\Delta}{\partial a} = \frac{I_1}{4} + \delta(\bar{Q} - S^m) > 0$ since $\bar{Q} > S^m$ i.e. the incentive to deviate is increasing in "a", holding everything else constant. Thus $\frac{dI^*}{da}$ has the sign of $\partial\Delta(I^*)/\partial I^*$. As discussed above, this is positive (negative) if I^* is relatively small (large).

Thus an increase in demand can lead to an increase or decrease in equilibrium inventories. However, there are two reasons why the former is the more relevant case. First, the incentive to deviate at the production stage often eliminates the lower equilibrium while leaving the upper equilibrium intact. Second, in the more realistic case where marginal costs are increasing studied in the the next section we find that equilibrium inventories sometimes increase with an increase in demand at both equilibria.

In the above analysis we made two simplifying assumptions for which we now provide sufficient conditions:

(i) A firm that deviates at the sales stage was assumed to sell its entire first-period production, $S^m + I_1*$, in the first period. The total revenue in the first period when the firm sells $S^m + I_i^*$, is TR = $(a-b(S^m + I_i^*)- bS^m)(S^m + I_i^*)$ and hence the marginal valuation of the last unit is $(a-2b(S^m+I_i^*)-bS^m)$. On the other hand, in the second period, a marginal unit is worth $a - 3b\bar{Q} - bI^*_2$ (since total revenue there is $(a-b(\bar{Q} + I_2^*)-b\bar{Q})\bar{Q})$. Thus firm 1 will choose to sell all of $S^m + I_i^*$ in period 1 if $a-2b(S^m+ I_i^*) -bS^m > \delta(a-3b\bar{Q} - bI_2^*) -d$. A sufficient condition for this is $\bar{Q} > S^m/\delta$.

ii) It was assumed that if firm 1 deviated that firm 2 would find it unprofitable to exceed capacity in period 2. If it does not exceed capacity firm 2 earns at least $(a-2b\bar{Q})\bar{Q}$. If it exceeds capacity it earns $(a-b\bar{Q} -bX_2^b(\bar{Q}))X_2^b(\bar{Q}) -f$. A sufficient condition for it to be unprofitable to exceed capacity is thus $f > \frac{1}{4b}\left[a^2-2b\bar{Q}(1+2a) + 9b^2\bar{Q}^{\,2}\right]$. We provide an example below in which equilibria exist and in which these conditions are satisfied.

c) The Production Stage in the First Period

We turn now to an analysis of the decision at the production stage. We assume that the amounts produced become common knowledge. This means not only that a firm is able to display its inventories to induce its rival to cooperate; but also that neither firm is able to secretly build up inventories to unleash on the opponent at the sales stage.[12]

Where production is common knowledge, a deviation at the production stage from $S^m + I^*$ launches the firms into a noncooperative game at the sales stage. Suppose that firm 2 has produced $Q_2 = S^m + I^*$ as planned but that firm 1 has produced $Q_1 \neq S^m + I^*$. In the noncooperative game that ensues firm 1 chooses how much of the Q_1 on hand to carry over to period 2 as inventory.[13]

That is, it chooses $\tilde{I}_1$ to

$$\max_{\tilde{I}_1} \; (a-b(Q_1-\tilde{I}_1)-b(Q_2-\tilde{I}_2))(Q_1-\tilde{I}_1) + \delta A(\bar{Q}+\tilde{I}_1)-d\tilde{I}_1$$

where $A \equiv a-b(\bar{Q}+\tilde{I}_1)-b(\bar{Q}+\tilde{I}_2)$.

There is a similar expression for firm 2.

The first-order conditions for $\tilde{I}_1$ and $\tilde{I}_2$ respectively are:

$$a-2b(Q_1-\tilde{I}_1)-b(Q_2-\tilde{I}_2) - \delta A + \delta b(\bar{Q}+\tilde{I}_1) + d = 0 \qquad (13)$$

$$a-b(Q_1-\tilde{I}_1) - 2b(Q_2-\tilde{I}_2) - \delta A + \delta b(\bar{Q}+\tilde{I}_2) + d = 0. \qquad (14)$$

Solving (13) and (14) simultaneously yields:

[12] In the armaments analogy the first of these underlies demonstrations of preparedness like the Soviet May Day Parade, while "mutual verifiability" seeks to achieve the second.

[13] Equivalently, firm 1 chooses its level of sales in period 1 which determines its inventory at the beginning of period 2.

$$\tilde{I}_i^* = \frac{3b(Q_i - \delta\bar{Q}) - a(1-\delta) - d}{3b(1+\delta)}.$$

The above analysis has assumed an interior solution for $\tilde{I}_i^*$ i.e. $\tilde{I}_i^* > 0$. However, if d is large enough, the firm sells its entire output in period 1. Thus, in fact,

$$\tilde{I}_i^* = \max\ \{0, [3b(Q_i - \delta\bar{Q}) - a(1-\delta) - d]/3b(1+\delta)\}. \tag{15}$$

Using (15) it is straightforward in numerical examples to calculate the profits from the noncooperative game for any given deviation in production. It is then possible to check numerically for the existence of profitable deviations in production.

We calculate the equilibria, and check the sufficient conditions and that no deviations are profitable, for the case where $P = 100 - 0.25S$, $\delta = 0.6$, $d=3$, $K = 20$ and $\bar{Q} = 117\ 2/3$. The equilibria have $I^{*u} = 23.14$ (with $Q_1 = 123.14 > \bar{Q}$) and $I^{*\ell} = 18.26$ (with $Q_1 = 118.26 > \bar{Q}$). The sufficient condition on f is $f > 138.06$.

