A Cost-side Analysis on Collusive Sustainability

Luca Lambertini
Department of Economics
University of Bologna
Strada Maggiore 45
I-40125 Bologna, Italy
lamberti@spbo.unibo.it

Dan Sasaki
Department of Economics
University of Exeter
Exeter, Devon EX4 4PU
United Kingdom
d.sasaki@exeter.ac.uk

September 1999

Acknowledgements : We are grateful to George Norman, Sougata Poddar, Christian Schultz and Xavier Wauthy for insightful discussions, without which this research would not have been possible. The authors assume the sole responsibility for any remaining defects.

A Cost-side Analysis on Collusive Sustainability

Abstract : In an oligopoly supergame, firms' actions in prices and quantities are subject to non-negativity constraints. These constraints can obstruct the practicability of optimal punishment (*à la* Abreu (1986), Lambson (1987), and Häckner (1996)) in sustaining tacit collusion. Noting that the prospect of single-period optimal punishment depends indispensably upon firms' ability to charge prices *strictly below marginal costs* (loss-making pricing), under the presence of positive price constraints, marginal costs can serve as a "fudge" to materialise single-period optimal punishment. In this paper we characterise the effects of profit-cost ratios (or mark-ups) on the sustainability of tacit collusion, in light of optimal punishment.

Keywords : Marginal costs, mark-up, positive price constraint, positive quantity constraint, product differentiation, substitutability.

JEL classification : L13, D43, C72.

1 Introduction

GAME-THEORETIC OLIGOPOLY LITERATURE has been widely receptive of the "simplifying assumption" to suppress supply costs and thereby to equate firms' profits with their revenues. Undoubtedly, this practice has been promoted by the innocuous intention to simplify the model as much as possible. This simplification, however, is accompanied by a dear sacrifice: with zero marginal costs, *negative prices* and *prices below marginal costs* are *synonymous*. In reality, on the other hand, it is conceivable that negative prices would encounter implementational difficulties, whilst prices below marginal costs would be free of any such complication insofar as their literal figures are non-negative.

Clearly, this distinction never materialises when the oligopoly game is *static*. For, both negative prices and prices below marginal costs are individually irrational, and thus can be dismissed from consideration. Even when the oligopoly game is *repeated* (an oligopoly supergame), insofar as firms apply trigger strategies with Nash reversion *à la* Friedman (1971), there is little need to consider any strategies or actions which are individually irrational either temporarily or intertemporally. In fact, the prospect of temporarily charging a price strictly below marginal costs received little theoretical attention until Abreu (1986, 1988) and Abreu, Pearce and Stacchetti (1986) pioneered the concept of *optimal punishment*.

In contrast with its substantial impact on game theory, economic applications of optimal punishment in more specialised fields, including industrial organisation, have been surprisingly few and far between. Among these few pathbreakers is Lambson (1987). Using a Bertrand duopoly supergame with two firms supplying *perfect substitute products*, Lambson constructs an optimal single-period penal code which is *temporarily individually irrational* yet *intertemporally precisely individually rational*.[1] However, the minimum discount factor required in sustaining price collusion in Lambson (1987) is the same 1/2 as that in Friedman (1971). This, as we discover in this paper, is due to Lambson's specific assumption that the two firms' products are *perfect substitutes*.

Literature abounds on the effects of product differentiation on the sustainability of

[1]Lambson calls this the "security level", i.e. the discounted flow of profits at which the individual rationality constraint precisely binds. In his more recent contributions (Lambson, 1994; 1995) it is shown that, if firms' *a priori* symmetry is waived, optimal punishments may no longer hit the security level. Specifically, if firms differ in size, the security level punishment is operative for the large firms but not for the smaller ones.

tacit collusion either in prices or in quantities (Deneckere, 1983; Chang, 1991, 1992; Rothschild, 1992; Ross, 1992; Friedman and Thisse, 1993; Häckner, 1994, 1995, 1996; Lambertini, 1997; *inter alia*). Except for Häckner (1996), however, all this literature uses the traditional Friedman (1971) formulation of folk theorem.

In this paper, we analyse the effects of marginal costs on sustainability of tacit collusion, in light of optimal punishment under both non-negative price and non-negative quantity constraints. Section 2 provides a basic general model. In sections 3 and 4 we present a linear duopoly model (as in Singh and Vives, 1984) to develop intuition by deriving subgame perfect equilibria explicitly. In linear Cournot duopoly supergame presented in section 3 the price floor is never a binding concern, whilst in linear Bertrand duopoly supergame presented in section 4 the price floor can bind if firms are supplying *close substitutes*. Section 5 briefly illustrates when the *duality* between price competition and quantity competition holds and when it fails, in terms of the magnitude of marginal costs relative to demand prices. Section 6 summarises the possible incidence of various industrial taxation and subsidisation on collusive sustainability, in light of their effects on private firms' cost structure. Section 7 concludes the paper.

2 General set-up

2.1 The model

We consider the following supergame. Two *a priori* identical firms operate in the market over the time horizon $t = 1, 2, 3, \cdots$, selling (possibly imperfect) substitute products. At each time t, firm i chooses its action α_i, which materialises either in a price p_i or in a quantity q_i in most of the existing theoretical models, although our general set-up does not preclude other action spaces insofar as they are subsets of well-defined metric spaces. We adopt the following assumptions (cf. Abreu, 1986, p. 195):

A1 : The discount factor $\delta \in (0, 1)$ is constant over time and common to both firms.

A2 : The unit production cost is constant and equal to k.

Therefore, the resulting per-period profit function for firm i is $\pi_i(\alpha_i, \alpha_j) = (p_i - k)q_i$, where $\{i, j\} = \{1, 2\}$.

A3 : Each inverse demand function $p_i(q_i\,, q_j)$ is strictly monotone and continuous in q_i given any q_j .

A4 : $p_i(0\,,0) > 0$ and, for any q_j there exists a $z(q_j) > 0$ such that $p_i(z(q_j)\,,q_j) = 0$.

