REGULATING OLIGOPOLY

Jonathan Cave

No.461

WARWICK ECONOMIC RESEARCH PAPERS

THE UNIVERSITY OF WARWICK

DEPARTMENT OF ECONOMICS

REGULATING OLIGOPOLY

Jonathan Cave
Department of Economics
University of Warwick
Coventry CV4 7AL
England

No.461

March 1996

Regulating Oligopoly

Jonathan Cave*

The standard analysis of single-product monopoly regulation with hidden information is adapted to an oligopoly setting using two simple approaches: a "virtual objective function" approach (which also provides a novel approach to oligopolistic collusion) and a Bayesian mechanism design analysis.

NOTE: This paper is highly preliminary and should not be cited or distributed without the author's consent. In particular, while the names of authors working on related topics have been noted where applicable, no systematic literature survey has been undertaken. All comments are, of course, welcome.

1 Introduction

The theory of optimal regulation of a single market has largely concentrated on the case of a monopolist incumbent, possibly facing an unregulated competitive fringe. Within this literature[1], the two main strands are the "hidden information" strand started by [Baron and Myerson (1982)] and the "hidden actions" strand associated with the work of Laffont and Tirole (see for instance [Laffont and Tirole (1993)]). Related work on "yardstick competition" considers separated monopolies linked by a common regulatory authority, while the general "Ramsey pricing" literature addresses a single firm operating in multiple markets. Much of this literature concentrates on the assessment or prescription of regulations, with added richness resulting from the complexity of the underlying regulations and the need to obey increasingly stringent incentive compatibility and participation constraints. Recently, however, attention has turned to competition policy as a substitute for or complement to regulatory policy. This literature addresses two questions: where is competition desirable; and when is it feasible? These are best answered by predicting the consequences of a specific liberalization policy and comparing results with the *status quo ante.*

Throughout this literature, only the briefest attention[2] is paid to market structures and forms of conduct that are neither monopolistic nor competitive, despite the fact that observed market structures are almost always oligopolistic and may be collusive to some degree. This has several implications:

- Analysis of optimal regulation may be substantially affected by market structure;
- Comparisons between pre- and post-liberalization outcomes may depend on the associated structure and conduct;
- Structure, conduct and the ensemble of regulatory and competition policies used to mediate them may exert reciprocal and even evolutionary influences on each other; and
- Limits on regulatory power (e.g. the inability to implement direct revelation mechanisms) may have severe consequences.

[1]An excellent survey can be found in [Armstrong, Cowan and Vickers (1994)].

[2]Two recent exceptions are working papers ([Wolinsky (1993)] and [Wang (1994)]) that develop models related to the second example of this paper.

This change in perspective has the potential to go beyond recomputation of standard results. Arguments about the force and effects of potential competition ("contestible markets") and potential regulation, for instance, are substantially modified. New phenomena also appear possible: the literature on the "capture" of regulators by incumbent monopolists can be expanded to include the possibility that actual, potential, or contingent regulation can facilitate or inhibit collusion. Finally, telecommunications markets have recently witnessed the emergence of overt and tacit cartels of firms based in different countries and subject to a mix of separated and overlapping regulations.

This paper has a modest goal: to construct and analyse a pair of simple models of optimal regulation of single-market oligopoly in a "hidden information" setting. Hopefully, the issues analyzed and the methods used will prove useful in the more general research program.

The first model is based on the construction of a "virtual objective function." This is a function of aggregate output that is maximised by the relevant equilibrium behaviour of the firms[3]. For a monopolist, it is simply the firm's profit function; for a sea of nonatomic competitors it is total (consumers' plus producers') surplus. For a Cournot oligopoly it is a weighted sum of (a measure of) aggregate profit and consumers' surplus. Collusive oligopolies are represented as placing relatively more weight on aggregate profits: the determination of the weights is also briefly examined.

1.1 Model variants

There are several important variants of the basic models under consideration, reflecting: i) underlying information conditions; ii) strategies available to the firms; iii) social objectives; and iv) the regulator's power to make transfers. The two examples in this paper take different positions with respect to all of these, and are thus intended to shed light on their aggregate impact. Clearly, much work remains to "fill in the gaps."

[3]This "virtual objective function" concept is not original with this paper. In one form or another, it has been developed by [Slade (1994)], [Bergstrom and Varian (1985)], [Cave and Salant (1995)], and [Loury (1986)]. [Slade (1994)] provides a survey of this literature. Our definition differs from hers, and our conditions are milder.

1.1.1 Information Conditions

The main information issues are whether the regulator's beliefs about firm marginal costs have finite or continuous support, and whether the firms' costs are common knowledge (among the firms).

