

TRIGGER-POINT MECHANISM AND CONDITIONAL COMMITMENT: IMPLICATIONS FOR ENTRY, COLLUSION, AND WELFARE

LARRY D. QIU, LEONARD K. CHENG, and MICHAEL K. FUNG*

When fixed, sunk investment costs are high, firms may not have sufficient incentive to enter the market unless future entry is constrained. In this case, the government faces a dilemma between a full commitment and noncommitment of restricted future entry. A way out is to consider a commitment conditional on the realization of the uncertain parameters, such as the trigger-point mechanism (TPM) that sets conditions on current production level, excess capacity, and demand growth under which future entry will be allowed. This article shows that the TPM facilitates the incumbents' collusion but may improve social welfare under certain circumstances. (JEL L13, L43, L50, H10, H54)

I. INTRODUCTION

In regulated industries, a key issue faced by the regulator is how to induce investment to increase social welfare. For example, a regulated firm (monopoly) can make a research and development investment to reduce its production costs. After the investment has taken place, it is socially optimal for the regulator to reduce the regulated firm's supply price, thus reducing the firm's profit from such an investment. But anticipating the regulator's opportunistic behavior, the firm may not make the investment in the first place.¹ This

*The authors thank Professors Joseph Harrington Jr., Jaehong Kim, Oliver Williamson, Keith Wong, and two referees for helpful comments. They are grateful for financial support by a grant from the Research Grant Council of Hong Kong (HKUST6211/97H).

Qiu: Department of Economics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. Phone 852-2358-7628, Fax 852-2358-2084, E-mail larryqiu@ust.hk

Cheng: Department of Economics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. Phone 1-852-2358-7620, Fax 1-852-2358-2084, E-mail leonard@ust.hk

Fung: Department of Business Studies, Hong Kong Polytechnic University, Kowloon, Hong Kong. Phone 1-852-2766-7102, Fax 1-852-2653-3947, E-mail afmikef@inet.polyu.edu.hk

1. This leads to the time consistency problem as in the macroeconomics literature (see Kydland and Prescott 1977) or the opportunism problem as described by Williamson (1975, 1984). Williamson argues that if the regulator cannot guarantee not acting opportunistically, ex ante inefficiency will result. Besankol and Spulber (1992) cite the following real-world example of Hesse (1989): The Federal Energy Regulatory Commission observes that capacity shortages result from the limited commitment of the system in which normal return to cost-reducing investment by the regulated firms is not guaranteed.

type of ex ante inefficiency, however, can be mitigated or completely avoided by using some properly designed mechanisms.² One of those mechanisms is to write a long-term contract between the regulator and the regulated firm, under which the regulator fully commits to no change of policy. But as the environment (other than the firm's investment) changes, the policy may turn out to be bad and the regulator will have no means to correct it under the long-term contract. This represents the danger of overcommitment. As pointed out by Laffont and Tirole (1993, p. 620), "The benefit of commitment is that the regulated firm's investment is not expropriated. The cost of commitment is that the government may identify with the firm and bind the nation to a bad outcome over the long run."

The trade-off between full commitment and noncommitment naturally suggests that a mechanism of conditional commitment may be preferable. Although this type of mechanism has not been studied in the regulation literature, it has already been put to

2. Laffont and Tirole (1993, pp. 101–103) briefly discuss four mechanisms that can help mitigate the underinvestment effect.

ABBREVIATIONS

CHH: Container Handling Committee
PDB: Port Development Board
TPM: Trigger-Point Mechanism

use in the real world.³ One such example is the so-called trigger-point mechanism (TPM) used to regulate entry into container terminal service in the world's busiest container port, Hong Kong. To protect the financial interests of the private investors who developed the terminals, the Hong Kong Port Development Board (PDB) adopted the TPM, under which "new berths were not triggered until forecast throughput equalled working capacity of existing and planned berths" (PDB 1992, p. 23). The PDB makes a conditional commitment under the TPM with regard to entry: Entry will be allowed if future demand is sufficiently strong but not allowed otherwise.⁴

Another example is the second-sourcing of defense systems adopted by the U.S. Department of Defense. As described and modelled by Anton and Yao (1987), the mechanism specifies that if the developer (the primary contractor) produces x units in the first stage, then the remaining quantity, $z - x$ units (where z is the prespecified quantity) will be reproced and the contract will be awarded to the winner in a bid between the developer and a second contractor. Hence, reprocurement occurs if and only if $x < z$. If no reprocurement is interpreted as full commitment and reprocurement as noncommitment, then the mechanism of second-sourcing represents a conditional commitment.

The book edited by Levy and Spiller (1996) contains many other examples of conditional commitment regulations of the telecommunications industry in Chile, Jamaica, the Philippines, and the United Kingdom.

This article constructs a multiperiod model to investigate the effects of the regulator's conditional commitment, in the form of the TPM,

3. Baron (1989, p. 1410) points out that one way for the government to make a commitment is to issue a bond to the regulated firm. If the government changes its initial policy, the firm can claim the specified amount of money from the government. If the bond's value is very small, then it is equivalent to noncommitment. If the bond's value is sufficiently large, then it is equivalent to full commitment. If the bond's value is in the intermediate range, then it is equivalent to conditional commitment, i.e., the government does not change the policy in some circumstances but does in others.

4. In the cartel literature, it is well known that in the presence of incomplete information, firms can use the trigger-price strategy to facilitate collusion (see Green and Porter 1984). The TPM considered in this article is a contract between a regulator and firms in the regulated industry, but it has some features similar to the trigger-price strategy that is used by firms to facilitate interfirm cooperation.

on entry, collusion, and social welfare. In the first period, two firms contemplate entering the market in which future entry by a third firm is regulated. Because entry in the first period incurs a large sunk cost, the government's entry regulation affects first period entry as well as future entry. To see commitment value as clearly as possible, we suppose that there will be entry in the second period in the absence of regulation. Because future entry heightens competition and lowers the incumbent firms' future profits, their discounted profits may not be large enough to recover the sunk cost of entry in the first period if the government does not prohibit future entry. That means the firms may not enter at all, resulting in a loss of welfare. This calls for commitment by the government to disallowing future entry. However, if future demand proves to be sufficiently strong, then permitting future entry may not discourage first-period entry. In that case, full commitment to absolutely no entry may generate lower social welfare than noncommitment. Therefore, in an environment with uncertain demand, a TPM, which specifies conditions for new entry, may be more desirable than both full commitment and noncommitment.

After entry in the first period, the two incumbents may collude by producing an output level lower than that without collusion. Without the TPM, however, future entry will destroy current collusion. The authors show that in many cases the TPM not only restores first-period collusion but also induces the incumbents to collude at an even lower output level for the purpose of reducing entry.⁵ The reason is simple. In addition to the usual profit-enhancing motivation for collusion, with the TPM the condition of entry depends on the gap between capacity and current demand. The incumbents have an additional incentive to lower the output level such that entry will not be triggered.⁶

5. This supports the following concern by the Hong Kong Centre for Economic Research (1992): "Since the amount of excess demand is determined in part by the level of handling charges set by terminal operators, this mechanism provides existing operators with a tool to limit competition by reducing the rate of entry."

6. Collusion under a TPM has its counterpart in international trade. Rotemberg and Saloner (1989) have analyzed the impact of import quota on collusion. Krishna and Morgan (1998) show that voluntary import expansion facilitates/induces collusion among the import competing firms.

In general, the effect of TPM on social welfare is ambiguous. It has a positive effect because it induces current entry, but it has a negative effect because it facilitates collusion, which reduces competition and blocks future entry. Nevertheless, the authors show that TPM improves overall welfare in some cases, depending (among other factors) on the level of entry cost. Therefore, although a simple TPM will not be able to achieve an efficient outcome in all cases, it may still be a preferred policy if the regulator does not have information about entry costs.