A5 : $q_i \in [0, \Omega(\delta)]$, where $\Omega(\delta)$ is such that producing $q_i > \Omega(\delta)$ would violate the individual rationality constraint to the continuation of the game.

This states that it will never be in the interest of firm i to produce any output larger than $\Omega(\delta)$, for any given $\delta \in (0, 1)$.

A6 : $\arg\max\{\pi_i(\alpha_i, \alpha_j) \mid \alpha_i \in [0, \Omega(\delta)]\} \equiv \alpha_i^*[\alpha_j]$ is a singleton. Firm i's profit $\pi_i(\alpha_i, \alpha_j)$ decreases unimodally as the distance between α_i and $\alpha_i^*[\alpha_j]$ increases.

A7 : The one-shot game in the space $\{\alpha_1, \alpha_2\}$ has a unique equilibrium in pure strategies, (α_1^N, α_2^N) .

We assume hereinafter that the unique static Nash equilibrium is *symmetric*, i.e., $\alpha_1^N = \alpha_2^N = \alpha^N$.

Throughout this paper, our focus is exclusively to discuss the sustainability of a path that consists of simple repetition of the same symmetric action profile (α^C, α^C) , referred to as **collusive actions** or simply **collusion** unless contextually confusing. Our premise is that this collusive profile entails profits that are *strictly above* the profits from repetition of the static Nash equilibrium (α^N, α^N) .

In general, for any prescribed profile (α^C, α^C) there exists a unique **critical discount factor** $\delta^*[\alpha^C]$ such that the profile is sustainable with subgame perfection if and only if the actual discount factor δ is equal to or higher than $\delta^*[\alpha^C]$. The general derivation of this critical discount factor has been established by Abreu (1986; 1988) and Abreu, Pearce and Stacchetti (1986). However, their symmetric optimal punishment schemes can be affected by the *positivity constraints* imposed on firms' actions (either quantities or prices). This shall be closely inspected in the following 2.2 and 2.3.

2.2 Single-period symmetric optimal punishment

Firms initially follow a prescribed collusive path, defined as a symmetric outcome yielding strictly higher profits than the static Nash equilibrium outcome, until any deviation is detected. If a deviation is detected in period t, then in the next period $t+1$, firms switch to the punishment phase where *both* firms adopt the punishment action $\alpha^{[1]}$ irrespective of which firm is punishing the other. If both firms follow the prescribed penal code at $t+1$, then they revert to the initial collusive path from $t+2$ onwards. Otherwise, the punishment phase continues until the penal code is adopted by both firms at the same time. Abreu (1986, Lemma 17, p. 204) proves that this symmetric penal code, when a^P satisfies the system of equations

$$\pi_i(\alpha_i^*[\alpha^C], \alpha^C) - \pi_i(\alpha^C, \alpha^C) = \delta^*[\alpha^C]\left(\pi_i(\alpha^C, \alpha^C) - \pi_i(\alpha^{[1]}, \alpha^{[1]})\right); \quad (1)$$

$$\pi_i(\alpha_i^*[\alpha^{[1]}], \alpha^{[1]}) - \pi_i(\alpha^{[1]}, \alpha^{[1]}) = \delta^*[\alpha^C]\left(\pi_i(\alpha^C, \alpha^C) - \pi_i(\alpha^{[1]}, \alpha^{[1]})\right); \quad (2)$$

is *optimal* in that it requires a lower discount factor $\delta^*[\alpha^C]$ to sustain collusion than any other punishment rule.

2.3 Constrained multi-period symmetric optimal punishment

Although system (1)-(2) consists of two equations with two unknowns, due to exogenous constraints imposed on action sets, a solution $\left\{\alpha^{[1]}, \delta^*[\alpha^C]\right\}$ may not always exist. When no solution exists, this reflects the situation where even the toughest single-period penal action admissible under the exogenous constraints is still not severe enough to make the firms indifferent between deviating *from the penal code* and complying. To fill this "gap" we consider the following modification. We adopt the toughest admissible instantaneous penal action α^P, where the exogenous constraints are already binding. Then we also impose a small amount of additional punishment *in the period after*. This serves to increase the toughness of the punishment (that is, reduce the incentives for the initial deviation) on one hand, and to increase the deviation incentive *from the punishment* on the other hand, whereby reducing the aforementioned gap. As we increase the amount of punishment in the second penal period, more of the gap is filled. If the gap is not completely filled by the two-period punishment, i.e., if both the first and the second penal periods carry the toughest penal action α^P and firms still have *strict* incentives to comply with the penal code *in the first penal period*, then we load additional punishment in the *third* penal period. The protraction of the penal phase continues until the gap is completely filled.

Thereby in general, when our penal code is ℓ periods long, the severest penal action α^P admissible under the exogenous constraints is inflicted during the first $\ell - 1$ periods, and then the final ℓ-th period of the penal phase picks up the remainder of the total prescribed amount of punishment. (See Appendix for auxiliary explanations.) More formally:

- At $t = 1$, both firms set $\alpha_i = \alpha^C$.
- At t ($t \geq 2$), both firms set $\alpha_i = \alpha^C$ if, at $t-1$, $\alpha_1 = \alpha_2 = \alpha^C$, i.e., if firms have previously adhered to the collusive prescription. Otherwise, firms switch to an ℓ-period punishment phase, during the first $\ell - 1$ periods of which *both* firms adopt the penal action α^P, in the final ℓ-th period they adopt the auxiliary penal action $\alpha^{P[\ell]}$. If any deviation from the penal code is detected during the punishment phase, then the whole ℓ periods of the punishment phase restart all over again; otherwise if both firms have adhered to the penal code throughout the entire ℓ periods, then they revert to the initial collusive path thenceforth.