The first issue has implications for the technical details of the analysis. The representation of firm behaviour in the finite support case is obtained from explicit consideration of the binding incentive compatibility and individual rationality conditions in the associated direct revelation game. In the continuous support case, incentive compatibility is replaced by an "envelope condition" obtained from firms' optimal behaviour in the original game.

When costs are common knowledge, the virtual objective function is obtained by summing the firms' first-order conditions in Cournot equilibrium. Under more general information conditions, the game between the firms must be expanded to include the different types of each firm, so that the virtual objective function becomes an "envelope" over the firm's Bayesian equilibrium behaviour. In this paper, we provide an examination of the Bayesian equilibrium that may be useful in obtaining such an objective function in a later paper.

1.1.2 Firm Strategies

In the moral hazard model developed by Baron and Myerson, the firm simply chooses a (common knowledge) level of output. In the adverse selection model of Laffont and Tirole, firms choose an unobserved level of effort as well. In both cases, the first-order conditions can be aggregated to conditions on aggregate output. Inspection of those conditions suggests a generalisation to partially collusive firms. However, the questions of how firms choose and sustain levels of collusion need further analysis. This paper suggests three possible approaches. The first derives from a repeated-game "trigger strategy" perfect equilibrium model, in which firms punish defections by immediate and permanent reversion to the non-collusive Cournot-Nash equilibrium. The resulting conditions define sustainable allocations of the collusive total output level and, as a result, sustainable levels of collusion. The second approach is to use a variant of the Rubinstein-Stahl bargaining solution according to which firms earn non-collusive levels of profit whilst negotiating an agreement. A third approach uses a variant of the [Cave and Salant (1995)] voting model, according to which firms select the preferred degree of competition by majority rule.

1.1.3 The Regulator's Objectives

Much of the literature assumes that the regulator wishes to maximise a weighted sum of consumers' surplus and firm profit. This leaves open the question of how the transfers used to ensure truthful revelation are treated. When consumers' surplus and profit are given equal weight, it seems best to ignore transfer payments. However, when distributional concerns are present, they should be reflected in social objectives. Thus, we are lead to consider the possibility that the regulator sets transfers T to maximise $V+\alpha\Pi-(1-\alpha)T$ as an alternative to the "standard" objective function: $V+\alpha\Pi$. The power of this approach comes from the relation between the regulator's objective (a weighted sum of V, Π, and T) and the firms' virtual objective function (a differently-weighted sum of the same elements).

Another approach to the regulators' objective[4] comes from the procurement literature and similar settings in which the government acts as the purchaser on behalf of the public. Social welfare is some function $\tilde{W}(Q)$ of aggregate output Q, and the regulator tries to maximise the expected value of $\tilde{W}(Q)-T$. This approach offers the advantage of computational simplicity.

1.1.4 The Regulator's Transfer Power

It is conventional to "simplify" mechanism design problems by recourse to the Revelation Principle: an outcome can be implemented if and only if there is a "direct revelation" mechanism that implements it. With such a mechanism, moreover, one can speak meaningfully of "truthful" revelation without much loss of generality. However, it does not follow that outcomes that can be implemented by direct revelation mechanisms can necessarily be implemented by specific mechanisms that do not have at least the "span" (in a precise sense) of the direct revelation game. It is therefore useful to distinguish between situations where the government's power to make transfers is as powerful as a direct revelation game and those where the power to make transfers is more limited. Indeed, a substantial literature deals with situations where the government cannot even make transfers but must rely

[4]This approach has been taken by (among others) [Cremer and McLean (1985)], [Wolinsky (1993)], and [Wang (1994)].

on e.g., price caps. It is worth noting that in the monopoly setting there is no difference between price and quantity as the firm's strategy, so transfers as a function of either market price or quantity produced are equivalent to each other and to a direct mechanism. The power of the virtual objective function approach derives in part from the fact that the reduction of the behaviour of separate firms to the maximisation of a specific function of *aggregate* output allows us to pass back and forth between price and quantity, both as strategic variables and as the basis for transfers. However, this may not be sufficient to span the possibilities covered by direct revelation mechanisms, especially when the joint distribution of individual firms' marginal costs is complex.

In this paper we examine two polar cases. The first extends the Baron-Myerson approach to the case of colluding oligopoly, while the second is an extension of the "procurement" model to the Bayesian equilibrium of non-collusive oligopoly. These are tentative first steps into a rich and broad area, and no attempt at generality has been made - these are simply examples.