The present study contributes to the regulation literature by showing the potential superiority of the regulator's conditional commitment over both full commitment and non-commitment and by deriving the implications of such a mechanism for the incumbents' collusive behavior and social welfare. It differs from the literature in a number of ways. Laffont and Tirole (1993, chapter 16) construct a model to illustrate "the trade-off between encouraging investment through commitment and correcting wrong policies through noncommitment." The present article goes further by examining the implications of conditional commitment for firms' collusive behavior. Though the existing studies in the literature have largely focused on the regulation of a franchised monopolist, in which case the complication of strategic interaction among the regulated firms is absent (see Baron 1989, p. 1350).⁷ This article assumes two firms in a regulated industry to explicitly analyze their strategic interaction (competition and collusion).

This article is related to Kim (1997), who assumes excess entry in the absence of regulation in contrary to the present assumption of insufficient entry without regulation. In Kim's model, a single incumbent sets its production capacity in the first period with a view of deterring entry in the second period. In between the two periods, the government decides whether to allow entry in the second period. Kim shows that the government's entry regulation lowers the incumbent's cost of deterring entry and hence it may deter even socially optimal entry. Two additional differences distinguish this article from Kim's. First, in the present model the government sets the TPM

7. Gilbert and Vives (1986) is one exception, but they do not consider collusion between incumbents to deter entry.

before the first-period investment takes place, so the incumbents' firms are not able to act strategically to influence the government's mechanism design. Second, there are multiple incumbents in the model so that the authors can investigate the issue of collusion among the incumbents for the purpose of deterring future entry under the TPM.⁸

The rest of the paper is organized as follows. The model is set up in section II. Section III analyzes the equilibrium under TPM-sanctioned entry. Section IV analyzes the first-period entry and welfare. Section V discusses robustness of the results and concludes.

II. THE MODEL

There are infinitely many periods. Consider an industry in which firms produce a homogeneous good. Demand in period t is assumed to be $p_t = a_t - X$. The authors focus on one particular scenario where there is a one-time (permanent) demand shock (expansion) that occurs in period 2. Specifically, assume $a_1 = 1$ and $a_t = a > 1$, for all $t > 1$. Assume that a is a random variable with distribution $F(a)$, $a \in (1, \bar{a}]$. Denote $E(x) = \int_1^{\bar{a}} x dF(x)$.

Before period 1, the government announces the TPM, which specifies the conditions under which new entry will be allowed in period 2. In period 1, two risk-neutral firms, 1 and 2, make their respective entry decisions. After the entry decisions are made, the firms know more about the demand (through market research). For simplicity the authors assume that they know the exact value of a , capturing the fact that firms already in the market have more information about demand than the government. But at the time of entry they need to make entry decision based on expected payoffs because they are not in the market yet. If both firms enter, each firm incurs a fixed and sunk cost s_2 . If only one firm enters, the fixed and sunk cost is s_1 . In the case of two-firm entry and if the firms do not collude, the authors assume the equilibrium in the product market is achieved through Cournot competition. If there is no entry, there will be no production

8. There is a large literature on entry deterrence (i.e., barriers to entry created by incumbents' strategic behavior, not by the government), e.g., limit pricing, overinvestment, extensive noncooperative competition, etc. See Gilbert (1989) and Tirole (1988) for a survey of this literature. In contrast, in this article, explicit entry barriers are created by the government.

in this period, and the game moves to the second period. The authors assume that if firm 1 or firm 2 does not enter in the first period, it never enters in the future because it will have invested in some other projects.

At the beginning of the second period, a is realized to all parties (including the regulator) and the TPM determines whether new entry is allowed. Assume that the potential entrant, firm 3, will always enter in period 2 if entry is allowed. This will be the case if the entry cost to firm 3 in period 2, denoted by s_3 , is much lower than the entry cost to firms 1 and 2 in period 1. It will soon become clear that if no firm enters in period 1, the TPM always allows firm 3 to enter. Then in period 2, all firms observe a and produce for the market. All subsequent periods are identical to period 2.

For simplicity and without loss of generality, assume that each firm can produce at constant marginal cost, equal to 0, up to a capacity level K . This assumption captures the lumpiness of investment in container terminals. Moreover, K is exogenous because it is to a large extent determined by the government.⁹

Before the authors lay down the mathematical representation of the TPM, they describe its institutional background and actual practice. In 1990, the PDB of Hong Kong was established for the strategic planning of Hong Kong's port. Under the PDB, the Container Handling Committee (CHC) plans and oversees the development of new container terminals. Its aim is to "ensure that supply would continue to meet demand and to obviate overprovision, or underprovision, of container terminal capacity" (PDB 1992, p. 23). When determining whether a new terminal will be built to meet future demand, the CHC uses the recent data of container throughput to forecast demand growth in the future. Based on the forecast demand, the CHC decides

9. In the container terminal business, the capacity of each operator is largely constrained by the quay front and back-up land, both of which are determined by the regulator. For instance, the newest terminal in Hong Kong (Terminal no. 9) features 320 meters of quay front and 15 hectares of back-up land per berth. The length of the quay front determines how many ships can be serviced at the same time. Containers must first be stored in the container yard (i.e., on the back-up land) before they are loaded onto the ships or after they are unloaded from the ships. Although it is not impossible for the terminal operators to make investment in equipment and yard to increase their container-handling capacity, this kind of investment is very costly and inevitably runs into diminishing returns.

whether to trigger a new terminal. The CHC has little room to manipulate the decision because the TPM is demand-led and a new terminal will not be triggered until forecast throughput equaled working capacity of existing and planned terminals. In particular, the trigger point is "set at the throughput level where forecast demand exceeded existing capacity" (PDB 1992, p. 23).

To sum up, the idea behind the TPM is that if the growth in demand is not strong enough, the incumbents will have sufficient capacity to satisfy future demand and so entry should not be allowed. If growth in demand is sufficiently strong, however, new entry is justifiable.¹⁰ In terms of the level of demand, let X_1 be the actual quantity demanded (the available data observed and used by the CHC) in the first period, X be the future quantity demanded, and K_1 be the total capacity in the first period. The TPM says that additional berths will be built and new entry will be triggered if and only if X is greater than or equal to the existing capacity K_1 . An equivalent way of capturing the essence of the TPM is in terms of demand growth and existing idle capacity. Note $(a - 1)$ captures the growth in demand. By definition, the idle capacity is given by $(K_1 - X_1)$. Accordingly, the TPM is specified mathematically as follows:

- (1) New entry is allowed if and only if

$$(a - 1) \geq \xi(K_1 - X_1),$$

where $\xi(\bullet) > 0$, $\xi'(\bullet) > 0$.

Because there is no growth after the second period, the authors assume that the TPM disallows entry after period 2.

The authors do not claim that condition (1) is the actual TPM in the real world but use it as a simple mechanism to capture the government's conditional commitment, namely, allowing entry if the growth in demand is relatively strong but not allowing entry if it is relatively weak.¹¹ They believe this simplification

10. It is justifiable for two reasons. First, allowing for new entry will not reduce incumbents' profit much because future demand is strong. Second, even if the incumbents can incur higher marginal production costs beyond their design capacity, a new entrant may result in lower total costs for the entire industry.

11. The first-period idle capacity has a key effect on the government's decision. This can also be seen in Anton and Yao's (1987) modeling of the Department of Defense's mechanism for regulating new entry and in Kim's (1997) optimal government decision on new entry.

is justified because the purpose of this article is not to design an optimal TPM but to investigate the implications of such a mechanism for collusion and welfare.

III. COLLUSION UNDER TPM-SANCTIONED ENTRY

To underscore the importance of TPM and perform an analysis of collusion in a repeated game, first consider the following case: Firms 1 and 2 have already entered the market and there is no future entry.

The authors use superscripts m , c , n , and d to denote monopoly, collusion, Cournot-Nash competition, and defection, respectively. In deriving the various output levels, for convenience the authors ignore the capacity constraints. Even though capacity is crucial to the TPM, for a large range of parameter values (specified later) the authors find that firms' outputs are not constrained by the capacity. In the case when capacity constraint is binding, firms' competition and collusion will be affected. The effects of binding capacity constraint will be briefly discussed in section V.