IV. The Model with Increasing Marginal Costs

In this section we carry out the analysis with the cost function given by (2). We show that, here too, there exist situations where the collusive level of sales cannot be sustained without inventories. The inventories serve to ensure that there is a sufficient fear of future punishments to keep the duopolists in line. Moreover, we show that increases in demand continue to tend to increase the need for credible punishments like those provided by inventories.

If demand is the same in the two periods there is now an even greater disincentive to accumulate inventories. Not only must inventory carrying costs be paid but costs are made larger by the increasing marginal costs of production and are paid earlier. Nonetheless it is still the case that firms are willing to accumulate inventories in the first period to have a credible punishment which ensures cooperation in that period. To show this we follow the organization of the previous section. We start with the noncooperative game in the second period and show how larger inventories in the opponents hands do reduce one's own profits. Next we study the sales stage and finally the production stage of the first period. After that we turn to some comparative statics and to a discussion of optimal sales.

a) The noncooperative subgame in the second period

Assuming nothing gets sold between models firm i maximizes with respect to its own production Q_i:

$$(\bar{a}-b(X_i+X_j+\tilde{I}_i+\tilde{I}_j))(X_i+\tilde{I}_i) -c/2(X_i)^2 \tag{16}$$

taking the production of firm j and both inventories as given. Here $\bar{a}$ denotes the level of a in the second period. It therefore produces:

$$X_i = (\bar{a}-b(X_j+\tilde{I}_j)-2b\tilde{I}_i)/(2b+c). \tag{17}$$

A similar expression obtains for firm j. Adding together both of these expressions and using the demand curve one obtains the following equilibrium price:

$$P_c = \bar{a}(b+c)/(3b+c) - bc(\tilde{I}_i+\tilde{I}_j)/(3b+c) \tag{18}$$

Therefore firm i's second period profits when the firms behave noncooperatively are:

$$\Pi_{ic}(\tilde{I}_i,\tilde{I}_j) = \frac{\bar{a}^2(b+c/2)}{(3b+c)^2} + \frac{\bar{a}c(2b+c)^2\tilde{I}_i - \bar{a}bc(2b+c)\tilde{I}_j}{(3b+c)^2(b+c)}$$
$$- \frac{bc(9b^3/2+8b^2c+5bc^2+c^3)\tilde{I}_i^{\,2} - (bc)^2(b+c/2)\tilde{I}_j^{\,2} + bc^2(2b+c)^2\tilde{I}_i\tilde{I}_j}{(3b+c)^2(b+c)^2}. \tag{19}$$

Starting from zero inventories, increases in $\tilde{I}_i$ raise profits while increases in $\tilde{I}_j$, by increasing firm j's sales, reduce them. This is what makes inventories a credible deterrent. As inventories rise, however, the marginal reduction in profits from increases in $\tilde{I}_j$ falls. This occurs because, as $\tilde{I}_j$ rises firm i sells less and less. Increases in I_i continue to (linearly) lower prices but the losses from lower prices are incurred over fewer units. Similarly, increases in $\tilde{I}_i$ have lower positive effects on profits as $\tilde{I}_i$ rises because each increase in $\tilde{I}_i$ lowers prices in equilibrium. These lower prices have a more pronounced effect on firm i's profits when firm i sells a lot.

b) <u>The sales stage in the first period</u>

We now consider optimal behavior at the sales stage in the first period. At this point firm i has already produced the monopoly level of sales S^m (given by $a/(4b+c)$) plus I_i the intended equilibrium level of inventories. It has two choices. First it can sell S^m. Under our assumptions this ensures that the collusive outcome will also prevail in the second period. Thus the present value of profits at the sales stage from following this strategy is:

$$\Pi^*_s\,(I_i) = \frac{a^2(2b+c) + \delta\bar{a}^2(2b+c/2)}{(4b+c)^2} - (d- \frac{\delta c\bar{a}}{4b+c})I_i - \delta c/2I_i{}^2 - g. \qquad (20)$$

The alternative is to sell an amount different from S^m. Selling less than S^m produces fewer profits even in the first period. Instead, selling more than S^m increases first period profits albeit at the expense of noncooperative behavior in the second period. The value of doing this depends in part on whether the firm that does this sells all the goods it has on hand or whether it keeps some of them as inventories. As long as g, the fixed costs of carrying over any inventories at all, is large enough, a defecting firm always chooses to sell all the output it has on hand. We first assume that this is the case and return at the end of the subsection to the conditions that ensure it. In this case, if firm i defects its profits from then on are:

$$(a-2bS^m-bI_i)(S^m+I_i) + r\Pi_{ic}(0,I_j). \tag{21}$$

We now turn to an analysis of the conditions under which, at the sales stage, firms are deterred from cheating by the inventories the other firm is carrying over. To analyze these conditions we focus on Δ, the difference between the profits from deviating by selling (S^m+I_i) and the profits from going along given in (20). As before Δ is the net incentive to deviate:

$$\Delta = g-K-\delta\bar{a}^2\left[\frac{2b+c/2}{(4b+c)^2} - \frac{b+c/2}{(3b+c)^2}\right] + \left[\frac{a(b+c)-\delta c\bar{a}}{4b+c} + d\right]I_i - [b-\delta c/2]I_i^2$$
$$- \frac{\delta bc(b+c/2)\ [2\bar{a}(b+c)I_j - bcI_j^2\ ]}{(3b+c)^2(b+c)^2}\ .$$