2 The Virtual Objective Function Approach

The idea of reducing an oligopoly problem to a maximization problem goes back to [Samuelson (1947)], who asked whether there existed a function of many variables whose critical values corresponded to the first-order conditions of a game. This question has a trivial positive answer: the Euclidean norm of the vector of derivatives of firm profits w.r.t. the same firms' strategy will clearly work. [Slade (1994)] and others sought a much more "informative" function: one whose derivative w.r.t firm i's strategy *always* equals the derivative of firm i's payoff w.r.t. its own strategy. In both cases, the function in question maps the *vector* of firm strategies into real numbers. This is clearly the right sort of function to use for questions involving non-critical points, such as (evolutionary or dynamic) stability or (subgame) perfectness.

The analysis we are attempting is simpler and has more modest requirements. The mechanism design approach is inherently a comparative statics one, and thus need not be concerned with reproducing the entire set of first-order conditions over all output vectors. Indeed, if the regulators' objectives can be written in terms of *aggregate* output, the "virtual objective" of the firms can usefully be written in the same way. This does not mean that the regulator does not care about the way output is divided between the firms, but only that this division is beyond his influence. This may be a result of a restriction on transfer power or deficient information. The point is that the resulting conditions are far milder and the form of the objective function is easier to relate to the regulator's objectives.

2.1 The symmetric case

We begin with a fairly straightforward extension of the Baron-Myerson analysis. There are n firms, with no fixed costs and a common (-knowledge) constant marginal cost θ, facing a downward-sloping demand Q(P) for a homogeneous product. The marginal cost is unknown to the regulator, who believes it is distributed on an interval $[\theta^-, \ \theta^+]$ according to a density f with cumulative distribution function F. The regulator imposes a (common) transfer scheme T(P) on each firm, in an attempt to maximize the expected value of a weighted average of consumer surplus and profit:

$$EW(P) = \int_{\theta^-}^{\theta^+} \left[\int_{P(\theta)}^{\infty} Q(p)dp + \alpha(P-\theta)Q(P) - (1-\alpha)nT(P) \right] f(\theta)d\theta \quad (1)$$

where $P = P(\theta)$ is the common price charged by the firms when they share the marginal cost θ.

2.1.1 The Cournot Problem

Faced with the transfer scheme T and aggregate output level $\overline{q}_{-i} = \sum_{j \neq i} q_j$ from other firms, firm i picks its output q_i to solve:

$$\max_{q_i}[P(\overline{q}_{-i} + q_i) - \theta]q_i + T(P(\overline{q}_{-i} + q_i)) \quad (2)$$

This leads to the following first-order condition:

$$P(\overline{q}_{-i} + q_i) - \theta + [q_i + T'(P(\overline{q}_{-i} + q_i)]P'(\overline{q}_{-i} + q_i) = 0 \quad (3)$$

Summing over n firms gives the following condition on total output Q:

$$nP(Q) - n\theta + [Q + nT'(P(Q)]P'(Q) = 0 \quad (4)$$

Under reasonable conditions[5], this has a unique solution, which corresponds in turn to a unique vector of individual firm output levels.

[5] See Appendix A

The derivatives w.r.t. total output of consumer surplus $V(P(Q))$ and total profit $\Pi(Q)$ are:

$$\frac{dV(P(Q))}{dQ} = -QP'(Q) \tag{5}$$

and:

$$\frac{d\Pi(Q)}{dQ} = P(Q) - \theta + [Q + nT'(P(Q))]P'(Q) \tag{6}$$

Comparing (4) with (5) and (6) shows that the Cournot equilibrium quantity maximizes

$$\Pi(Q) + (\frac{n-1}{n})V(P(Q)) + nT(P(Q)) \tag{7}$$

another weighted average of consumer surplus and profit, plus total transfer payments.

This can be generalised still further by observing that, if firms maximize profit plus a multiple of consumer surplus V, their total profits are monotone decreasing in the weight attached to V. We are thus lead to define a "collusion parameter" σ: firms *collude to degree* σ if their total output maximizes the virtual objective function

$$\Pi(Q) + (1-\sigma)V(P(Q)) + nT(P(Q)) \tag{8}$$

The advantages of this approach are:

- It allows the use of "envelope" methods by casting oligopolists' behaviour as a maximisation problem;
- It provides a simple parameterisation of the degree of collusion shown by firms; and
- It allows us to switch between price and quantity as the variable of choice, greatly simplifying subsequent calculations.

Maximizing w.r.t. price P, gives the following first-order condition:

$$[P - \theta]Q'(P) + \sigma Q(P) + nT'(P) = 0 \tag{9}$$

The solution to this equation gives the firms' pricing strategy $\widehat{P}(\theta, \sigma, T)$.