In the absence of collusion, the equilibrium is given by that of Cournot. They are (for each firm), respectively, $x_1^n = 1/3$, $p_1^n = 1/3$, and $\pi_1^n = 1/9$ for period 1, and $x^n = a/3$, $p^n = a/3$, and $\pi^n = a^2/9$ for every other period onward. If there is a single monopoly or if the two firms maximize joint profit, the optimal (total) output, price and (total) profit are, respectively, $x_1^m = 1/2$, $p_1^m = 1/2$, and $\pi_1^m = 1/4$ for period 1, and $x^m = a/2$, $p^m = a/2$, and $\pi^m = a^2/4$ for every other period onward.

When firms tacitly collude at output x_1^c in period 1: that is, each produces $x_1^c/2$, then the price and each firm's profit are, respectively, $p_1^c = 1 - x_1^c$ and $\pi_1^c = (1 - x_1^c)x_1^c/2$. Similarly, if they collude at output x^c in every other period, then the price and each firm's profit are, respectively, $p^c = a - x^c$ and $\pi^c = (a - x^c)x^c/2$. The colluding firms will choose $x_1^c \in [x_1^m, 2x_1^n)$ and $x^c \in [x^m, 2x^n)$ as low as possible.

As it is common in the literature, the authors follow Friedman (1971) to use grim strategy to model punishment schemes on defection.¹² Specifically, when a firm defects from x_1^c or x^c , the other firm will not continue to produce at the collusive level. However, as-

sume that the other firm is not able to respond until one period later, implying that the defector's optimal output and the resulting price and profit during the defection period are, respectively: $x_1^d = p_1^d = (2 - x^c)/4$, and $\pi_1^d = (2 - x_1^c)^2/16$; or $x^d = p^d = (2a - x^c)/4$, and $\pi^d = (2a - x^c)^2/16$.

The authors can analyze the defection/collusion behavior for every period, but given that they are most interested in the first-period collusion between the incumbent firms in the main model with TPM, they just focus on the first-period equilibrium. When contemplating defection in the first period, a firm must weigh the one-time gain for defection ($G(x^c) \equiv \pi_1^d - \pi_1^c$) against the discounted future losses from the breakdown in collusion ($L(x^c) \equiv \pi^c - \pi^n$). Assuming that firms have the same discount factor, $\delta \in (0, 1)$, a firm's *defection incentive (DI)* is given by

$$(2) \quad DI(x_1^c) \equiv G(x_1^c) - \frac{\delta}{1 - \delta} L(x^c)$$

Collusion at x_1^c is sustainable if and only if $DI(x_1^c) \leq 0$. When firms discount future profits less (larger δ), they have less incentive to defect from current collusion.

The authors now return to the model in which firm 3 is waiting to enter in period 2. Suppose there is no TPM in place. Then, if firms 1 and 2 collude in the first period and firm 3 enters in the second period, a natural question is how the incumbent firms respond to the entry. As pointed out by Harrington (1989), there are generally two classes of response. First, the incumbent firms break up their collusion in response to entry, that is, all three firms will play the Cournot competition game in all periods after entry. Second, the incumbent firms accommodate the entrant by establishing a new collusion including firm 3.¹³ To make the implications of TPM for collusion as sharp as possible, this article adopts the first class of response, in which entry is not accommodated. Section V discusses the robustness of results when entry is accommodated.

It turns out that with future entry and in the absence of the TPM, collusion in the first period cannot be sustained. The reason is that because the outcome in the second period and

12. For other more appealing and sophisticated punishment schemes to support collusion, see Green and Porter (1984) and Abreu (1986).

13. Within the second class of response, Harrington (1989) also considers the case in which the incumbent firms coordinate in punishing the entrant (predation).

beyond is always a three-firm Cournot-Nash equilibrium, there is no future loss due to current defection ($L(x^c) = 0$ because there is no collusion anyway), but there is always a one-time gain from defection in the first period. Thus, $DI(x_1^c) > 0$ for all x_1^c , and so collusion in the first period cannot be sustained.

The authors shall show that the TPM not only restores the first period collusion but in some cases also makes it more collusive, because the incumbents restrict output further to reduce the threat of potential entry.

First the authors derive the most collusive output after the first period under the assumption that firm 3 does not enter in period 2. This result will be used repeatedly later. From period 2 onward, the defection incentive becomes $DI(x^c) \equiv G(x^c) - \delta/(1 - \delta)L(x^c)$. Let $x^o = 2a(9 - 5\delta)/3(9 - \delta)$ and x^* be the optimal (lowest) sustainable output for collusion between firms 1 and 2 in every period from period 2 onward. It can be shown that $x^* = x^m$ for $\delta \geq 9/17$ and $x^* = x^o$ for $\delta < 9/17$. If $\delta < 9/17$, $x^o > x^m$. Accordingly, in this article, the discount factor is considered high if $\delta \geq 9/17$, and low if $\delta < 9/17$.

Suppose the TPM as specified by (1) takes the following specific form:

- (3) entry is allowed if $X_1 > v$,
but entry is not allowed if $X_1 \leq v$,

where the “entry threshold” $v \equiv 2K - (a - 1)$. Thus, if $K > 3/8$, firms’ outputs will not be constrained by the capacity in all relevant circumstances analyzed. Moreover, if demand growth is expected to be sufficiently high (e.g., $a \geq 2$), it is reasonable to believe that existing incumbents’ capacity will not be sufficient to meet the demand, and therefore the TPM is justifiable.

Depending on the value of v , the authors can distinguish between three cases:

- Case A: $v \geq 2x_1^m$; Case B: $v \in [x_1^m, 2x_1^m)$;
and Case C: $v < x_1^m$,

where $x_1^m (= 1/2)$ is the monopoly output in the first period and $x_1^n (= 1/3)$ is the Cournot-Nash equilibrium output of each of the two incumbent firms in the first period. The rest of this section assumes that under the TPM as given by (3) two firms will enter in the first period. Firms 1 and 2’s entry decisions in period 1 will be considered in section IV.

Defection under TPM-Sanctioned Entry

When TPM as specified in (3) is in place, the first-period collusion will be different. Defection not only triggers punishment by the colluding partner but also results in new entry if the first-period output becomes sufficiently high due to defection. However, if the threat of entry is low (i.e., v is very large), the increased output as a result of defection will not trigger new entry. On the other hand, if the threat of entry is high (i.e., v is very small), even a small deviation will trigger new entry. In these two cases, the defector’s behavior will not be different from that in the absence of TPM. In other cases, TPM constrains the defector’s output because the defector does not wish to trigger new entry. The authors demonstrate this intuition in this subsection.

Suppose in the absence of future entry firms 1 and 2 collude at $x_1^c \in [x_1^m, 2x_1^n)$ in the first period and at x^c in the second period and beyond. Without the risk of future entry, a defector’s first-period defecting optimal output is $x_1^d = (2 - x_1^c)/4$, implying that the total output in the first period is equal to $X_1(x_1^c) = (2 + x_1^c)/4$. Whether this “unconstrained cheating output” is optimal under the TPM depends on whether X_1 will trigger entry. If $X_1 \leq v$, then the unconstrained cheating output is optimal.

However, if $X_1 > v$, then the defect will induce entry in the second period, implying that the three firms will play the Cournot competition game after period 1. To prevent entry by firm 3, the defector must choose a lower constrained output. The optimal “constrained cheating output” is equal to $x_1^v \equiv v - x_1^c/2$, implying that the first period’s total output is equal to v . After cheating, the two incumbents play the Cournot competition game in all future periods.