By deviating, the firm avoids the inventory carrying cost g but incurs the punishment K. The next term in (22) gives the second period loss in the excess of collusive over noncooperative profits when inventories are absent. Increases in I_i reduce the incentive to cooperate, at least for low values of I_i while increases in I_j tend to increase it. Equation (22), when equated to zero can be thought of as giving the minimum level of I_j which deters firm i from cheating. This is the lowest value of I_j which makes Δ zero for a given

I_i. The plot of this minimum I_j as a function of I_i is given in Figure 2 for the parameter combination given by:

$$a=\bar{a}=2500 \qquad b=75 \qquad c=1 \qquad d=.2 \qquad \delta=.945 \qquad g=.23 \qquad K=195.23 \qquad (23)$$

For this example S^m is given by 8.31. The minimum I_j which deters cheating is very small for low values of I_i and then rises steeply only to decline again.[14] As before it crosses the 45° line at two points $I^{*\ell}$ (4.08) and I^{*u} (4.20). These points can be obtained as solutions to $\Delta=0$ for a common I (equal to both I_i and I_j). These points are symmetric with respect to the point at which, for a common I, the derivative of Δ with respect to the common I is equal to zero. As long as $b-rc/2-r(bc)^2\,(b+c/2)/[(3b+c)^2\,(b+c)^2]$ is positive, this second deriviative is negative. Then, just as in the previous section $I^{*\ell}$ has a positive derivative of Δ with respect to I while I^{*u} has a negative one. An increase in production by both firms at the lower point raises Δ thus precipitating competition while the same increase at the upper equilibrium maintains collusion. This can be seen directly by inspecting Figure 2. Below $I^{*\ell}$ any common I deters cheating since less than the common I is needed to stop the other firm from deivating. Between $I^{*\ell}$ and I^{*u} more than the common I is needed once again. Thus at the sales stage all points in which both firms produce an excess over S^m below $I^{*\ell}$ and those in which the excess over S^m is above I^{*u} induce cooperation.

To conclude this subsection we study the conditions under which a firm which defects at the sales stage chooses to carry no inventories. To do this we must first analyze the equilibrium in the case in which it does choose to

---

[14] In this section, a firm that deviates may be willing to sell its entire first period production even when the marginal revenue of the last unit is less than what it would be in the second period. This is because, in addition to the variable inventory carrying costs, it also saves the fixed costs, g.

carry inventories and then compare the resulting profits to the case in which it doesn't. If firm i chooses to carry inventories it will pick a level I_i to maximize:

$$(a - b(2S^m + I_i - \tilde{I}_i))(S^m + I_i - \tilde{I}_i) - d\tilde{I}_i + \delta\Pi_{ic}(\tilde{I}_i, I_j). \qquad (24)$$

This leads to inventories given by:

$$\tilde{I}_i = [\phi_0 + 2bI_i + \phi_2 I_j]/\phi_1 \qquad (25)$$

where:

$$\phi_0 = -a(b+c)/(4b+c) - d + \delta\bar{a}c(2b+c)^2/[(3b+c)^2(b+c)] \qquad (26)$$

$$\phi_1 = 2b + 2\delta bc(9b^3/2+8b^2c+5bc^2+c^3)/[(3b+c)^2(b+c)^2] \qquad (27)$$

$$\phi_2 = -\delta bc^2(2b+c)^2/[(3b+c)(b+c)]^2. \qquad (28)$$

Substituting this level of inventories in to (24) one obtains a level of profits which can be compared with the profits in (21). If this latter is bigger no inventories will be carried over.

c) The production stage in the first period

We now consider the production stage of the first period. Let the amount that firms are implicitly expected to produce be (S^m+I). We first assume that this amount is such that cheating at the sales stage is deterred. Thus if the firm goes along with this level of production it earns:

$$\Pi^*_p = \Pi^*_s(I) - c(S^m+I)^2/2. \qquad (29)$$

Instead, the firm can produce either less or more than S^m+I. For the moment we assume that underproduction, even though it is always detected, is not penalized. Thus, the only deterrent to underproduction is that it might lead the other firm to deviate at the sales stage. However, if firm i is expected to produce more than S^m, which is either strictly below $I^{*\ell}$ or strictly above I^{*u}, then firm j can be sure that firm i won't deviate from collusion at the sales stage even if firm j produces slightly less than firm i. Since production for inventory is costly, as before, such underproduction is worthwhile. Thus, situations in which output of either firm is either above $S^m + I^{*u}$ or strictly between S^m and $S^m+I^{*\ell}$ cannot be equilibria.[15] This leaves as candidate equilibria at the production stage only S^m, $S^m+I^{*\ell}$ and $S^m+ I^{*u}$.

We now consider the possibility that firm i produces an amount different from firm j, which is assumed, for the moment, to produce one of these three levels of output. As discussed in section II it makes a big difference whether overproduction by i can be hidden from j or not. We first assume that it can. Then i's first consideration is whether it wants to overproduce in order to sell a lot in the first period. If it will not be detected, it can safely assume that j will sell S^m.[16] Thus as long as i does not intend to carry over any inventories (which is assured for even moderate inventory carrying costs since the marginal revenue from one unit of a good in the second period is low under competition) it should produce the best response to S^m, or $(a-bS^m)/(2b+c)$. This leads to profits equal to:

$$(b+c)[(a-bS^m)/(2b+c)]^2+ \delta\Pi_{ic}(0, I_j). \qquad (30)$$

Clearly, these profits are decreasing in the production of firm j. In particular, for inventories to be held at all, (30) for I_j equal to zero must exceed (29) plus K. Otherwise, the duopoly can sustain collusion without any

---

[15] This is, in part a consequence of our exclusive focus on pure strategies. If mixed strategies were allowed it might be that equilibria existed in which output sometimes exceeded this level.