The incentive compatibility of this penal code can be expressed by the system of simultaneous inequalities:

$$\begin{aligned}\pi_i(\alpha_i^*[\alpha^C], \alpha^C) - \pi_i(\alpha^C, \alpha^C) \;\leq\; & \sum_{h=1}^{\ell-1} \delta^h \left(\pi_i(\alpha^C, \alpha^C) - \pi_i(\alpha^P, \alpha^P)\right) \\ & + \delta^\ell \left(\pi_i(\alpha^C, \alpha^C) - \pi_i(\alpha^{P[\ell]}, \alpha^{P[\ell]})\right); \end{aligned} \tag{3}$$

$$\begin{aligned}\pi_i(\alpha_i^*[\alpha^P], \alpha^P) - \pi_i(\alpha^P, \alpha^P) \;\leq\; & \delta^{\ell-\tau} \left(\pi_i(\alpha^{P[\ell]}, \alpha^{P[\ell]}) - \pi_i(\alpha^P, \alpha^P)\right) \\ & + \sum_{h=\ell-\tau+1}^{\ell-1} \delta^h \left(\pi_i(\alpha^C, \alpha^C) - \pi_i(\alpha^P, \alpha^P)\right) \\ \tau = 1, \cdots, \ell - 1 \qquad & + \delta^\ell \left(\pi_i(\alpha^C, \alpha^C) - \pi_i(\alpha^{P[\ell]}, \alpha^{P[\ell]})\right) \end{aligned} \tag{4}$$

$$\begin{aligned}\pi_i(\alpha_i^*[\alpha^{[\ell]}], \alpha^{[\ell]}) - \pi_i(\alpha^{[\ell]}, \alpha^{[\ell]}) \;\leq\; & \sum_{h=1}^{\ell-1} \delta^h \left(\pi_i(\alpha^C, \alpha^C) - \pi_i(\alpha^P, \alpha^P)\right) \\ & + \delta^\ell \left(\pi_i(\alpha^C, \alpha^C) - \pi_i(\alpha^{P[\ell]}, \alpha^{P[\ell]})\right); \end{aligned} \tag{5}$$

of which (3) represents the incentive not to deviate from the collusive path, (4) represents the incentive not to deviate from the penal code at the τ-th period of the punishment phase ($\tau = 1, \cdots, \ell - 1$), and finally (5) the incentive not to deviate from the final ℓ-th period of the penal code. From the expression in the right-hand sides of these inequalities it emerges that among these $\ell + 1$ inequalities only the first two, i.e., (3), and (4) with $\tau = 1$, can bind whilst the other $\ell - 1$ inequalities ($\tau = 2, \cdots, \ell - 1$ of (4), as well as (5)) are slack.

In addition, the *intertemporal individual rationality condition*

$$\sum_{h=1}^{\ell-1} \delta^h \pi_i(\alpha^P, \alpha^P) \ + \ \delta^\ell \pi_i(\alpha^{[\ell]}, \alpha^{[\ell]}) + \sum_{h=\ell+1}^{\infty} \delta^h \pi_i(\alpha^C, \alpha^C) \ \geq \ 0 \tag{6}$$

must always be satisfied. If constraint (6) were violated, firms would find it preferable to cease all the production activities permanently, because the harshness of the punishment would outweigh the discounted value of the reinstalment of collusive profits.

To summarise, *when exogenous constraints bind*, the **constrained critical discount factor** $\delta^{**}[\alpha^C]$ is determined through the system of two equations

$$\begin{aligned} \pi_i(\alpha_i^*[\alpha^C], \alpha^C) - \pi_i(\alpha^C, \alpha^C) \ &= \ \sum_{h=1}^{\ell-1} (\delta^{**}[\alpha^C])^h \left(\pi_i(\alpha^C, \alpha^C) - \pi_i(\alpha^P, \alpha^P)\right) \\ &\quad + (\delta^{**}[\alpha^C])^\ell \left(\pi_i(\alpha^C, \alpha^C) - \pi_i(\alpha^{P[\ell]}, \alpha^{P[\ell]})\right); \end{aligned} \tag{7}$$

$$\begin{aligned} \pi_i(\alpha_i^*[\alpha^P], \alpha^P) - \pi_i(\alpha^P, \alpha^P) \ &\leq \ (\delta^{**}[\alpha^C])^{\ell-1} \left(\pi_i(\alpha^{P[\ell]}, \alpha^{P[\ell]}) - \pi_i(\alpha^P, \alpha^P)\right) \\ &\quad + (\delta^{**}[\alpha^C])^\ell \left(\pi_i(\alpha^C, \alpha^C) - \pi_i(\alpha^{P[\ell]}, \alpha^{P[\ell]})\right) \end{aligned} \tag{8}$$

with two unknowns $\delta^{**}[\alpha^C]$ and $\alpha^{[\ell]}$ (note that the principal penal action α^P is no longer an unknown because it is now determined by the constraints not by incentives), subject to individual rationality (6).

This observation immediately implies the following.

Theorem : When the set of admissible actions enlarges (resp., diminishes), the required duration of punishment ℓ shortens (resp., lengthens) and the constrained critical discount factor $\delta^{**}[\alpha^C]$ decreases (resp., increases), at least weakly.

2.4 Cost effects

Having constrained penal actions characterised in general abstract terms, now we return to our main intention in this paper, that is, the effects of *marginal costs* on the sustainability of collusion.

The basic premise we accept is that firms are allowed to charge prices strictly below marginal costs, but not literally negative prices. This automatically implies that, when marginal costs are negligible, no loss-making actions are available, which restricts the set of implementable penal codes. In our notation from 2.3, this is expressed as $\pi_i(\alpha^P, \alpha^P) = 0$;

in words, the toughest admissible penal profit is equal to zero. Clearly from the right-hand side of (7), this restriction serves to prolong the required duration of punishment on one hand, and to raise the required discount factor $\delta^{**}[\alpha^C]$ on the other hand. As marginal costs increase, the amount of losses that can be inflicted in one period also increases. This makes $\pi_i(\alpha^P, \alpha^P)$ substantially negative, serving to keep the duration of punishment short and also to keep the required discount factor $\delta^{**}[\alpha^C]$ low.

Viewed differently, appropriate rescaling of prices and costs enables us to associate relatively high (resp., low) marginal costs with relatively low (resp., high) *demand prices in excess of marginal costs*. A high-cost industry can be reinterpreted as a low-markup industry, where profit opportunities from deviation are scarce. In such an industry, "collusion" (albeit not enormously profitable) is easy to sustain.