If the defector produces the unconstrained cheating output, then its total discounted profit is

$$\begin{aligned} \Pi^{un} &= \frac{(2 - x_1^c)^2}{16} + \frac{\delta}{1 - \delta} \frac{a^2}{16} \quad \text{if } X_1 > v, \text{ or} \\ \Pi^d &= \frac{(2 - x_1^c)^2}{16} + \frac{\delta}{1 - \delta} \frac{a^2}{9} \quad \text{if } X_1 \leq v. \end{aligned}$$

If the defector produces the constrained cheating output, then its total discounted profit is

$$\Pi^{con} = (1 - v) \left(v - \frac{x_1^c}{2} \right) + \frac{\delta}{1 - \delta} \frac{a^2}{9}.$$

Note, $\Pi^d > \max\{\Pi^{un}, \Pi^{con}\}$, but the comparison of Π^{un} and Π^{con} is indeterminate. Intuitively, when v is small, the constrained cheating output might be too small to make it worthwhile to deter entry (i.e., $\Pi^{un} > \Pi^{con}$). However, if future demand is high and the firms do not discount future profits too much, it could be worthwhile to prevent entry by choosing the constrained cheating output. If v is very large, there is no need to worry about entry at all. These intuitive results are stated more precisely in Lemma 1.

Lemma 1. *Suppose a firm defects from a collusive output $x_1^c \in [x_1^m, 2x_1^c]$ in the first period.*

- i. *If $v \geq 2x_1^m$ or $v < x_1^m$, then the defector's output is equal to x_1^d .*
- ii. *If $v \in (5/8, 2x_1^m)$, then there exists $z_0 \in (2(2v - 1), v]$ such that the defector's output is given by*

$$\begin{cases} x_1^v & \text{for } x_1^c \in (2(2v - 1), z_0), \\ x_1^d & \text{otherwise.} \end{cases}$$

- iii. *If $v \in [x_1^m, 5/8)$, then there exists $z_1 \in [x_1^m, v]$ such that the defector's output is given by*

$$\begin{cases} x_1^v & \text{for } x_1^c \in [x_1^m, z_1), \\ x_1^d & \text{otherwise.} \end{cases}$$

Proof. See appendix. □

The intuition behind Lemma 1(i) is simple. The defector chooses the unconstrained cheating output when there is either always no entry (in Case A) or always entry (in Case C). The authors will explain the intuition behind Lemma 1(ii) and (iii), which belong to Case B. In (ii), v is relatively large. One can divide the interval $[x_1^m, v]$ into three regions, $[x_1^m, 2(2v - 1)]$, $(2(2v - 1), z_0)$ and $[z_0, v]$. If x_1^c falls into the first region, it is far below v , and so the unconstrained cheating output does not induce entry. The defector gets the highest discounted profit Π^d by producing this amount. If x_1^c falls into the second region, the unconstrained cheating output will induce entry while the constrained cheating output is not too restrictive, and so the defector chooses this latter output level to get Π^{con} . However, if x_1^c falls into the third region, the unconstrained cheating output will induce

entry, but the constrained cheating output will be too restrictive to be worthwhile. Hence, the defector chooses the unconstrained cheating output and gets Π^{un} .

In Lemma 1(iii), v is already small, so the first region does not exist. As in the case of (ii), the defector chooses the constrained cheating output when x_1^c is in the second region $[x_1^m, z_1)$, but chooses the unconstrained cheating output when x_1^c is in the third region $[z_1, v]$.

In the next two subsections, the authors analyze the defection incentive under TPM to determine whether the two incumbent firms will collude and at which output level. Because the TPM may constrain the defector's output, the one-time gain from defection could be reduced, which in turn reduces the incentive to defect at a given collusive output. As a result, the first-period collusion output could be lower as a result of the TPM. This intuition will be borne out below for both high and low discount factors.

Collusion When Discount Factor Is High

This subsection analyzes collusion when the discount factor is relatively large (i.e., $\delta \geq 9/17$), and the next subsection deals with the case of a smaller discount factor. First, consider Case A where v takes on a large value. Following Lemma 1, the authors know that at any first period collusive output x_1^c the defector always produces the unconstrained cheating output and entry does not occur under the TPM. Hence, the defection incentive is

$$(4) \quad DI_H(x_1^c) \equiv \frac{(2 - x_1^c)^2}{16} - \frac{(1 - x_1^c)x_1^c}{2} - \frac{\delta}{1 - \delta} \frac{a^2}{72}.$$

Noting that $\partial DI_H / \partial x_1^c < 0$ for all $x^c < 2x_1^m$ and $DI_H(x_1^m) \leq 0$ given that $a \geq 1$ and $\delta \geq 9/17$, the authors conclude that in Case A the optimal sustainable collusive output is the monopoly output x_1^m .

Then, the authors turn to Case B where v takes on a moderate value. Recalling that collusion never occurs at any output within the interval $(v, 2x_1^m)$, the authors focus on $x_1^c \in (1/2, v]$. If the defector defects from collusion at x_1^m but keeps $X_1 \leq v$, then the future loss from the current defection is the same as in Case A, but the current one-time gain cannot be higher than $G_1(x_1^m)$ because the defector

may not be able to choose the unconstrained cheating output. Alternatively, if the defector produces the unconstrained cheating output, then the one-time gain is the same as in Case A, but the future loss is at least as large as that in Case A. In both cases, the defection incentive is no greater than that in Case A, and thus the authors conclude that in Case B collusion at the monopoly output x_1^m is sustainable.

It is interesting to note that although collusion at the monopoly output is sustainable, collusion at some higher output level ($x_1^c > v$) is not sustainable. This counterintuitive result arises because any higher output level induces entry and consequently a breakdown of collusion.

Last, the authors consider Case C, where any sustainable collusive output must be below $v (< x_1^m)$. Each firm's profit in the first period with collusion output x^c is concave in x^c and reaches its maximum at x_1^m if entry is ignored. In addition, the defector's profit $(2 - x_1^c)^2/16$ (and so the defection incentive) decreases as x^c increases. These two observations together imply that if the firms collude, they should collude at $x_1^c = v$. Knowing this, the authors next derive conditions under which collusion at v is both beneficial and sustainable.

If the firms collude in the first period, they will collude at x^m in every subsequent period, giving each incumbent firm a total discounted profit $\Pi_c \equiv (1 - v)v/2 + a^2\delta/8(1 - \delta)$. On the other hand, if there is no collusion in the first period, they play the Cournot competition game in all periods, giving each incumbent firm a total discounted profit $\Pi_n \equiv 1/9 + a^2\delta/16(1 - \delta)$. The defection incentive is given by $[(2 - 3v)^2(1 - \delta) - a^2\delta]/16(1 - \delta)$, which is nonpositive if and only if $v \geq v_0 \equiv [2 - a\sqrt{\delta/(1 - \delta)}]/3$, where $v_0 < x_1^m$. In other words, the incumbents have no incentive to defect from v unless v is very small (i.e., at $v < v_0$). But is colluding at v beneficial to the incumbent firms? Because $(\Pi_c - \Pi_n)$ increases in v and is positive when v is close to x_1^m , it follows that in Case C $\Pi_c > \Pi_n$ for all $v \geq v_0$.

The authors now summarize the foregoing analysis in Proposition 1.

Proposition 1 (Collusion with a Larger Discount Factor). *Suppose the discount factor is sufficiently large (i.e., $\delta \geq 9/17$). The introduction of the TPM not only restores collusion between*

the incumbents but in some cases also induces the incumbent firms to restrict output below the monopoly output.

More precisely, if $v \geq x_1^m$, then the incumbent firms collude at the monopoly output x_1^m . If $v_0 \leq v < x_1^m$, then the incumbent firms collude at the output level v , which is below the monopoly output x_1^m . If $v < v_0$, then no collusion between the incumbent firms in the first period can be sustained.