---

[16] This is guaranteed in the case in which firm j produces only S^m whether overproduction is observable or not.

costly inventories. Indeed in the example of (23) the expression in (29) is 20193 for I equal to zero while the expression in (30) minus K is 20203. Consider next the possibility that j produces $S^m + I^{*\ell}$. This is still a relatively samll number. Moreover, now overproduction with the intention of cheating in the first period also leads to a saving of g. For the example of (23) Π_p^* at $I^{*\ell}$ is 20174.2 while the expression in (30) is 20369.4. The latter barely exceeds the former by more than K-g (195) so that, if overproduction can be hidden, each firm knows that it will deviate from $S^m + I^{*\ell}$. Now consider the upper candidate equilibrium. It has an output of 13.26 which exceeds 12.43, the best response to S^m. Thus no firm would want to overproduce.[17]

This leaves the possibility of underproducing and allowing the other firm to deviate at the sales stage. This is worthwhile if the cost of carrying inventories is very high. We consider the effects of underproduction in a more general setting. Underproduction is always detected. Thus deviations resulting from underproduction are simply a special case of those deviations in production which lead to a breakdown of the collusive arrangement in the first period. These include overproduction in the case in which overproduction cannot be hidden from one's competitor. If firm i deviates in this manner, it knows firm j will act noncooperatively even in the sales stage of the first period.[18]

We now study this noncooperative subgame. In this subgame firms i and j

[17]We have considered a number of parametric simulations and, in all of them, $S^m + I^{*u}$ exceeds the best response to S^m while $S^m + I^{*\ell}$ falls short of it. However, we are unable to prove this result analytically.

[18]It must be noted, however, that this is not imperative. A slightly different (and more "cooperative") equilibrium would allow the firms to "modify" their behavior at the sales stage to sustain whatever profits they can, given the actual outputs that have been produced.

must decide upon their sales in the first period given productions of Q_i and Q_j respectively. What complicates this problem is that both firms must decide whether to carry over inventories at this stage or not. Thus comparisons of profits like those carried out earlier between (20) and (21) are necessary. We carry out only the analysis for the case in which both firms in fact carry over inventories. The other cases to which this one must be compared are analogous. Then, firm i, taking $\tilde{I}_j$ as given, maximizes with respect to $\tilde{I}_i$:

$$(a-b(Q_j-\tilde{I}_j+Q_i-\tilde{I}_i))(Q_i-\tilde{I}_i) - d\tilde{I}_i + \delta\Pi_{ic}(\tilde{I}_i, \tilde{I}_j) - c/2\, Q_i^2 \qquad (31)$$

Thus:

$$\tilde{I}_i = [2bQ_i+\phi_0-3bS^m+bQ_j+(\phi_2-b)\tilde{I}_j]/\phi_1 \qquad (32)$$

and, equivalently for firm j:

$$\tilde{I}_j = [2bQ_j+\phi_0-3bS^m+bQ_i+(\phi_2-b)\tilde{I}_i]/\phi_1. \qquad (33)$$

These two expressions can be solved for the equilibrium sales in the first period. If firm i deviates at the production stage and this leads to noncooperation at the sales stage it will choose a Q_i that maximizes (31) knowing that the inventories will be picked according to these formulae. The first order condition for this problem is:

$$a - (2b+c)Q_i + 2b\tilde{I}_i - b(Q_j - \tilde{I}_j) = 0. \qquad (34)$$

Which, using (32) and (33) yields the level of optimal production. This level of production can then be substituted back into (31) to obtain the profits from this deviation.

For the example of (23) these deviations are never worthwhile. In particular, if firm j produces $S^m + I^{*u}$, then firm i would produce 12.41 units if it decided to produce the optimal level of output consistent with noncooperation at the sales stage. This would lead to profits of 17965 which are substantially less than even the profits from going along with S^m+I^{*u} (21073). In this case it is optimal for both firms to carry over some inventories from period to period.

The level of production $S^m+I^{*\ell}$ is not an equilibrium if overproduction is not verifiable. Suppose instead that it is. Then, any production different from this level generates a noncooperative outcome at the sales stage. The optimal such deviation for firm i if firm j does indeed produce $S^m+I^{*\ell}$ (12.39) is to produce 12.41. Now, however, this leads only to profits of 17967 which are substantially below the profits from going along (20174). Therefore this lower level of production is indeed an equilibrium when overproduction is detectable.[19] Indeed it may be a more natural equilibrium in this case than S^m+I^{*u} since it has higher profits.

d) Comparative statics

We have now shown that the two solutions to $\Delta=0$ are equilibria; one when overproduction is detectable and the other when it isn't. Since the conditions that make these points equilibria are satisfied more than locally we can conduct local comparative statics by assuming that the new equilibria also solve this equation. We are particularly interested in the effect of increases in demand on equilibrium inventories. We consider two types of increases in demand. The first raises it in both periods (thus raising a and $\bar{a}$ simultaneously). The second raises first period a only.