These observations can be summarised as follows, which is indeed a corollary to Theorem in 2.3.

Corollary : Under the positive price constraint, if the demand structure is unchanged in terms of *prices in excess of marginal costs*, then adding a positive constant to marginal costs serves to enlarge

- the set of sustainable collusive outcomes given δ ;
- the set of admissible δ in sustaining a given collusive outcome.

3 Linear Cournot duopoly supergames

To be concrete, in this section we follow much of existing literature by adopting the linear duopoly model *à la* Singh and Vives (1984). Duopoly interaction unfolds over $t = 1, 2, 3, \cdots$. In every period, each firm i ($\{i, j\} = \{1, 2\}$) chooses a positive quantity $q_i \geq 0$ and the market price realises according to the inverse demand function:

$$p_i \;=\; \max\{1 - q_i - \gamma q_j\,, 0\} \qquad\qquad q_1 \geq 0\,,\; q_2 \geq 0 \tag{9}$$

in which $\gamma \in (-1, 1]$ measures the degree of substitutability. Notice that the normalisation of the reservation price to one implies $k \in [0, 1)$. Throughout the game, the discount factor $\delta \in [0, 1)$ is common to both firms.

Conforming to our previous notation, we define profit functions and quantity-reaction functions as

$$\begin{aligned} \pi_i(q_i\,, q_j) &= (1 - q_i - \gamma q_j - k) q_i \\ q_i^*[q_j] &= \max\left\{\frac{1-k-\gamma q_j}{2}\,,\,0\right\} \end{aligned} \qquad \{i,j\} = \{1,2\}\,. \tag{10}$$

Obviously the most profitable outcome, if sustainable, is for the two firms to split the *monopoly quantity* evenly by setting

$$q_1 = q_2 = q^M = \frac{1-k}{2(1+\gamma)}\,. \tag{11}$$

3.1 Single-period symmetric optimal punishment

Any symmetric profile $q_1 = q_2 = q^C$ is sustainable by means of single-period optimal punishment (see Abreu, 1986, 1988 ; Abreu, Pearce and Stacchetti, 1996 ; Häckner, 1996) if and only if the discount factor δ is greater than or equal to the threshold $\delta^*[q^C]$ which is determined by

$$\pi_i(q_i^*[q^C], q^C) - \pi_i(q^C, q^C) = \delta^*[q^C]\left(\pi_i(q^C, q^C) - \pi_i(q^{[1]}, q^{[1]})\right); \tag{12}$$

$$\pi_i(q_i^*[q^{[1]}], q^{[1]}) - \pi_i(q^{[1]}, q^{[1]}) = \delta^*[q^C]\left(\pi_i(q^C, q^C) - \pi_i(q^{[1]}, q^{[1]})\right). \tag{13}$$

In particular, the most profitable collusion $q^C = q^M$ (see notation (11)) is associated with the solution

$$q^{[1]} = \frac{(1-k)(2+3\gamma)}{2(1+\gamma)(2+\gamma)}\,; \qquad \delta_C^*[q^M] = \frac{(2+\gamma)^2}{16(1+\gamma)} \tag{14}$$

which meets the positive quantity requirement within, and only within, the range $\gamma \in \left[-\frac{2}{3}\,, 1\right]$. Note also that this penal quantity $q^{[1]}$ always entails the price strictly above the marginal cost k.[2]

Over the range $\gamma \in \left(-1\,, -\frac{2}{3}\right)$, the most profitable collusion q^C sustainable by single-period punishment *subject to the positive quantity constraint* is derived by plugging the binding constraint $q^{[1]} = 0$ into the system of simultaneous equations (12)-(13). It is clear that this will not drive the penal price down below the marginal cost k either.

[2] Easy although tedious algebra reveals the penal price $p_i(q^{[1]}, q^{[1]})$ to be

$$p_i(q^{[1]}, q^{[1]}) = 1 - \frac{(1-k)(2+3\gamma)}{2(2+\gamma)}$$

which is non-negative all over the relevant range, i.e., $\gamma \in \left[-\frac{2}{3}\,, 1\right]$.

Economic intuition : When the two firms' products are mutually substitutes (perceived from the demand side, likewise heretofore), punishment is inflicted by *expanding* each firms' quantity. This is reflected upon the fact that the penal quantity $q^{[1]}$ increases monotonically in the substitutability parameter γ. This causes the resulting penal price to *decrease* in γ, However, even when the two firms sell perfect substitutes ($\gamma = 1$), the penal price is still strictly above the marginal cost k. On the other hand, when the two firms supply complementary products, punishment is inflicted by *contracting* their quantities. This serves to raise the price, hence the price floor is never a binding concern.

3.2 Possibly multi-period symmetric optimal punishment

When $\gamma \in \left(-1\,,-\dfrac{2}{3}\right)$, the monopoly-level collusion (11) may require *multi-period punishment* defined in 2.3. When the duration of punishment is ℓ periods ($\ell = 2, 3, \cdots$), the penal quantity during the first $\ell - 1$ periods is constrained at $q^P = 0$ whilst the penal quantity in the final ℓ-th period $q^{[\ell]}$ and the resulting minimum admissible discount factor $\delta^{**}[q^M]$ are determined by the system of simultaneous equations:

$$\begin{aligned}\pi_i(q_i^*[q^M], q^M) - \pi_i(q^M, q^M) &= \frac{\delta^{**}[q^M] - (\delta^{**}[q^M])^\ell}{1 - \delta^{**}[q^M]}\Big(\pi_i(q^M, q^M) - \pi_i(0,0)\Big) \\ &\quad + (\delta^{**}[q^M])^\ell\Big(\pi_i(q^M, q^M) - \pi_i(q^{[\ell]}, q^{[\ell]})\Big)\,; \qquad (15)\end{aligned}$$

$$\begin{aligned}\pi_i(q_i^*[0], 0) - \pi_i(0,0) &= (\delta^{**}[q^M])^{\ell-1}\Big(\pi_i(q^{[\ell]}, q^{[\ell]}) - \pi_i(0,0)\Big) \\ &\quad + (\delta^{**}[q^M])^\ell\Big(\pi_i(q^M, q^M) - \pi_i(q^{[\ell]}, q^{[\ell]})\Big) \qquad (16)\end{aligned}$$

which corresponds to (3), and (4) with $\tau = 1$, in section 2. The duration of punishment ℓ can be chosen depending upon the complementarity γ between the two firms' products. Namely, it can be verified that there is a *strictly decreasing sequence* $\{\gamma_{[\nu]}\}_{\nu=0}^{\infty}$ such that