Collusion When Discount Factor Is Small

This subsection analyzes the case for $\delta < 9/17$. As shown before the introduction of (3), in this case the lowest sustainable output for collusion from second period onward is given by x^o if there is no entry in the second period. The firms' defection incentive in period 1 is given by

$$(5) \quad DI_L(x_1^c) \equiv \frac{(2 - 3x_1^c)^2}{16} - \frac{\delta}{1 - \delta} \left(\pi^o - \frac{a^2}{9} \right),$$

where $\pi^o \equiv \pi^c(x^o) = a^2(9 - 5\delta)(9 + 7\delta)/9(9 - \delta)^2$. Noting that $x_1^c \in [x_1^m, 2x_1^m]$, we obtain x_1^o as the unique solution to $DI_L(x_1^c) = 0$:

$$x_1^o \equiv \frac{2(9 - \delta - 4a\delta)}{3(9 - \delta)}.$$

Letting x_1^* denote the first period's collusion output in the absence of future entry. The analysis is summarized as

Lemma 2. *Suppose $\delta < 9/17$ and there is no future entry. Then the first period's collusive output is given by*

$$x_1^* = \begin{cases} x_1^m & \text{for } \delta \geq \delta_1, \\ x_1^o \in (x_1^m, 2x_1^m) & \text{for } \delta < \delta_1, \end{cases}$$

where $\delta_1 \equiv \frac{9}{16a + 1} < \frac{9}{17}$.

Moreover, x_1^ is lower than the collusion output (x^o) in every other period.*

The authors now examine the defection incentive under the condition that any defection in the first period leads to new entry. The defection incentive in this case is given by

$$\frac{(2 - 3x_1^c)^2}{16} - \frac{\delta}{1 - \delta} \left(\pi^o - \frac{a^2}{16} \right).$$

The DI is equal to 0 at

$$x_1^e \equiv \frac{1}{3} \left[2 - \sqrt{\frac{\delta}{1-\delta}(16\pi^o - a^2)} \right].$$

Finally, if the incumbents can only sustain their collusion at $v < x_1^m$, the authors need to verify if it is beneficial, that is, $(\tilde{\Pi}_c - \Pi_n) > 0$, where $\tilde{\Pi}_c \equiv (1-v)v/2 + \delta\pi^o/(1-\delta)$ is the counterpart of Π_c as given previously.

Armed with these preliminary results, the authors now state the counterpart of Proposition 1.

Proposition 2 (Collusion with a Small Discount Factor). *Suppose $\delta < 9/17$. The first period collusive output is characterized as follows.*

(i). *If (a) $v \geq 2x_1^m$, or (b) $v \in [x_1^m, 2x_1^m]$ with $\delta \in [\delta_1, 9/17]$, or (c) $v \in (\frac{2}{3}(1 - \frac{a\delta}{9-\delta}), 2x_1^m)$ with $\delta < \delta_1$, then the incumbents collude at x_1^* .*

(ii). *If (a) $v < \min\{x_1^e, x_1^*\}$, or (b) $\tilde{\Pi}_c - \Pi_n < 0$ and $v < x_1^m$, then there will be no collusion at all.*

(iii). *In all other cases, the incumbents collude at some output level below x_1^* .*

Proof. See appendix. \square

Although the conditions for various outcomes described in Proposition 2 look complicated, they are in fact quite intuitive. First, the same collusive output without entry (i.e., x_1^*) can be restored by TPM if entry is automatically deterred by TPM without the incumbent firms choosing their first-period output strategically to deter entry. This happens if v is very large (i.e., under condition (i)(a) entry is deterred in any case), or if δ is relatively large (i.e., condition (i)(b), so the firms value collusion more), or if $(v - x_1^*)$ is sufficiently large (i.e., condition (i)(c), which is equivalent to $v \in (5/8, 2x_1^m)$ and $x_1^* \in (x_1^m, 2(2v - 1))$), so v does not affect the defection incentive at x_1^* . In all of these cases, the threat from TPM-sanctioned entry is very weak.

Second, at the other extreme where the pressure of TPM-sanctioned entry is very strong, the incumbent firms have to collude at output levels that are too low to be sustainable or profitable to prevent entry. This occurs when v is not only lower than x_1^* but also lower than x_1^e (i.e., condition (ii)(a), so the tempta-

tion to defect is too high), or condition (ii)(b) holds (so it is too costly to deter entry as the firms have to produce at an output level much below the monopoly level).

Last, when TPM-sanctioned entry is neither very weak nor very strong in the sense described, the firms can utilize the TPM to support more collusive outcome if $v > x_1^m$. The reason is that the TPM lowers defection incentive because the defector will receive more punishment (if the defector chooses the unconstrained cheating output) or get less one-time gain (if the defector chooses the constrained cheating output). This explains why under the TPM the firms go further in output restriction for the purpose of entry deterrence.¹⁴

IV. ENTRY AND WELFARE ANALYSIS

This section analyzes the first-period entry decision under TPM and TPM's welfare results. For simplicity, the authors focus on the case that the demand size a is uniformly distributed over $[1, 2]$. They shall focus on the role of entry costs, which affect the firms' entry decisions together with the pressure of future entry, TPM, and their effects on collusion. The first subsection characterizes the conditions for single entry and two-firm entry. The second subsection performs some welfare analysis based on the result obtained from the first subsection. The authors highlight an important welfare trade-off that is inherent in the TPM: a positive effect because TPM encourages first-period entry and a negative effect because TPM facilitates first-period collusion and blocks future entry. Examples in which TPM improves social welfare will be presented.

Entry

To see the impact of TPM on entry, the authors first examine conditions for entry in the absence of TPM. For simplicity, let $s_1 = s_2 = s$ be the entry cost incurred by each firm. Recall that firm 3 always enters (which imposes condition on s_3 in the welfare analysis) in the second period. If only one firm enters in

14. In limit pricing, the incumbent(s) will set a lower price to deter future entry. This strategic behavior changes in the presence of government entry regulation. With the TPM, the first period price is higher (output is lower) than otherwise to deter future entry.

the first period, the entrant's expected profit is equal to

$$\frac{1}{4} + \frac{\delta}{1-\delta} \int_1^2 \frac{a^2}{9} da - s = \frac{1}{4} + \frac{7\delta}{27(1-\delta)} - s.$$

Similarly, if two firms enter in the first period, each entrant's expected profit is equal to $1/9 + 7\delta/48(1-\delta) - s$. Hence, there is single entry if and only if

$$\frac{1}{9} + \frac{7\delta}{48(1-\delta)} \leq s < z_3 \equiv \frac{1}{4} + \frac{7\delta}{27(1-\delta)};$$

and there is double entry if and only if $s < z_6 \equiv 1/9 + 7\delta/48(1-\delta)$.

Now consider the impact of the TPM. Suppose $v \geq 1/2$. In the case of single entry, the entrant just produces the monopoly output in the first period and successfully block firm 3's entry. Its expected profit is

$$\frac{1}{4} + \frac{\delta}{1-\delta} \int_1^2 \frac{a^2}{4} da - s = \frac{1}{4} + \frac{7\delta}{12(1-\delta)} - s.$$

In the case of two-firm entry, Propositions 1 and 2 show that the output level at which the two firms collude depends on δ . The authors first derive entry condition with high discount factor $\delta \geq 9/17$. The firms collude at the monopoly output level in the first period and successfully block future entry. Each firm's expected profit is equal to $1/8 + 7\delta/24(1-\delta) - s$. As a result, there is no entry if $s \geq z_1 \equiv 1/4 + 7\delta/12(1-\delta)$; there is single entry if $z_4 \equiv 1/8 + 7\delta/24(1-\delta) \leq s < z_1$; and there is two-firm entry if $s < z_4$.

Suppose $v < 1/2$. For single entry, the entrant compares its expected profit from setting monopoly output in the first period while allowing future entry, and the expected profit from setting lower first-period output at v while blocking future entry. For $\delta \geq 9/17$, it always choose the latter and its expected profit as $v(1-v) + 7\delta/12(1-\delta) - s$. For two-firm entry, if $v \geq v_0$, they collude to block future entry and each firm's expected profit is $v(1-v)/2 + 7\delta/24(1-\delta) - s$; if $v < v_0$, they do not collude, allowing future entry, and each's expected profit is $1/9 + 7\delta/48(1-\delta) - s$. Hence, there is no entry if $s \geq z_2 \equiv v(1-v) + 7\delta/12(1-\delta)$; there is single entry if $v \geq v_0$ and $z_5 \equiv v(1-v)/2 + 7\delta/24(1-\delta) \leq s < z_2$, or if $v < v_0$ and $z_6 \leq s < z_2$; there is two-firm entry if $v \geq v_0$ and $s < z_5$, or if $v < v_0$ and $s < z_6$.