Writing the optimised value of the firms' objective as:

$$\Phi(\theta,\sigma,T)=\max_{Q}\Pi(Q,\theta)+(1-\sigma)V(P(Q))+nT(P(Q)) \tag{10}$$

we observe that it satisfies the "envelope condition:"

$$\frac{\partial\Phi(\theta,\sigma,T)}{\partial\theta}=-Q(P(\theta))=\frac{dV(P)}{dP} \tag{11}$$

Integration by parts gives the following expression for the expected value of the firms' objective:

$$\begin{aligned}\int_{\theta^-}^{\theta^+}\Phi(\theta,\sigma,T)f(\theta)d\theta &= [\Phi(\theta,\sigma,T)F(\theta)]_{\theta^-}^{\theta^+}+\int_{\theta^-}^{\theta^+}\frac{\partial\Phi(\theta,\sigma,T)}{\partial\theta}F(\theta)d\theta\\ &= -\int_{\theta^-}^{\theta^+}Q(P(\theta))F(\theta)d\theta \qquad (12)\end{aligned}$$

since $\Phi(\theta^+,\sigma,T)=F(\theta^-)=0$ under general conditions.

2.1.2 The Regulatory Problem

To solve the regulator's problem, we write expected welfare in terms of profit and consumer surplus:

$$EW=\int_{\theta^-}^{\theta^+}\{[\Pi+V]-(1-\alpha)[\Pi+(1-\sigma)V+nT]+(1-\alpha)(1-\sigma)V\}\,f(\theta)d\theta(13)$$

Maximizing pointwise w.r.t. price P, gives the following first-order condition:

$$\begin{aligned}0 &= [P-\theta]Q'(P)f(\theta)-(1-\alpha)Q'(P)F(\theta)\\ &\quad -(1-\alpha)(1-\sigma)Q(P)f(\theta) \qquad (14)\end{aligned}$$

which simplifies to the following expression for the socially-optimal price:

$$\widetilde{P}=\theta+(1-\alpha)\frac{F(\theta)}{f(\theta)}+(1-\alpha)(1-\sigma)\frac{Q(\widetilde{P})}{Q'(\widetilde{P})} \tag{15}$$

Rewriting this in terms of the elasticity of demand η:

$$\widetilde{P}\left[1+\frac{(1-\alpha)(1-\sigma)}{\eta}\right]=\theta+(1-\alpha)\frac{F(\theta)}{f(\theta)} \tag{16}$$

This resembles the standard Ramsey pricing formula and the other single-market optimal pricing formulae: it reduces to marginal cost pricing when there is no uncertainty or when distributional concerns are absent ($\alpha = 1$), and gives the Baron-Myerson price formula when collusion is perfect ($\sigma = 1$).

2.2 The Nonsymmetric case

Here, we consider the possibility that the firms may have different marginal costs: θ is replaced with a vector $\vec{\theta} = (\theta_1, ...\theta_n)$. To keep the analysis manageable, we maintain the standard mechanism design assumption that all elements of $\vec{\theta}$ are common knowledge among the firms. In this setting, the regulator may be able to exercise additional power by tuning transfers to the output levels of specific firms. However, we begin with the "standard" setting in which each firm chooses a quantity to produce and the transfer is determined by the market-clearing price.

2.2.1 The Cournot Problem

Faced with transfer scheme T and aggregate output $\bar{q}_{-i}$ by the other firms, firm i picks its output q_i to solve:

$$\max_{q_i}[P(\bar{q}_{-i}+q_i)-\theta_i]q_i+T(P(\bar{q}_{-i}+q_i)) \tag{17}$$

leading to the following first-order condition:

$$P(\bar{q}_{-i}+q_i)-\theta_i+[q_i+T'(P(\bar{q}_{-i}+q_i)]P'(\bar{q}_{-i}+q_i)=0 \tag{18}$$

Summing over the n firms as before, and substituting $\bar{\theta} = \frac{\sum_i \theta_i}{n}$ gives the following condition on total output Q:

$$nP(Q)-n\bar{\theta}+[Q+nT'(P(Q)]P'(Q)=0 \tag{19}$$

Under reasonable conditions[6], this has a unique solution, which corresponds in turn to a unique vector of individual firm output levels.

[6] See Appendix A

The derivative w.r.t. total output of consumer surplus $V(P(Q))$ remains:

$$\frac{dV(P(Q))}{dQ} = -QP'(Q) \tag{20}$$

In this case, we must be careful when writing aggregate profit, since it will depend on the distribution of total output across the firms. The obvious approach is to assume that aggregate output will be produced at the unweighted average $\bar{\theta}$ of the marginal costs, in which case the (net of transfers) profit function becomes:

$$\tilde{\Pi}(Q) = [P(Q) - \bar{\theta}]Q \tag{21}$$

we obtain the following expression for the derivative of aggregate profit w.r.t. aggregate output:

$$\frac{d\tilde{\Pi}(Q)}{dQ} = P(Q) - \bar{\theta} + [Q + nT'(P(Q))]P'(Q) \tag{22}$$