- $\gamma_{[0]} = 1\,,\ \gamma_{[1]} = -\dfrac{2}{3}\,,$
- Whenever $\gamma_{[\nu]} \le \gamma < \gamma_{[\nu-1]}$ ($\nu = 2, 3, \cdots$), there exists a solution $q^{P[\ell]} \in [0, q^M)$, $\delta^{**}[q^M] \in (0,1)$ to the system (15)-(16) when $\ell = \nu$ but there exists no such solution when $\ell < \nu$.
- $\lim_{\nu \uparrow \infty} \gamma_{[\nu]} = -1\,.$

As we have already confirmed in 3.1, prices are above the marginal cost k when each firm produces nil, and when each firm produces $q^{P[\ell]} \in [0, q^M]$. Thus once again, the price floor is never a binding constraint when multi-period punishment is inflicted.

Hereby both under single-period punishment and multi-period punishment, the following observation holds.

Proposition i : In a *quantity-setting* supergame with linear stage-demand (9), sustainability of the most profitable sustainable collusion, defined by the admissible range of the discount factor, is unaffected by the marginal cost k .

4 Linear Bertrand duopoly supergames

Now we consider the same linear duopoly supergame as in section 3 *except* that, at the beginning of every stage game, the two firms simultaneously set their *prices* $p_1 \geq 0$, $p_2 \geq 0$. By inverting (9), the direct demand function is obtained as

$$q_i = \frac{1}{1+\gamma} - \frac{1}{1-\gamma^2} p_i + \frac{\gamma}{1-\gamma^2} p_j \qquad \{i,j\} = \{1,2\} \tag{17}$$

as long as $\gamma < 1$. If the selected prices result in positive quantities $\widehat{q}_1 \geq 0$, $\widehat{q}_2 \geq 0$ according to (17), then these quantities realise unmodified. Otherwise, if $\widehat{q}_i = \min\{\widehat{q}_1 , \widehat{q}_2\} < 0$, the non-negative quantity constraint forces the actual quantity realisation to be $q_i = 0$ and then q_j is determined by (9) setting $q_i = 0$, or, if this q_j is also negative, then realised quantities are $q_1 = q_2 = 0$.

Note that, as $\gamma \uparrow 1$, the demand function (17) approaches the undifferentiated Bertrand demand function. Hereby without infringing the continuity of the model (with respect to γ), we can assume that when $\gamma = 1$ and $p_i < p_j$ the whole demand would be taken by firm i whilst firm j sells nil, and that when $\gamma = 1$ and $p_i = p_j$ the two firms share the demand evenly.

Based upon these resulting quantities q_i , we again conform to our previous notation in defining profit functions $\pi_i(p_i , p_j)$ and price-reaction functions

$$p_i^*[p_j] = \max\left\{\frac{1-\gamma+k+\gamma p}{2} , 0\right\} \qquad \{i,j\} = \{1,2\} . \tag{18}$$

As the demand structure is identical to our previous model in section 3, the most profitable outcome, if sustainable, is also identical. Namely, both firms set the *monopoly price*

$$p_1 = p_2 = p^M = \frac{1+k}{2} . \tag{19}$$

4.1 Single-period symmetric optimal punishment

Any symmetric profile $p_1 = p_2 = p^C$ is sustainable by means of single-period optimal punishment (also see Lambson, 1987, 1994, 1995) if and only if the discount factor δ is greater than or equal to the threshold $\delta^*[p^C]$ which is determined by

$$\pi_i(p_i^*[p^C], p^C) - \pi_i(p^C, p^C) = \delta^*[p^C]\left(\pi_i(p^C, p^C) - \pi_i(p^{[1]}, p^{[1]})\right); \tag{20}$$

$$\pi_i(p_i^*[p^{[1]}], p^{[1]}) - \pi_i(p^{[1]}, p^{[1]}) = \delta^*[p^C]\left(\pi_i(p^C, p^C) - \pi_i(p^{[1]}, p^{[1]})\right). \tag{21}$$

Unlike in linear Cournot duopoly (see section 3), the most profitable price collusion $p^C = p^M$ (see notation (19)) has a solution without violating the *positive quantity constraint* over the whole range $\gamma \in (-1, 1]$, taking into account the following qualifications. First, over $\gamma \in \left(\sqrt{3} - 1\,, 1\right]$, deviation from the monopoly price $p^M = \dfrac{k+1}{2}$ drives the cheated firm's output to zero. Second, over $\gamma \in \left(\dfrac{3\sqrt{5}-5}{2}\,, 1\right]$, the optimal deviation in the punishment phase produces zero profits for the cheating firm. Introducing these constraints into (20)-(21) yields the following solutions:

$$p^{[1]} = \frac{2 - 3\gamma + k(2+\gamma)}{2(2-\gamma)}; \qquad \delta^*[p^C] = \frac{(2-\gamma)^2}{16(1-\gamma)} \qquad \gamma \in \left(-1\,, \sqrt{3}-1\right];$$

$$\begin{aligned} p^{[1]} &= k + \frac{(1-\gamma)\gamma - (1-k)\sqrt{-1+2\gamma-\gamma^3}}{(2-\gamma)\gamma} \\ \delta^*[p^C] &= \frac{(2-\gamma)^2(\gamma^2+\gamma-1)}{(\gamma^2+2\sqrt{-1+2\gamma-\gamma^3})^2} \end{aligned} \qquad \gamma \in \left(\sqrt{3}-1\,, \frac{3\sqrt{5}-5}{2}\right];$$

$$p^{[1]} = \frac{1+k}{2} - \frac{(1-k)\sqrt{2\gamma^2+\gamma-1}}{2\gamma}; \qquad \delta^*[p^C] = \frac{\gamma^2+\gamma-1}{2\gamma^2+\gamma-1} \qquad \gamma \in \left(\frac{3\sqrt{5}-5}{2}\,, 1\right].$$

However, this penal price $p^{[1]}$, denoted by $p^{[1]}[\gamma]$ hereinbelow in order to be distinguishable from the generic penal price $p^{[1]}$, falls below the marginal cost k when $\dfrac{\sqrt{5}-1}{2} \approx 0.618 < \gamma \leq 1$, and hits its minimum $\dfrac{1+k-(1-k)\sqrt{2}}{2}$ at $\gamma = 1$. Notice that the latter boundary is non-negative iff $k \in [3 - 2\sqrt{2}\,, 1)$. Under the positive price constraint, this implies the following.