Comparing related critical levels yields $z_1 > z_2 > z_3$ and $z_4 > z_5 > z_6$. The results about first-period entry are summarized in Figure 1. In the figure, for each scenario there are three regions: two-firm entry when s is small, single entry when s is at medium levels, and no entry when s is large. These three regions corresponding to the case of no TPM are shown by dotted lines/arrows, and those corresponding to the case of TPM are shown in solid lines/arrows. With the TPM, there are also three scenarios representing different levels of v . For example, in the scenario of $v \geq 1/2$, the two critical levels which define the corresponding three regions are z_4 and z_1 : two-firm entry when $s < z_4$, single entry when $s \in [z_4, z_1)$, and no entry when $s \geq z_1$. The figure shows that relaxing the TPM in the sense of raising v encourages first period entry. Moreover, this entry is more likely under TPM than under no TPM.¹⁵

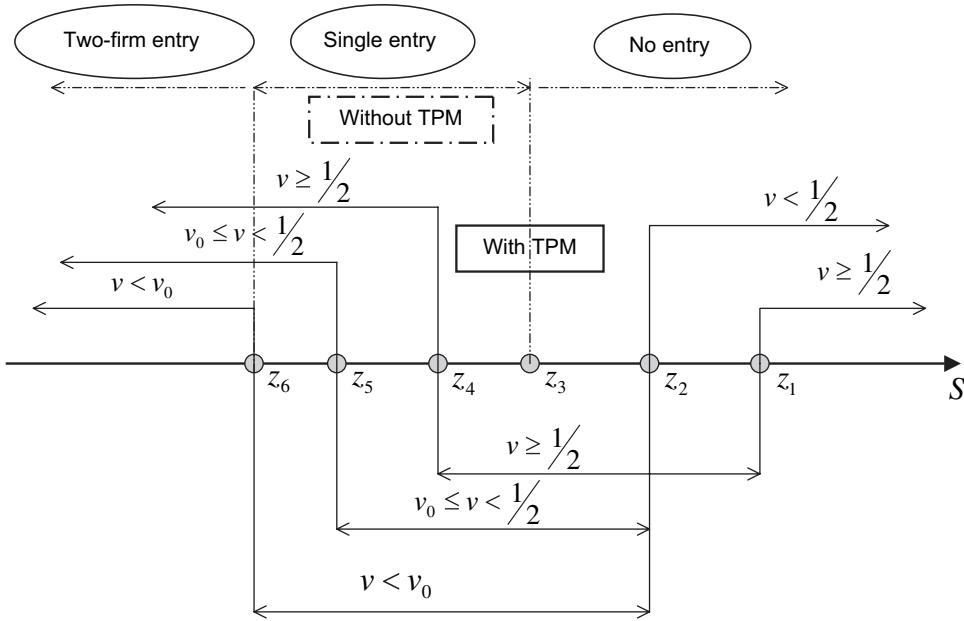
The authors next show that the entry-inducing results also hold for small discount factors. Recall that in the analysis, the critical points z_2, z_4, z_5 , and z_6 are derived under the assumption of a large discount factor. Now the authors will see how these critical value in Figure 1 will change when δ switches from above $9/17$ to below this critical level. First, if v is sufficiently small, a single-entry firm will choose to produce the monopoly profit (rather than v) in the first period and let entry occurs in the second period. In this case, $z_2 = z_3$. Otherwise, $z_2 > z_3$ still holds. Second, according to Proposition 2, two incumbents still collude at monopoly output in the first period even under low discount factor ($\delta < 9/17$) so long as $v > 2/3$, or $v \in [1/2, 2/3)$ but $\delta \geq \delta_1$. Thus, under these conditions, the same z_4 is still the critical value for two-firm entry. Finally, applying the same analysis as for $\delta \geq 9/17$ and making use of Proposition 2, it is easy to obtain corresponding critical levels of z_5 and z_6 under $\delta < 9/17$, of course with different magnitudes but the same ranking. Therefore, a figure similar to Figure 1 also applies to low discount factor and the result about TPM and first-period entry is still valid.

Welfare

This subsection shows that even though the TPM facilitates the incumbents' collusion, it

15. If $\delta > 27/34$, then $z_3 < z_4$. The result becomes even stronger.

FIGURE 1
Critical Levels for Entry



may nevertheless improve social welfare under certain conditions. Welfare is defined as the sum of discounted consumer surplus and the firms' discounted profits less entry costs.

When TPM induces entry, the first-period welfare generally increases. However, TPM facilitates collusion, which deters future entry and so reduces future competition. This has a negative impact on welfare. Thus, inducing entry is a necessary condition for TPM to increase welfare. It is not difficult to give a full welfare analysis, but it is quite tedious to do so because there are many cases to consider, depending on the magnitudes of s , δ , v , and other parameters. The purpose is to identify some cases where welfare is higher with TPM than without TPM and provide reasons behind them.

The authors first look at the case where $s \in [z_3, z_1)$. In this case, there is no first-period entry in the absence of TPM and so the expected (present value of) welfare is

$$W^{\text{NO}} = \frac{\delta}{1-\delta} \int_1^2 \left[\frac{a^2}{4} + \frac{1}{2} \left(a - \frac{a}{2} \right) \frac{a}{2} \right] da - \delta s_3$$

$$= \frac{7\delta}{8(1-\delta)} - \delta s_3.$$

With TPM, there is single entry for all δ and $v \geq 1/2$. The resulting welfare is

$$W^{\text{TPM}} = \frac{1}{4} + \frac{7\delta}{12(1-\delta)} + \frac{1}{2} \left(1 - \frac{1}{2} \right) \frac{1}{2}$$

$$+ \frac{\delta}{1-\delta} \int_1^2 \frac{1}{2} \left(a - \frac{a}{2} \right) \frac{a}{2} da - s$$

$$= \frac{3}{8} + \frac{7\delta}{8(1-\delta)} - s.$$

Thus, $W^{\text{TPM}} > W^{\text{NO}}$ if and only if $s - \delta s_3 < 3/8$. A sufficient condition for this to hold, for all $s \in [z_3, z_1)$ and $s_3 \geq 0$, is $\delta < 3/17$. Thus, when the discount factor is sufficiently small, TPM yields higher welfare than no TPM. Three factors together explain this welfare-improving result. First, s is large and so without TPM's protection in future periods, there will be no first-period entry. Second, δ is small so first period welfare is important. Third, v is large so TPM is effective in blocking future entry and inducing first period entry.

Next, the authors examine a case where TPM induces two-firm entry and facilitates collusion by considering a more general case where $s_1 > s_2$. Suppose $s_2 \in [z_6, z_4)$ and $s_1 > z_3$, but $s_1 < 2s_2$. It is easy to verify the existence of

a range of s_1 and of a range of s_2 satisfying these conditions. Under these conditions and referring to Figure 1, there is no first-period entry in the absence of TPM so the welfare is W^{NO} . However, with TPM and $v \geq 1/2$, there is two-firm entry in the first period and collusion blocks firm 3's entry under the following conditions: (i) $v \geq 2/3$, or (ii) $v \in [1/2, 2/3)$ and $\delta \geq \delta_1$. The resulting welfare is equal to

$$\begin{aligned} W_2^{TPM} &= 2 \left[\frac{1}{8} + \frac{7\delta}{24(1-\delta)} \right] + \frac{1}{8} \\ &\quad + \frac{\delta}{1-\delta} \int_1^2 \frac{1}{2} \left(a - \frac{a}{2} \right) a da - 2s_2 \\ &= \frac{3}{8} + \frac{7\delta}{8(1-\delta)} - 2s_2. \end{aligned}$$

The authors have $W_2^{TPM} > W^{NO}$ if and only if $2s_2 - \delta s_3 < 3/8$. A sufficient condition for this to hold, for all $s_2 \in [z_6, z_4)$ and $s_3 \geq 0$, is $\delta < 6/13$. Thus, although TPM facilitates the two incumbents' collusion, which blocks firm 3's entry, it induces first-period entry. Welfare is higher with TPM than without TPM because present welfare carries a large weight relative to future welfare ($\delta < 6/13$).