Differentiating Δ with respect to a and $\bar{a}$ one obtains:

---

[19] In the context of this paper detectability is the same as verifiability. This suggests that extending this analysis to situations in which there is imperfect information may be necessary to understand the language of arms races.

$$d\Delta = \frac{(b+c)I}{4b+c}\,da - \left[\frac{2b+c/2}{(4b+c)^2} - \frac{b+c/2}{(3b+c)^2} + \frac{\delta cI}{4b+c} + \frac{\delta bc(2b+c)I}{(3b+c)^2(b+c)}\right] d\bar{a}. \quad (35)$$

An increase in a alone raise Δ. This is the result of Rotemberg and Saloner (1985) that, for linear demand, an outward shift in demand makes defections more attractive. It raises the price that a defector can get for the additional units that he sells when he defects. On the other hand an increase in $\bar{a}$ actually lowers Δ, the incentive to deviate. This occurs for two reasons. First, even in the absence of inventories it make the benefits from colluding in the second period higher relative to the profits from noncooperation. Second, it increases the damage caused by the other firms' inventories. On net the effect of an increase in both a and $\bar{a}$ is therefore ambiguous.

For the example of (23) both $d\Delta/da$ and the sum of $d\Delta/da$ and $d\Delta/d\bar{a}$ are positive at both equilibria. Since $d\Delta/dI$ is positive at the lower one and negative at the upper one, I increases at the upper equilibrium when demand increases in both periods while it falls at the lower one. An increase in the demand in both periods makes collusion less attractive at both equilibria. To induce more collusion inventories must fall at the lower equilibrium while they must rise at the upper equilibrium. However, we have also constructed numerical examples in which the lower equilibrium has a level of inventories that rises when a and $\bar{a}$ both rise. This occurs when δ is .9 and K is 200 with the other variables as given in (23). Then, an increase in the demand in both periods lowers the incentive to deviate at the lower equilibrium.[20]

e) Optimal sales

So far we have considered only duopolies who try to sustain sales of S^m in the first period. Since this is costly, it is natural for duopolies to also try to reduce costs by selling a different amount. However, whether more or less should be sold in the first period is in general uncertain. Selling more reduces the incentive to deviate as discussed in Rotemberg and Saloner (1985). However, since more than S^m is produced, marginal costs exceed S^m thus the duopoly might wish to obtain a higher marginal revenue. This would require a lower level of sales.

In the second period the duopoly is assumed to be able to sustain any level of sales. However, even then S^m is not generally optimal in the presence of inventories. These reduce marginal cost thus making it attractive to sell more. On the other hand such increases in second period sales, reduce the deterrent effect of inventories.

To optimize over sales we use an iterative technique. For each candidate pair of optimal sales S_1 and S_2 we compute the inventories I^* necessary to sustain these sales. This gives each firm profits of:

$$(a - 2bS_1)S_1 - c/2(S_1 + I^*(S_1, S_2))^2 - g - dI^*(S_1, S_2) + \delta((\bar{a} - 2bS_2)S_2 - c/2(S_2 - I^*(S_1, S_2))^2). \quad (36)$$

This expression can then be optimized over sales.

To obtain $I^*(S_1, S_2)$ we proceed as before. We compute the levels of inventories that make firms just indifferent, at the sales stage of the first

[20] This example however has no upper equilibrium. The higher solution to $\Delta=0$ has a level of inventories so high that firms would never sell them all when they cheat at the sales stage. Moreover, if they carry over inventories when they cheat at the sales stage, no level of inventories above $I^{*\lambda}$ balances the benefits from deviating with those going along. This illustrates the lack of robustness of our numerical examples in certain respects.

period, between selling their entire production and going along with the agreed upon level of sales. These levels of inventories satisfy the quadratic equation analogous to the one that equated Δ to zero:

$$- \left(b-\delta c/2- \frac{\delta(bc)^2(b+c/2)}{(3b+c)^2b+c)^2}\right) \quad I^2 + \left(a+d-3bS_2-\delta cS_1 - \frac{\delta\bar{a}bc(2b+c)}{(3b+c)^2(b+c)}\right) I +$$

$$g- K+ \frac{\delta\bar{a}^2(b+c2)}{(3b+c)} - \delta(\bar{a}S_2 - (2b+c/2)S_2^2) \; . \tag{37}$$

Again, there are two solutions to this equation. The higher one corresponds to unobservable overproduction while the lower one corresponds to observable overproduction.

In practice the solution to the numerical maximization of (36) using (37) is very similar to the inventories obtained when trying to sustain S^m. For the example in (23) the inventories are 3.15 when overproduction is observable and 4.16 when it is not. In the former case sales in the first period are 8.20 while they are 8.32 in the second period. This can be compared to S^m which equals 8.31. Thus the main effect on sales comes from changes in marginal cost brought about by the inventory accumulation.

The main difference between the case in which sales are fixed as S^m and the case in which they are not is that now carrying no inventories at all is more likely to be an equilibrium. This is particularly true in the case in which g, the fixed cost of carrying inventories is high. For the example in (27) raising S_1 to 8.32 eliminates the incentive to cheat at the production stage even when planned inventories are zero. This leads to profits of 20193 which exceed even the profits obtained from the low inventories equilibrium (20180).

To ensure that inventories are held even when sales are allowed to be different from S^m the incentive to cheat must be bigger. Since K is already very small in our example of (23) we change the parameters to $a = 2500$, $\bar{a} = 1000$, $b = 75$, $c = 1$, $d = 0.2$, $\delta = 0.9$, $K = 200$ and $g = 0$. The lower δ and $\bar{a}$ has the effect of reducing the losses from not cooperating in the second period.[21] Then optimal sales in the first period are 8.24 when we choose the equilibrium with observable overproduction. This equilibrium has inventories of 0.63 and profits of 11873. Instead, sales in the first period must be raised to 9.59 lowering profits to 11627 to sustain an equilibrium with no inventories. It must be noted that in this example inventories are also 0.63 when overproduction is observed and the duopoly sells S^m. Moreover, when a and $\bar{a}$ rise, equilibrium inventories rise whether sales are chosen optimally or not.