Proposition ii : In a *price-setting* supergame with linear stage-demand (9), the monopoly-level collusion can be sustained by means of single-period optimal punishment (20)-(21):

- over the whole range $-1 < \gamma \leq 1$ when $k \in [3 - 2\sqrt{2}\,, 1)$;

- only within the range $-1 < \gamma \leq \big(\gamma \,|\, p^{[1]}[\gamma] = 0\big)$ when $k \in [0\,, 3 - 2\sqrt{2})$.

In other words, when $k \in [0\,, 3 - 2\sqrt{2})$ and $\big(\gamma \,|\, p^{[1]}[\gamma] = 0\big) < \gamma \leq 1$, the system (20)-(21) fails to accommodate the monopoly price p^M without infringing the positive price constraint. In such cases, the most profitable collusion p^C sustainable by single-period punishment *subject to the positive price constraint* is derived by imposing the binding constraint $p^{[1]} = 0$ on the system (20)-(21). It is clear that this constrained collusion will not make the positive quantity constraint bind.

Economic intuition : When the two firms' products are mutually substitutes, punishment is inflicted by *reducing* the price. This is why the penal price $p^{[1]}[\gamma]$ decreases monotonically in the substitutability parameter γ. In particular, the penal price falls *below the marginal cost* k, driving profits temporarily negative, when the two firms' products are *close substitutes*. How far below k the penal price $p^{[1]}$ can fall without being negative, obviously depends upon the level of k. When the two firms supply complementary products, on the other hand, punishment is inflicted by *raising* the price. This serves to reduce the quantity. However, the positive quantity constraint is independent of the level of k, hence this does not create any dependency upon k.

4.2 Possibly multi-period symmetric optimal punishment

When $p^{[1]}[\gamma] < 0$, the monopoly-level collusion $p^C = p^M$ may require multi-period punishment defined in 2.3. When punishment lasts for ℓ periods ($\ell = 2, 3, \cdots$), the penal price during the first $\ell - 1$ periods is constrained at $p^P = 0$ whilst the penal price in the final ℓ-th period $p^{[\ell]}$ and the resulting minimum admissible discount factor $\delta^{**}[p^M]$ are determined by the system of simultaneous equations:

$$\begin{aligned} \pi_i(p_i^*[p^M], p^M) - \pi_i(p^M, p^M) &= \frac{\delta^{**}[p^M] - (\delta^{**}[p^M])^\ell}{1 - \delta^{**}[p^M]} \Big(\pi_i(p^M, p^M) - \pi_i(0,0)\Big) \\ &\quad + (\delta^{**}[p^M])^\ell \Big(\pi_i(p^M, p^M) - \pi_i(p^{[\ell]}, p^{[\ell]})\Big)\,; \end{aligned} \tag{22}$$

$$\begin{aligned} \pi_i(p_i^*[0], 0) - \pi_i(0,0) &= (\delta^{**}[p^M])^{\ell-1} \Big(\pi_i(p^{[\ell]}, p^{[\ell]}) - \pi_i(0,0)\Big) \\ &\quad + (\delta^{**}[p^M])^\ell \Big(\pi_i(p^M, p^M) - \pi_i(p^{[\ell]}, p^{[\ell]})\Big) \end{aligned} \tag{23}$$

which corresponds to (3), and (4) with $\tau = 1$, in section 2. The duration of punishment ℓ can be chosen depending upon the marginal cost k and the substitutability between the two firms' products γ. Namely, it can be verified that,

- for any pair k, γ, there is a unique integer $\underline{\ell}$ such that there exists a solution $q^{P[\ell]} \in [0, q^M)$, $\delta^{**}[p^M] \in (0,1)$ to the system (22)-(23) when $\ell = \underline{\ell}$ but not when $\ell < \underline{\ell}$;

- for any γ, the minimum admissible discount factor $\delta^{**}[p^M]$ resulting from $\ell = \underline{\ell}$ monotonically decreases k.

This, in conjunction with Proposition ii, implies the following result which is common to both single- and multi-period symmetric punishment.

Proposition iii : In a *price-setting* supergame with linear stage-demand (9), the set of sustainable collusive outcomes $\{\gamma, \delta, p^C\}$ (at least weakly) increases in the marginal cost k.

5 Reexamination of price-quantity duality

In this section, we investigate the similarities and differences between price competing and quantity competing supergames. In general, there exists duality between a price setting game and a quantity setting game. For instance, in linear duopoly with demand structure (9), the duality between price setting games and quantity setting games can be expressed essentially by flipping the sign of γ. This duality has been well known in the literature, and therefore connoisseurs may have been puzzled why in our paper, on the contrary, these two forms of competition entail substantially distinct game-theoretic characteristics.

The asymmetry between price competition and quantity competition is not inherent in Abreu's optimal punishment rule, but stems purely from non-negativity constraints. For instance, Deneckere (1984) shows the similarity between Bertrand games with positive quantity constraints but without positive price constraints, and Cournot games with positive price constraints but without positive quantity constraints. Note that, since Deneckere's analogy between Cournot and Bertrand games is about the one-shot deviation *from the collusive path*, all the results there must be independent of what penal codes are employed in order to sustain the collusion. Instead, the key is: when we interchange firms' strategic variables between prices and quantities, we also need to interchange likewise all the relevant constraints such as non-negativity, in order to maintain the duality.