Comparing (19) with (20) and (21), we observe that as before the Cournot equilibrium quantity maximizes

$$\tilde{\Pi}(Q) + (\frac{n-1}{n})V(P(Q)) + nT(P(Q)) \tag{23}$$

which generalises to the virtual objective function

$$\tilde{\Pi}(Q) + (1-\sigma)V(P(Q)) + nT(P(Q)) \tag{24}$$

2.2.2 The Regulator's Problem

There is an additional wrinkle to the Baron-Myerson stye of analysis. Although the firms behave collectively as if they were maximizing the virtual objective function, their true profits are given by

$$\hat{\Pi}(\vec{q}) = P(\sum q_i)\sum q_i - \sum \theta_i q_i \tag{25}$$

The revenues given by the two profit functions are the same (as are the consumer surplus levels), but the costs differ. More precisely, (19) can be used to write:

$$\bar{\theta}Q = \frac{[P(Q) - T'(P(Q))P'(Q)]\sum \theta_i - n\bar{\theta}^2}{-P'(Q)} \tag{26}$$

By contrast, using (18), we obtain:

$$\sum \theta_i q_i = \frac{[P(Q) - T'(P(Q))P'(Q)]\sum \theta_i - n\sum \theta_i^2}{-P'(Q)} \tag{27}$$

The envelope result refers to the "virtual" objective function (using the costs given in (23), while the regulator's objective function refers to the costs in (24). With linear demand, this is purely a matter of accounting, but when $P'(Q)$ varies with Q, the formula for the optimal transfer schedule is modified.

2.2.3 Other Forms of Regulation

The results of this section are very sensitive to the form of the transfer function. This is demonstrated quite starkly by the example in the second section, where the regulator's transfer power is equivalent to a direct revelation mechanism, which allows him to implement the "no distortion at the bottom" regime (see e.g. [Ireland (1991)], [Armstrong and Vickers (1993)], or [Vagliasindi (1995)]) that converges to the first-best outcome under mild conditions. Of course, exotic forms of regulation can upset the conditions for existence of a virtual objective function, even in the weak sense employed here. However, a certain degree of extension is possible. Suppose that we replace the uniform transfers used so far by a two-part transfer scheme, $\langle T(Q), R_i(Q)\rangle$ where firm i gets

$$(P(Q) - \theta_i)q_i + T(Q)q_i + R_i(Q) \tag{28}$$

The first-order conditions

$$P(Q) - \theta_i + P'(Q)q_i + T(Q) + T'(Q)q_i + R'(Q) = 0$$

average to produce

$$P(Q) - \bar{\theta} + [P'(Q) + T'(Q)]\frac{Q}{n} + T(Q) + \bar{R}'(Q) = 0 \tag{29}$$

where $\bar{R}(Q) = \frac{R(Q)}{n}$. The net (including transfers) profit and consumer surplus functions for this case are

$$\begin{aligned} \Pi^n &= [P(Q) - \bar{\theta}]Q + T(Q)Q + n\bar{R}(Q) \\ V^n &= \int_0^Q [P(q) + T(q) + \frac{n\bar{R}(Q)}{q}]dq - [P(Q) + T(Q)]Q - n\bar{R}(Q) \end{aligned}$$

from which we conclude that the average first-order condition maximises

$$\Pi^n + \frac{(n-1)V^n}{n}$$

as before.

However, more general transfer schemes do not necessarily lead to maxmimization of a virtual objective function of *aggregate* output. With a transfer scheme of the form $t(q_i)$, for instance, the average first-order condition is

$$P(Q) - \bar{\theta} + \frac{P(Q)Q + \sum_i t'(q_i)}{n}$$

which does not necessarily reduce to a function of Q.

A similar issue arises when we relax the assumption that the vector of marginal costs is common knowledge among the firms. For some situations, we can find a function of $Q(\vec{\theta})$ that is maxmised by the Bayesian equilibrium among the firms (the function critically involves the joint distribution of the firms' marginal costs), while for other cases it is impossible.

2.3 Selecting The Optimal Degree of Collusion

In previous sections, we assumed that the degree of collusion, σ, was exogenously fixed. We now consider some alternatives; that collusive arrangements result from a repeated-game equilibrium, a bargain between the firms, or a voting procedure. In this section we sketch each of those possibilities for the hidden-information case, with an eye to eventually being able to investigate the effect of optimal regulation on the set of sustainable collusive schemes, and conversely how the possibility of influencing collusion affects the optimal regulatory scheme.