V. Conclusions

We have presented two rather special examples in which inventories are accumulated exclusively to deter deviations from an implicitly collusive arrangement. They are special in that they are based on explicit functional forms, sometimes even on explicit parameterizations, and in that we rely strongly on the presence of differing "models" to simplify the analysis. Nonetheless we feel that the main insight of these examples would carry through to both more general and more standard formulations.

The next step is to consider a nondurable good that is not subject to model years. This is important because the paradox that the variance of production exceeds that of sales is found in seasonally adjusted data as well. For products whose model year is a divisor of the calendar year seasonal adjustment would eliminate the variations in production that our

[21] An additional advantage of considering a low value for $\bar{a}$ is that this makes it more plausible that the collusive outcome can be sustained in the second period. This occurs because cheating is not very attractive when demand is low.

theory predicts in the case of constant demand. A theory without model years would have to consider the possibility that inventories are kept for a very long time. Its solution, for a stationary process of demand, would probably be a stationary inventory policy that makes inventories a function of current demand and lagged inventories. What makes this inventory policy substantially more complicated than the one that is obtained for competitive industries is that the oligopolistic industries we consider maximize profits subject to a much more intricate set of constraints. These require that firms find deviations unpalatable in every conceivable state of nature.

Consider for example the following two complications. First, these constraints may well require a high level of inventories when demand is high in order to deter deviations. However, production smoothing considerations may induce the oligopoly to stockpile inventories even when demand is low. So, for any level of demand below the highest level of demand, optimal inventories may exceed the level required to enforce cooperation. Second, the oligopoly faces difficulties when inventories have been stockpiled and current demand turns out to be unusually low. In that case, the incentive to deviate induced by these inventories may be significant.

Nonetheless, it is plausible that an increase in demand which raises the incentive to deviate in the absence of inventories, still raises the level of inventories necessary to deter deviations. This positive relationship between demand and inventories in a stationary setting would explain both the positive correlation between sales and inventories and the "excess" volatility of production.

A related issue which deserves exploration is the role of the inventory accumulation we predict when demand is high in accentuating fluctuations in aggregate output. It has often been pointed out (see Blinder and Fischer

(1981) for example) that the fluctuations in inventories are a large fraction of these output fluctuations. Thus the inventory fluctuations we predict are potentially important. However, it is not the case that only concentrated industries increase their output during booms. For our model to explain that together with aggregate inventory fluctuations, we must integrate the oligopolistic sector we consider with more competitive sectors in a general equilibrium model as is done in Rotemberg and Saloner (1985).

One question that such an integration must answer is whether the competitive industries would also accumulate inventories in a boom. The motivation for this might be the following. As in Rotemberg and Saloner (1985), booms would be consequences of shifts in demand towards oligopolistic sectors. These sectors would respond to these shifts by increasing their inventories and (possibly) lowering their prices as they do in our previous paper. This in turn raises real wages in terms of the goods produced by the oligopolistic sector and thus may, in equilibrium, lower real wages in terms of the goods produced competitively. This might make the competitive sectors increase production and increase inventories.

In this paper we have assumed that quantities are the strategic variable. However, it may well be the case that inventories perform an even stronger deterrent role when prices are the strategic variable. In this case deviations may be more tempting because a deviating firm captures a substantial fraction of the market. On the other hand, its rivals may respond more promptly in this case. The difficulty with analyzing inventories when prices are the strategic variable is that the optimal strategies, at least after a deviation has occurred, involve the use of mixed strategies.

Another issue which deserves exploration is the tradeoff between strategic buildups of capacity and inventories. Both have desirable consequences for a noncompetitive industry. As in Brock and Scheinkman (1981) increased capacity enhances the punishment to rivals who deviate from an implictly collusive arrangement. On the other hand, in our paper, it is the presence of costly capacity constraints which makes inventories necessary to keep rivals in line. Thus an optimizing oligopoly must invest in excess capacity knowing that inventories can, to some extent, be sustituted as a threat to deviators. It seems plausible that in an optimal arrangement most of the required fluctuations of this threat over the business cycle take the form of inventory fluctuations while capacity is picked with longer run considerations.

Capacity is also useful to a monopolistic incumbent firm as it tries to ward off entry (See Spence (1977) and Dixit (1980)). Such an incumbent can credibly threaten to use this extra capacity if another firm does enter. To some extent this is also useful to an oligopoly if it is believed that it will revert to competitive behavior after entry has occurred. Inventories can also be used in this manner. An oligopoly with more inventories will behave more aggressively after entry has occured. The question is whether inventories, which are shorter lived, will in fact be used for this purpose along with excess capacity. Their advantage is that they may be cheaper than capacity, particularly if they can be built up in periods of low demand.

The other important result we obtain is that the observability of overproduction can lower the equilibrium level of inventories and make the firms in the industry better off. If one is willing to draw the analogy between inventories and stockpiles of weapons (which is developed in several of our footnotes), this may shed light on the official US policy on simultaneous reductions of nuclear weapons by the US and the Soviet Union. That policy imposes as a prerequisite for these reductions that they be "verifiable". In other words it requires that any overproduction be not only observable but observable with little error. While error of observation is not explicitly considered in our paper, one can view the case in which overproduction is unobservable as a situation with large uncertainty as to the presence of overproduction. On the other hand the other case we consider has complete observability. The analysis of intermediate cases of partial observability is naturally of interest as well.