In our context, it is only the positive price constraint of which the implication to collusive sustainability varies with the marginal cost k, not the positive quantity constraint. Insofar as the marginal cost k is large enough to push the price floor, $p \geq 0$, sufficiently far below the marginal cost ($k \geq 3 - 2\sqrt{2} \approx 0.1716$ in our linear duopoly supergame example), the positive price constraint never binds. As k decreases, the positive price constraint starts to bind in a Bertrand supergame with firms producing close substitutes. This serves to reduce the sustainability of collusion.

In the extreme, when $k = 0$, it is intuitively clear that the duality between Bertrand and Cournot games is restored. Note in particular that, as supermodularity increases to the limit, the constrained multi-period punishment asymptotically coincides with one-shot Nash reversion, thereby becomes security-level punishment (see Lambson, 1987, 1994, 1995) even though the duration of punishment tends to infinity. This is obvious in a Bertrand supergame with firms producing perfect substitutes. In a Cournot supergame with firms supplying perfect *complements*, on the other hand, the duration of punishment again grows to infinity. It is straightforward to verify that, in our linear duopoly example, in the limit where $\gamma \downarrow -1$,

$$\frac{\pi_i(q_i^*[q^M], q^M)}{\pi_i(q^M, q^M)} \to 2, \qquad \pi_i(0,0) = 0$$

so that the critical discount factor in sustaining monopoly-level collusion converges to 1/2, the same level as in the Bertrand supergame with perfect substitute products.

6 Policy and regulatory implications

As we have shown in sections 2 and 4, in Bertrand supergames with firms supplying close substitutes, sustainability of tacit price collusion hinges upon the prospect of charging temporarily a price below the marginal cost, and how far below it can be. This leads to the following straightforward policy implications.

- **Production subsidies** and **sales subsidies** serve to reduce marginal production costs perceived from the firm's private profit concern, whereby increase the mark-up. This, in light of our foregoing analysis, can help destabilise price collusion if firms' products are close substitutes. By the same token, **sales taxes** do the opposite, bringing an adverse effect of protecting collusive sustainability.

- **Corporate taxation** at a flat marginal tax rate is known to be non-distortionary, but this is true only insofar as the flat rate extends to negative profits, i.e., only if losses are proportionately subsidised. In reality, losses are most typically not subsidised. This serves to magnify, in relative terms, the effective losses incurred by penal prices below marginal costs. It hence increases the effectiveness of punishment and thereby serves to protect collusive sustainability.

- **Subsidies and tax credits given to innovative investment**, which is typically aimed at reducing marginal production costs, encourage cost reducing *process innovation*. Cost reduction, as aforementioned, can bring an affirmative effect of curtailing collusive sustainability.

- **Anti-dumping regulation**, legally prohibiting prices below marginal costs, is a straightforward way to curtail the prospect of price collusion especially when the environment is unsuitable for direct implementation of antitrust price regulations. It deserves heightened attention that this regulation serves affirmatively if and only if it is legally enforced. If, on the other hand, dumping is subjected to "economic" sanctions such as fines or penal taxes instead of "legal" prosecution, then such "disincentive schemes" can magnify the effectiveness of punishment and thereby help sustain tacit collusion.

- Presence of sizeable *fixed costs* per firm per period affects the security level, i.e., the intertemporal individual rationality condition. Namely, it makes exit relatively more attractive, rendering harsher punishment no longer individually rational. This can bring an affirmative effect of reducing the sustainability of collusion by narrowing the choice set of available punishment. In this sense, various **subsidies and tax concessions given to fixed industrial expenses** have a welfare-reducing side-effect of encouraging tacit price collusion. It should also be noted that only *fixed costs per period*, not *sunk costs upon entry*, can serve to destabilise tacit price collusion.

- **Corporate taxation on a lump-sum basis**, such as compulsory registration or licensing, is known to be neutral in a static setting, but not necessarily so in our repeated-game framework. For, it augments fixed costs perceived from private firms' point of view, which can curtail collusive prospects as aforementioned. Note especially that only a lump-sum **subscription tax**, i.e., a lump sum per firm *per period*, can serve to destabilise price collusion, not a once-and-for-good sunk payment upon entry, the latter being truly neutral even in our dynamic oligopoly games.

7 Concluding discussion

The impact of marginal costs on collusive sustainability has frequently been overlooked in many of the existing theoretical studies on repeated oligopoly games. Using technically straightforward application of optimal punishment (Abreu, 1986, 1988 ; Abreu, Pearce and Stacchetti, 1986 ; Lambson, 1987 ; Häckner, 1996) and modifying it in light of positive price and quantity constraints, this paper has characterised some of the systematic relations between the magnitude of marginal costs and the sustainability of collusive outcomes.

Marginal costs affect oligopolists' capability of sustaining tacit collusion in two fronts. On one hand, the presence of substantial marginal costs facilitates collusion *at any price-quantity level* by lowering the constrained critical discount factor. On the other hand, when the discount factor is not high enough to sustain monopoly-level collusion, the presence of sizeable marginal costs enables collusion to be sustained at a higher (closer to monopoly) level.

Aside from its highly practical implications, our result offers an extra advantage, that is, it abounds in potential connexions to future research subjects. The effect of cost reduction immediately gives rise to the question of endogenising marginal costs by means of some precommitted investment. Such investment can be made either as capacity instalment, technological innovation, or the choice of different methods in corporate financing. Policy implications are another direction to probe. The prospect of endogenising marginal costs also relates to the issue of *joint ventures in R&D*, including legislative attitude either for or against such jointing. Last, but not least, the widely studied subject of *organisational design*, especially those questions such as why firms appear to retain inefficient production technologies on purpose, and why firms tend to opt for and adhere to inefficient managerial practices and seemingly unnecessarily bureaucratic corporate structures, can possibly be disentangled in light of cost-side theory on tacit collusion.