2.3.1 Repeated Game Collusion

We use a very simple model for illustrative purposes. The basic idea is that firms attempt to secure collusion by employing perfect equilibrium threats. This limits sustainable degrees of collusion. The policy-relevant observation is that the transfer scheme may affect the maximum sustainable degree of collusion. This has both negative and positive connotations; it suggests that ill-chosen regulation may facilitate collusion, while holding open the possibility that the above analysis could be modified in a way that allows the regulator to use the regulatory mechanism as an ani-collusive device.

The Basic Model To fix ideas, consider a group of firms with different marginal costs θ_i, who are considering adopting a collusive scheme with parameter σ. In other words, they will choose the total output quantity $\hat{Q}(\vec{\theta}, \sigma, T)$ that maxmises

$$\Phi(Q, \vec{\theta}) = [P(Q) - \overline{\theta}]Q + (1-\sigma)\left\{\int_0^Q P(q)dq - P(Q)Q\right\} + nT(P(Q)) \quad (30)$$

where $\overline{\theta}$ is the average marginal cost across firms.

This leads to the following first-order condition for $\hat{Q}(\vec{\theta}, \sigma)$:

$$P(\hat{Q}(\vec{\theta}, \sigma)) - \overline{\theta} + \sigma P'(\hat{Q}(\vec{\theta}, \sigma))\hat{Q}(\vec{\theta}, \sigma) + nT'(P(Q))P'(\hat{Q}(\vec{\theta}, \sigma)) = 0 \quad (31)$$

But this leaves open the question of how this total will be allocated across firms. If firms had identical characteristics, equal division would seem most likely. If sidepayments were allowed, all production "should" go to the lowest-cost firm. But in either of these cases there seems to be no *a priori* reason why σ would ever fall below 1. Indeed, in the sidepayments case the maximand is not the virtual objective given in but rather the profits accruing to the lowest cost firm operating as a monopolist.

The approach taken in this section is to asume that firms are playing an infinitely repeated, discounted game in which collusion is enforced by the (perfect) threat of permanent reversion to the *status quo ante* (formally, the state $\sigma = 0$). We can then characterise a degree σ of collusion as *feasible* if there exists a vector $\hat{q}(\vec{\theta}, \sigma)$ of outputs such that:

$$\sum_i \hat{q}_i(\vec{\theta}, \sigma) = \hat{Q}(\vec{\theta}, \sigma) \quad (32)$$

and, $\forall$ firm i,

$$\begin{aligned} 0 \geq & -\frac{[P(\hat{Q}(\vec{\theta}, \sigma)) - \theta_i]\hat{q}_i(\vec{\theta}, \sigma) + T(P(\hat{Q}(\vec{\theta}, \sigma)))}{1 - \delta_i} \quad (33) \\ & + \max_q\{[P(\hat{Q}_{-i}(\vec{\theta}, \sigma) + q) - \theta_i]q + T(P(\hat{Q}_{-i}(\vec{\theta}, \sigma) + q))\} \\ & + \left(\frac{\delta_i[P(\hat{Q}(\vec{\theta}, \sigma^p) - \theta_i]\hat{q}_i(\vec{\theta}, \sigma^p) + T(P(\hat{Q}(\vec{\theta}, \sigma^p))}{1 - \delta_i}\right) \end{aligned}$$

where $\hat{q}_i(\vec{\theta}, \sigma^p)$ is firm i's output in a "punishment equilibrium with collusion level σ^p. For example, we could set $\sigma^p = \frac{1}{n}$ to use the non-collusive Cournot-Nash equilibrium for punishment. Such a vector of outputs consitutes a *viable collusion scheme.*

As we shall show by example, () limits the allocation of output for a given level of collusion and also defines conditions under which no such allocation can be found. Such levels of collusion are not sustainable.

A Simple Example

Example 1 : *Consider the case of a linear demand curve $P = 1 - Q$ and the linear transfer scheme $T(P) = \alpha + \tau P$. Given degree of collusion σ, firms will produce in aggregate*

$$\hat{Q}(\vec{\theta}, \sigma) = \frac{1 - \bar{\theta} - n\tau}{1 + \sigma - n\tau} \tag{34}$$

which sells at a price of:

$$\hat{P}(\vec{\theta}, \sigma) = \frac{\sigma + \bar{\theta}}{1 + \sigma - n\tau} \tag{35}$$

where $\bar{\theta} = \frac{\sum \theta_i}{n}$. Letting $\sigma^p = \frac{1}{n}$, the conditions for $\hat{q}(\vec{\theta}, \sigma)$ to constitute a viable collusion scheme are:

$$\sum \hat{q}_i(\vec{\theta}, \sigma) = \frac{1 - \bar{\theta}}{1 + \sigma} \tag{36}$$

and (after some algebra):