This analogy is incomplete for at least two reasons. First, countries do not necessarily find it in their interests to wage war on defenseless enemies. Second, wars are unlike deviations in sales in that it is less likely that the entire arsenal of the country will be instantaneously unleashed on the enemy at the outbreak of hostilities.

References

Blinder, A. S. "Retail Inventory Behavior and Business Fluctuations", Brookings Papers on Economic Activity, 1981:2 443-520.

————————: "Can the Production Smoothing Model of Inventory Be Saved" National Bureau of Economic Research Working Paper #1257 (January 1984).

———— and S. Fischer: "Inventories, Rational Expectations and the Business Cycle", Journal of Monetary Economics, 8, (November 1981), 277-304.

Brock, W. and J. Scheinkman, "Price Competition in Capacity-Constrained Supergames," University of Wisconsin, Working Paper #8130, September 1981 (Review of Economic Studies, forthcoming).

Dixit, A.K., "The Role of Investment in Entry Deterrence," Economic Journal, 90(1980), 95-106.

Eichenbaum, M. S.: "Rational Expectations and the Smoothing Properties of Inventories of Finished Goods", Journal of Monetary Economics, 14, (July 1984), 71-96.

Friedman, J.W., "A Non-Cooperative Equilibrium for Supergames," Review of Economic Studies, 28 (1971), 1-12.

Green, E.J. and R.H. Porter, "Noncooperative Collusion Under Imperfect Price Information", Econometrica, 52 (January 1984), 87-100.

Porter, R.H., "Optimal Cartel Trigger-Price Strategies", Journal of Economic Theory, 29 (1983), 313-338.

Radner, R., "Collusive Behavior in Noncooperative Epsilon-Equilibria of Oligopolies with Long but Finite Lives", Journal of Economic Theory, 22 (1980), 136-154.

Rotemberg, J.J. and G. Saloner, "A Supergame-Theoretic Model of Business Cycles and Price Wars During Booms," M.I.T. Working Paper #349, July 1984 (Revised February 1985).

Saloner, G, "The Role of Obsolescence and Inventory Costs in Providing Commitment", M.I.T., mimeo, October 1984.

Spence, A.M., "Entry, Capacity, Investment and Oligopolistic Pricing," Bell Journal of Economics, Vol. 8 (Autumn 1977), 534-544.

West, K. D.: "A Variance Bounds Test of the Linear Quadratic Inventory Model" Mimeo, Princeton University, (1983).

Table 1

Coefficients of Sales in Inventory Regressions*

SIC CODE		Coefficent+
20	Food & Kindred products	.404 (.099)
26	Paper & Allied Products	-.208 (.033)
28	Chemical & Allied Products	.932 (.048)
29	Petroleum & Coal Products	.185 (.063)
30	Rubber	.348 (.033)
	Other Nondurable Manufacturing	.135 (.043)
33	Primary Metals	-.185 (.025)
34	Fabricated Metals	.015 (.029)
35	Machinery Except Electrical	.262 (.059)
36	Electrical Machinery	.412 (.048)
371	Motor Vehicle and Parts	.198 (.027)
372-9	Other Transportation Equipment	.238 (.096)
	Other Durable Manufacturing	.013 (.050)