Appendix

A more general form of multi-period punishment can be described as follows. Firms stay on a prescribed collusive path α^C, until any deviation is detected. Once a deviation is detected in period t, then in the next period $t+1$, firms switch to a multi-period punishment phase where *both* firms adopt the punishment action $\alpha^{P[1]}$ irrespective of which firm is punishing the other. In the punishment phase, whenever both firms follow the prescribed penal code $\alpha^{P[z]}$ at $t+z$ $(z=1,\cdots,\ell)$, then they play $\alpha^{P[z+1]}$ in the next period $z+1$. After ℓ periods, where ℓ is the spell of the penal code, firms revert to the initial collusive path forever, i.e., $\alpha^{P[\ell+1]}=\alpha^C$. If any deviation from the penal code is detected during the punishment phase, then the punishment phase restarts anew from $\alpha^{P[1]}$ all over again.

The incentive compatibility conditions in sustaining this ℓ-period penal code can be described by the following system of $\ell+1$ inequalities:

$$\pi_i(\alpha_i^*[\alpha^{P[z]}],\alpha^{P[z]})-\pi_i(\alpha^{P[z]},\alpha^{P[z]}) \le \sum_{h=1}^{\ell}\left(\delta^{**}[\alpha^C]\right)^h\left(\pi_i(\alpha^{P[z+h]},\alpha^{P[z+h]})-\pi_i(\alpha^{P[h]},\alpha^{P[h]})\right)$$

$$z=1,\cdots,\ell+1 \tag{24}$$

where, for notational convenience, let $\alpha^{P[L]}=a^C$ whenever $L>\ell$, that is, the last of these $\ell+1$ inequalities (with $z=\ell+1$) represents firms' incentive not to deviate *from the original collusive path*.

It is intuitively clear, however, that not all of these $\ell+1$ inequalities should bind (i.e., hold with strict equalities). For, a deviation from earlier stages of the penal phase generally gains stronger incentives than that from later stages. This implies that the first inequality, with $z=1$, together with the initial non-deviation condition from the collusive prescription, with $z=\ell+1$, should bind, whilst other $\ell-1$ inequalities, with $z=2,3,\cdots,\ell$, need not bind.

As we see from Abreu (1986, 1988), the optimality of punishment stands on two counterforces: on one hand the initial deviation from the prescribed collusive path should be discouraged by making the punishment as severe as possible, whilst on the other hand the deviation *from the first penal period* should also be prevented, in order for which the remainder of the punishment, i.e., the portion of punishment inflicted from the second penal period onwards, should not be too harsh to comply. The counterbalance between these two forces necessitates that the first penal period must carry as much punishment as feasible.

Whilst the incentive condition not to deviate from the first penal period binds, the conditions do not bind as for the latter $\ell - 1$ penal periods. This implies that there remain certain degrees of freedom on the allocation of punishment among these $\ell - 1$ periods *without affecting the overall optimality of the penal code*, insofar as the total discounted amount of punishment remains unchanged (which also implies that the optimal penal tenure ℓ may not be unique). In the presence of such multiplicity of optimal penal codes, in this paper we select the "front loaded" punishment, inflicting as much punishment as immediately as possible. This clearly leads to the penal code described in subsection 2.3, whereby the system (24) reduces into (3)-(4). Our assertion is simply that this is *an* optimal punishment, not necessarily *the unique* optimal punishment.

References

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

Abreu, D.J. (1988), "On the Theory of Infinitely Repeated Games with Discounting", *Econometrica*, **56**, 383-96.

Abreu, D.J., D. Pearce and E. Stacchetti (1986), "Optimal Cartel Equilibria with Imperfect Monitoring", *Journal of Economic Theory*, **39**, 251-69.

Albæk, S. and L. Lambertini (1998), "Collusion in Differentiated Duopolies Revisited", *Economics Letters*, **59**, 305-08.

Chang, M.H. (1991), "The Effects of Product Differentiation on Collusive Pricing", *International Journal of Industrial Organization*, **9**, 453-69.

Chang, M.H. (1992), "Intertemporal Product Choice and Its Effects on Collusive Firm Behavior", *International Economic Review*, **33**, 773-93.

Deneckere, R. (1983), "Duopoly Supergames with Product Differentiation", *Economics Letters*, **11**, 37-42.

Dixit, A.K. (1979), "A Model of Duopoly Suggesting a Theory of Entry Barriers", *Bell Journal of Economics*, **10**, 20-32.

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

Friedman, J.W. and J.-F. Thisse (1993), "Partial Collusion Fosters Minimum Product Differentiation", *RAND Journal of Economics*, **24**, 631-45.

Häckner, J. (1994), "Collusive Pricing in Markets for Vertically Differentiated Products", *International Journal of Industrial Organization*, **12**, 155-77.

Häckner, J. (1995), "Endogenous Product Design in an Infinitely Repeated Game", *International Journal of Industrial Organization*, **13**, 277-99.

Häckner, J. (1996), "Optimal Symmetric Punishments in a Bertrand Differentiated Product Duopoly", *International Journal of Industrial Organization*, **14**, 611-30.

Lambertini, L. (1997), "Prisoners' Dilemma in Duopoly (Super)Games", *Journal of Economic Theory*, **77**, 181-91.

Lambson, V.E. (1987), "Optimal Penal Codes in Price-Setting Supergames with Capacity Constraints", *Review of Economic Studies*, **54**, 385-97.

Lambson, V.E. (1994), "Some Results on Optimal Penal Codes in Asymmetric Bertrand Supergames", *Journal of Economic Theory*, **62**, 444-68.

Lambson, V.E. (1995), "Optimal Penal Codes in Nearly Symmetric Bertrand Supergames with Capacity Constraints", *Journal of Mathematical Economics*, **24**, 1-22.

Ross, T.W. (1992), "Cartel Stability and Product Differentiation", *International Journal of Industrial Organization*, **10**, 1-13.

Rothschild, R. (1992), "On the Sustainability of Collusion in Differentiated Duopolies", *Economics Letters*, **40**, 33-7.

Singh, N. and X. Vives (1984), "Price and Quantity Competition in a Differentiated Duopoly", *RAND Journal of Economics*, **15**, 546-54.