$$\begin{aligned} 0 \ \geq \ & \frac{\hat{q}_i^2}{4} + \hat{q}_i \frac{(K - \theta_i)(3 - \delta_i) - \tau(1 - \delta_i)}{2(1 + \sigma)} \\ & + \frac{(1 - \delta_i)[\sigma + \bar{\theta} - (1 + \sigma)\theta_i]^2}{2(1 - \delta_i)} + L\delta_i \Pi_i^{CN} \end{aligned} \tag{37}$$

where

$$K = \frac{\bar{\theta} + \sigma}{1 + \sigma - n\tau}$$

and

$$L = \delta_i \Pi_i^{CN} + \frac{K}{4} \left[K - 4\tau \left(\frac{3 - \delta_i}{2(1 - \delta_i)} \right) - 2\theta_i \right] + \frac{(\tau + \theta_i)^2}{4}$$

and where the "punishment profit" Π_i^{CN} is:

$$\Pi_i^{CN} = \frac{[1+\tau+n\bar{\theta}-(n+1)\theta_i]^2}{(n+1)^2} \tag{38}$$

The RHS of inequality (37) is a quadratic in $\hat{q}_i(\vec{\theta},\sigma)$ and, combined with (36), gives a fairly simple geometric condition for the sustainability of σ. Providing σ is large enough, the allowable values for $\hat{q}_i(\vec{\theta},\sigma)$ lie in a compact interval $[q_i^-(\vec{\theta},\sigma), q_i^+(\vec{\theta},\sigma)]$, and σ is sustainable iff:

$$\sum_i q_i^-(\vec{\theta},\sigma) \leq \frac{1-\bar{\theta}-n\tau}{1+\sigma-n\tau} \leq \sum_i q_i^+(\vec{\theta},\sigma) \tag{39}$$

Using the Cournot-Nash threats, the maximum sustainable degree of collusion is an increasing concave function of the individual discount rates, and the minimal discount rate is an increasing convex function of the degree of collusion.

Regulation and Collusion The conditions derived in this section can be modified to include the possibility of regulation. If the regulatory scheme is insensitive to the degree of collusion, the operative questions are: first, how optimal regulation in the sense defined above (maximal welfare taking the degree of collusion as fixed) changes the range of sustainable collusive schemes; and second, how a regulatory scheme that took collusion as endogenous might differ from one in which collusion was exogenous. As we shall see, these are not trivial questions. In answering the first, we need to compare ranges of sustainable collusion parameters. We may also wish to take account of the fact that each corresponds to a range of sustainable allocations of the optimal total output. For various reasons we may wish to include inter-firm distributional considerations into the social welfare function.

With regard to the second issue, that of designing regulatory schemes to influence the degree of collusion, it is necessary to develop a theory of how firms choose from among the sustainable collusive schemes.

The example suggests that the ambiguity regarding the way the output corresponding to a given degree of collusion is distributed among the firms may disappear at the maximal sustainable collusion level. This certainly happens in the linear-demand example. It should then be possible to make the degree of collusion endogeneous from the regulators' point of view, at least under the pesimistic assumption that firms will collude to the maximum sustainable degree.

2.3.2 Other Methods

There are various other possibilities to consider. For instance, one could construct a Rubinstein-Stahl bargaining model acording to which firms negotiate a collusive regime. In this model, the firms would accrue profits according to some *status quo ante* (e.g. the non-collusive Cournot-Nash equilibrium) whilst negotiating an appropriate degree of collusion and division of the associated output. In technical terms, they would bargain over the increase in their profits relative to Cournot-Nash, constrained by their relative rates of time discount; the feasible allocations might be further restricted to sustainable collusion schemes in the sense defined bove.

Another interesting possibility, albeit a static one, is that firms vote for the desired degree of collusion, following the "Cartels That Vote" model of [Cave and Salant (1995)]. The alternatives are the sustainable degrees of collusion, and the profits corresponding to a given alternative are those derived from, for example, a bargaining solution.

3 The Procurement Approach

In this section, we consider the problem faced by a regulator who purchases a good on behalf of the public from a number of competing suppliers whose costs are unknown to the regulator and only partially known to the firms (each of whom knows its own marginal cost). As before, firm i's constant marginal cost is denoted θ_i, aggregate output is Q, and government transfers are denoted T. For this section, we limit θ_i to two possible values: $\theta_i \in \{\theta^-, \theta^+\}$ $\forall i$. The social value of aggregate production Q is given by an increasing concave function $\tilde{W}(Q)$ with the following properties:

- $\tilde{W}(0) = 0$
- $\lim\limits_{Q \to 0} \tilde{W}'(Q) = +\infty$
- $\lim\limits_{Q \to +\infty} \tilde{W}'(Q) = 0$

Before modelling the oligopoly situation, it is useful to derive the first-best and optimal monopoly outcomes to provide a benchmark.