*Monthly Data 1967-1980
+Standard errors in parentheses

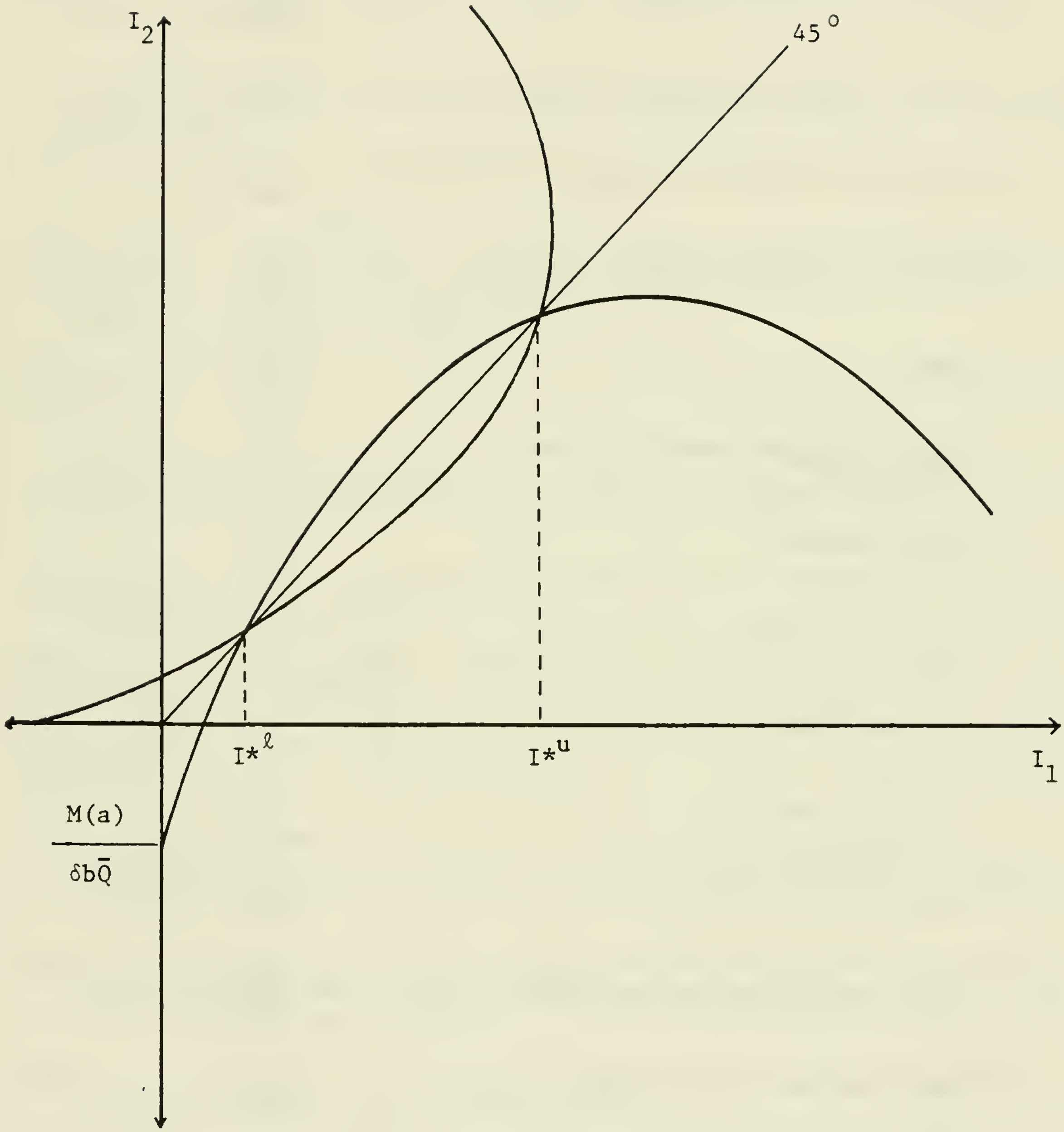

Figure 1: The smallest level of I_i that firm i requires to deter firm j from deviating when its inventories are I_j.

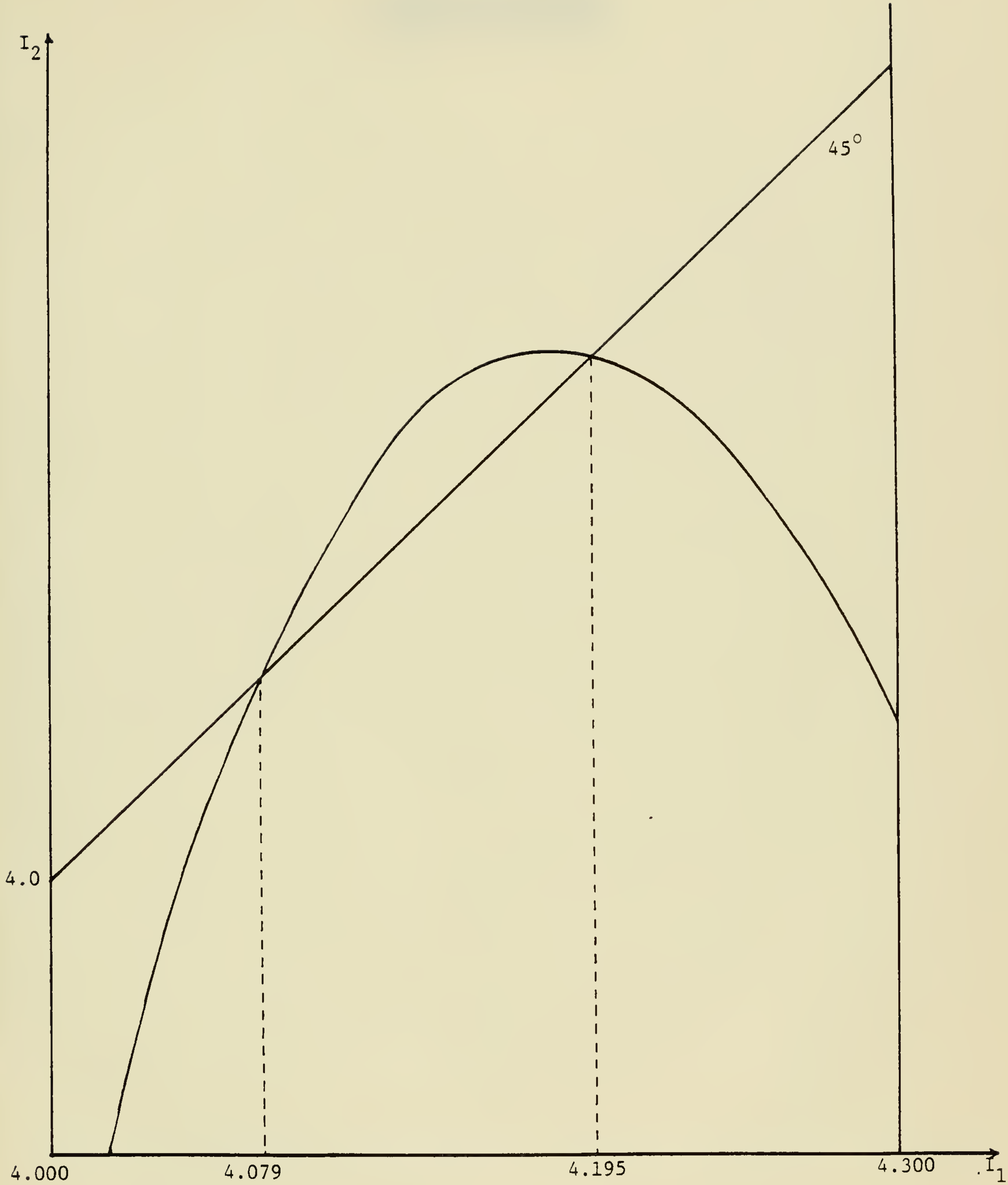

Figure 2: The smallest level of I_2 that deters firm 1 from deviating when its inventories are I_1.