V. DISCUSSION AND CONCLUDING REMARKS

This article has shown that in the case of insufficient entry, the regulator's conditional commitment may outperform both a full commitment to blocking future entry and no commitment to entry restraint. However, a conditional commitment mechanism such as TPM may facilitate the incumbent firms' collusion. The results obtained are based on a number of assumptions. The authors briefly discuss their implications next.

Note that this article has focused on a large enough capacity constraint so that firms' output will not be effectively constrained. If the constraint is binding, how will it affect the results? It is known that in general the impact of a capacity constraint on collusion is ambiguous (see Tirole 1988, pp. 242–43). In the absence of potential entry, a binding capacity constraint affects collusion in two ways. On one hand, it reduces the one-time gain of defection because the defector's output would be lower than that without capacity constraint. That facilitates collusion. On the other hand,

the constraint may limit the degree of punishment for the defector because the nondefecting parties could not produce above capacity. This increases collusion incentives.

The focus is whether TPM facilitates collusion with or without capacity constraints, not whether collusion is more likely with or without capacity constraints. Because in both cases the incumbents always have the incentive to keep the first-period output low to avoid triggering the TPM, it follows that even with a capacity constraint the TPM serves to facilitate collusion in the sense that the first-period collusive output would be lower than that without TPM and potential entry.

This article makes two important supplementary assumptions about which equilibrium will be played in the dynamic game of the model. The first one is about how punishment will be carried out after defection from collusion is detected. In particular, the authors assume that the incumbent firms use grim strategy to support collusion and they collude at the lowest output level which is sustainable. The grim strategy, which is most commonly assumed in the literature, is associated with very severe punishment and therefore supports high levels of collusion. In comparison, other strategies with more (less) severe punishment will support higher (lower) levels of collusion. Though different punishment schemes support different levels of collusion, the main results obtained here based on grim strategy will not change. It is clear that (under the second supplementary assumption to be discussed later) regardless of the punishment scheme, in the absence of TPM, future entry by firm 3 always destroys the incumbents' collusion in the first period. TPM helps restore collusion because it allows the incumbents to block future entry. As a result, the dynamic game will be similar to that assumed herein because the qualitative effects of TPM do not depend on how severe the punishment scheme is, which affects only the critical levels of v for various degrees of collusion.

The second supplementary assumption is about the colluding firms' response to new entry. The authors assume that entry is not accommodated. This yields more severe punishment for defection than if entry is accommodated, that is, the incumbents collude with the entrant. As a result, in the case of entry accommodation, for a given level of collusion output, the one-time gain of defection

from the collusion between the two incumbents in the first period is the same as that analyzed in the present model, but the future loss is smaller. Therefore, the first-period outcome is less collusive with entry accommodation than without. However, so long as entry occurs, the incumbents will get less profit. Hence, even with entry accommodation, the incumbents still have an incentive to lower the first-period output to deter future entry. That is, the TPM can still facilitate collusion in this case, even though at different levels.

One potential direction for further research would be to check the robustness of these results about collusive behavior under alternative assumptions. Another direction would be to design an optimal TPM.

APPENDIX

Proof of Lemma 1

To economize space, let UCO stand for the unconstrained cheating output and CCO stand for the constrained cheating output.

First, note that in Case A, if the defector chooses UCO, the total output $X_1(x_1^c) < 2x_1^m$. The defector gets the first period profit Π^d . Thus, in Case A, choosing UCO is the optimal strategy.

Let us now turn to the most interesting case, that is, Case B. Denoting $\Delta\pi(x_1^c) \equiv \Pi^{um} - \Pi^{com}$, then

$$\frac{\partial \Delta\pi(x_1^c)}{\partial x_1^c} = \frac{1}{8}(2 - \nu + x_1^c) > 0, \quad \frac{\partial^2 \Delta\pi(x_1^c)}{\partial x_1^{c2}} = \frac{1}{8} > 0.$$

The monotonicity comes from the fact that as x_1^c increases, while both Π^{um} and Π^{com} decrease, the constraint becomes more restrictive, reducing Π^{com} much more.

Given any ν , one only has to consider collusive output $x_1^c \leq \nu$ because if $x_1^c > \nu$, entry always occurs and collusion can never be sustained. In other words, if $x_1^c > \nu$, it is optimal for the defector to choose UCO. The authors examine two types of ν . First, ν is relatively large such that $\nu \in (5/8, 2x_1^m)$. In this case, choosing UCO will induce entry if $x_1^c \in (2(2\nu - 1), \nu]$, but not if $x_1^c \in (1/2, 2(2\nu - 1)]$. Second, ν is relatively small such that $\nu \in (x_1^m, 5/8]$. In this case, choosing UCO always induces entry.

In the case of a relatively large ν , because $\Delta\pi(2(2\nu - 1)) < 0$ and $\Delta\pi$ strictly increases in x_1^c , we can define $z_0 \in (2(2\nu - 1), \nu]$ as the solution to $\Delta\pi(z) = 0$ if $\Delta\pi(\nu) \geq 0$ but $z_0 = \nu$ if $\Delta\pi(\nu) < 0$. Hence, for all $x_1^c \in (2(2\nu - 1), z_0]$, $\Delta\pi(x_1^c) < 0$ and the defector chooses CCO, but for all $x_1^c \in [z_0, \nu]$, $\Delta\pi(x_1^c) \geq 0$ and the defector chooses UCO. Because choosing UCO does not induce entry if $x_1^c \in (x_1^m, 2(2\nu - 1)]$, the defector chooses UCO. However, if a and δ are sufficiently large, the following condition will hold:

$$\Delta\pi(\nu) = \frac{(2 - 3\nu)^2}{16} - \frac{\delta}{1 - \delta} \frac{7a^2}{144} < 0,$$

implying that choosing CCO is the preferred strategy.

In the case of relatively small ν , supposing $\Delta\pi(x_1^m) \leq 0$, one can define $z_1 \in [x_1^m, \nu]$ as the solution to $\Delta\pi(z_1) = 0$ if $\Delta\pi(\nu) \geq 0$ but $z_1 = \nu$ if $\Delta\pi(\nu) < 0$. Hence, for all $x_1^c \in [x_1^m, z_1)$, $\Delta\pi(x_1^c) < 0$ and the defector chooses CCO. For all $x_1^c \in [z_1, \nu]$, $\Delta\pi(x_1^c) \geq 0$ and the defector chooses UCO. If a and δ are sufficiently large, the following condition will hold:

$$\Delta\pi(x_1^m) = \frac{9}{64} - \frac{(1 - \nu)(4\nu - 1)}{4} - \frac{\delta}{1 - \delta} \frac{7a^2}{144} < 0,$$

implying that choosing CCO is preferred. If in contrast, $\Delta\pi(x_1^m) > 0$, then $\Delta\pi(x_1^c) > 0$ for all x_1^c and choosing UCO dominates choosing CCO. In this case the authors define $z_1 = x_1^m$.

Finally we turn to Case C where $\nu < x_1^m$. The authors show that the incumbent firms may collude at an output below x_1^m . However, if they collude only within the normal range $[x_1^m, 2x_1^m)$, it is obvious that the defector always chooses UCO to get a large first-period profit because entry always occurs when ν is small. \square

Proof of Proposition 2

The authors analyze collusion in the three cases first, and then summarize them to get the three results stated in the proposition. Denote the first period's collusion output as x_1^T (T stands for TPM).

In Case A, because ν is sufficiently large such that entry never occurs, the situation analyzed in the text prior to the proposition is the same as this case. Thus, $x_1^T = x_1^*$. The collusion is restored as in the case of no entry.

In Case B, it is easily seen that if they can collude at a lower output level, the firms never collude at $x_1^c \in (\nu, 2x_1^m)$ because collusion in this range always leads to entry. So the authors focus on $x_1^c \in [x_1^m, \nu]$.

Recall from Lemma 2 that $x_1^* = x_1^m$ for $\delta_1 \leq \delta < 9/17$. Here one can apply the same argument used in the analysis of Proposition 1 that compared to the case without entry, defection either leads to lower one-time gain or higher future loss, so collusion at x_1^m is also sustainable with potential entry regulated by TPM. Thus, $x_1^T = x_1^*$.

Suppose $\delta < \delta_1$ and recall from Lemma 2 that $x_1^* = x_1^o \in (x_1^m, 2x_1^m)$. The authors first examine the case $x_1^o \leq \nu$. They want to show that $x_1^T \leq x_1^o$ and the strict inequality holds in some cases. First, using the same argument, collusion at x_1^o is sustainable. Second, they examine whether collusion at an output level x_1^c strictly lower than x_1^o is sustainable. The authors are going to make extensive use of results in Lemma 1. First, suppose $\nu \in (5/8, 2x_1^m)$. Examine collusion at x_1^c . The authors know from Lemma 1(i) that the only case when the defector choosing UCO does not induce entry is when $x_1^c \in (x_1^m, 2(2\nu - 1))$. In this case, collusion at output lower than x_1^o is not sustainable because (a) the one-time gain from defection and future loss are the same as those when there is no entry and (b) x_1^o is already the lowest sustainable level. Combining the two conditions $\nu \in (5/8, 2x_1^m)$ and $x_1^c \in (x_1^m, 2(2\nu - 1))$ with $\delta < \delta_1$ produces the conditions as stated in Proposition 2(i)(c). In other cases, either the one-time gain from defection is strictly lower (choosing CCO) than that without entry or the future loss is larger (choosing UCO) than that without entry. Hence, collusion at an output level strictly lower than strategy II but arbitrarily close to x_1^o is sustainable. Second, suppose $\nu \in (x_1^m, 5/8)$. From Lemma 1(ii) the authors know that

there is no chance that the defector can choose UCO without inducing entry. Hence, as analyzed, collusion at an output level strictly lower than but arbitrary close to x_1^o is sustainable.

Suppose now $v < x_1^o$ and consider $x_1^c \leq v$. Because in Case B one need only to consider collusion within $[x_1^m, 2x_1^c]$, the authors define $x_1^{TO} \equiv \max\{x_1^m, x_1^c\}$. Let $DI_L^c(x_1^c)$ denote the DI corresponding to the case where any first period defection leads to entry. Since $DI_L^c(x_1^c) - DI_L(x_1^c) = -7a^2\delta/144(1 - \delta) < 0$, the authors know that $x_1^c < x_1^{TO}$. Note also if collusion at output level strictly below v is sustainable, then collusion at v is also sustainable. Therefore, the analysis shows that if $v < x_1^{TO}$, there will be no collusion in the first period, but if $v \geq x_1^{TO}$, then collusion at an output level strictly lower than x_1^o is sustainable (say, at v).

Finally, the authors turn to Case C. Because $v < x_1^m$, the incumbent never colludes at $x_1^c < v$. The authors just need to see if collusion at v is beneficial and sustainable. Following the analysis in the last paragraph, they note that collusion at v is sustainable if and only if $v \geq x_1^c$. Collusion at v is beneficial if and only if $(\bar{\Pi}_c - \Pi_n) > 0$, or

$$\frac{1}{18}(9v - 9v^2 - 2) + \frac{\delta}{1 - \delta} \left(\pi^o - \frac{a^2}{16} \right) > 0.$$

The analysis shows the first-period outcomes under various circumstances. The authors just need to summarize them to get the proposition. This completes the proof. \square

REFERENCES

Abreu, D. "Extremal Equilibria of Oligopolistic Supergames." *Journal of Economic Theory*, 39, 1986, 191-225.

Anton, J. J., and D. A. Yao. "Second-Sourcing and the Experience Curve: Price Competition in Defense Procurement." *RAND Journal of Economics*, 18, 1987, 57-76.

Baron, D. P. "Design of Regulatory Mechanisms and Institutions." In *Handbook of Industrial Organization*, vol 2, edited by R. Schmalensee and R. D. Willig. Amsterdam: North-Holland Elsevier Science Publishers, 1989, pp. 1347-447.

Besankol, D., and D. F. Spulber. "Sequential-Equilibrium Investment by Regulated Firms." *RAND Journal of Economics*, 23, 1992, 153-70.

Chatterjee, S., and R. W. Cooper. "Entry and Exit, Product Variety and the Business Cycle." *NBER Working Paper no. 4562*, 1993.

Chevalier, J. A., and D. S. Scharfstein. "Capital-Market Imperfections and Countercyclical Markups: Theory and Evidence." *American Economic Review*, 86(4), 1996, 703-25.

Davison, C., and R. Deneckere. "Excess Capacity and Collusion." *International Economic Review*, 31, 1990, 521-42.

Friedman, J. W. "A Non-cooperative Equilibrium for Supergames." *Review of Economic Studies*, 38, 1971, 1-12.

Gilbert, R. J. "Mobility Barriers and the Value of Incumbency." In *Handbook of Industrial Organization*, vol 1 edited by R. Schmalensee and R. D. Willig. Amsterdam: North-Holland Elsevier Science Publishers, 1989, pp. 475-535.

Gilbert, R. J., and X. Vives. "Entry Deterrence and the Free Rider Problem." *Review of Economic Studies*, 53, 1986, 71-83.

Green, E. J., and R. H. Porter. "Noncooperative Collusion under Imperfect Price Information." *Econometrica*, 52, 1984, 87-100.

Harrington, J. E. Jr. "Collusion and Predation Under (Almost) Free Entry." *International Journal of Industrial Organization*, 7, 1989, 381-401.

Hesse, M. O. "A New Era in Energy Regulation." *Public Utilities Fortnightly*, 123, 1989, 18-21.

Hong Kong Centre for Economic Research. "Competition in Container Terminals." *HKCER Letters*, 16, 1992.

Kim, Jaehong. "Inefficiency of Subgame Optimal Entry Regulation." *RAND Journal of Economics*, 28, 1997, 25-36.

Krishna, K., and J. Morgan. "Implementing Results-Oriented Trade Policies: The Case of the US-Japanese Auto Parts Dispute." *European Economic Review*, 42(8), 1998, 1443-67.

Kydland, F. E., and E. C. Prescott. "Rules Rather than Discretion: The Inconsistency of Optimal Plans." *Journal of Political Economy*, 85, 1977, 473-91.

Laffont, J.-J., and J. Tirole. *A Theory of Incentives in Procurement and Regulation*. (Cambridge, MA: MIT Press, 1993).

Levy, Brian, and Pablo T. Spiller. *Regulations, Institutions, and Commitment*. (New York: Cambridge University Press, 1996).

PDB (Port Development Board). *Annual Report 1992* (Hong Kong, 1992).

Rotemberg, J. J., and G. Saloner. "Tariffs vs. Quotas with Implicit Collusion." *Canadian Journal of Economics*, 22, 1989, 237-44.

Tirole, J. *The Theory of Industrial Organization*. (Cambridge, MA: MIT Press, 1988).

Williamson, O. E. *Markets and Hierarchies: Analysis and Antitrust Implications*. (New York: Free Press, 1975).

———. "Credible Commitments: Using Hostages to Support Exchange." *American Economic Review Proceedings*, 83, 1984, 519-40.