3.1 Benchmarks

3.1.1 Full Information Optimum

First, we derive the full-information (first-best) optimal outcome. Since firms face no fixed cost[7], all production should be done by (one or more of) the lowest-cost firm(s). If the lowest marginal cost is θ, the associated optimal output level, $Q^o(\theta)$, solves:

$$\max_Q \tilde{W}(Q) - \theta Q \tag{40}$$

which leads to the first-order condition:

$$\tilde{W}'(Q^o(\theta)) = \theta \tag{41}$$

Thanks to the conditions on $\tilde{W}$, this equation has a unique solution that is decreasing in θ.

3.1.2 The Optimal Monopoly Outcome

If the regulator must purchase the commodity from a single firm, there is no difference between a direct regulation mechanism and a transfer schedule based on the quantity supplied. In other words, we can either use a direct mechanism in which the firm reports its marginal cost c, produces the quantity $Q(c)$ and receives the payment $T(c)$ and an indirect mechanism in which the firm is faced with a transfer schedule $T(Q)$ and chooses the quantity Q it wishes to produce. For consistency with later development we shall work with the direct mechanism. There is no a priori reason why the regulator should accept abitrary production levels when costs have a finite support, so we may limit our attention to two "acceptable" production levels, $\underline{Q}$ and Q^+, corresponding to the two possible marginal costs. As usual, the mechanism T must meet incentive compatibility and individual rationality conditions. Moreover, the incentive compatibility (individual rationality) condition is binding on the low-cost (high-cost) type of the firm:

$$T(\theta^-) - \theta^- Q^- = T(\theta^+) - \theta^- Q^+ \tag{42}$$

[7]Fixed costs are almost immaterial providing they are a weakly increasing function of marginal cost.

and

$$T(\theta^+) - \theta^+ Q^+ = 0 \tag{43}$$

These expressions let us solve for the optimal transfers as functions of the desired output levels Q^+ and Q^- - this reduces the regulator's problem to one of finding the optimal production levels for each reported value of marginal cost. Letting $F = prob(\theta = \theta^+)$ denote the *a priori* probability that the firm has high marginal costs, we have the following expressions for the optimal transfers and the expected cost to the regulator:

$$T(\theta^+) = \theta^+ Q^+ \tag{44}$$

$$T(\theta^-) = \theta^+ Q^+ - \theta^-(Q^+ - Q^-) \tag{45}$$

and

$$\begin{aligned} ET &= FT(\theta^+) + (1-F)T(\theta^-) \\ &= \theta^+ Q^+ - (1-F)\theta^-(Q^+ - Q^-) \end{aligned} \tag{46}$$

The regulator chooses output to maximise expected net welfare.

$$\max_{Q^+,Q^-} F\tilde{W}(Q^+) + (1-F)\tilde{W}(Q^-) - ET \tag{47}$$

leading to the first-order conditions:

$$F\tilde{W}'(Q^+) - \theta^+ + (1-F)\theta^- = 0 \tag{48}$$

and

$$(1-F)\tilde{W}'(Q^-) - (1-F)\theta^- = 0 \tag{49}$$

It will be noticed that (41) is the same as (49), so the regulator can reduce the low-cost firm's surplus to 0. We can rewrite (48) as

$$\tilde{W}'(Q^+) = \theta^+ + \frac{(1-F)(\theta^+ - \theta^-)}{F} \tag{50}$$

which shows that the high-cost firm earns some "informational rents." Going further, let us write the output produced by an optimally-regulated monopolist as $Q^M(\theta, F)$. Clearly, $Q^M(\theta^-, F) = Q^0(\theta^-)$, while $Q^M(\theta^+, F)$ is an increasing function of F, reaching its maximal value $Q^0(\theta^+)$ when $F = 1$.

The expected cost of optimal monopoly regulation can be written:

$$ET^M(F) = \theta^+ Q^M(\theta^+, F) - (1-F)\theta^-[Q^M(\theta^+, F) - Q^0(\theta^-)] \tag{51}$$

which is an increasing function of F, while the expected net welfare:

$$E(\tilde{W}-T)^M(F) = F\tilde{W}(Q^M(\theta^+, F)) + (1-F)\tilde{W}(Q^0(\theta^-)) - ET^M(F) \tag{52}$$

is decreasing and convex in F.

3.1.3 The Optimal Oligopoly Outcome

We now turn to the oligopoly problem. By contrast with the example in the first section, we relax the assumption that the firms' marginal costs are common knowledge among them, and deal directly with the Bayesian equilibrium of the game. We further assume that the regulator uses a direct mechanism according to which the firms report a vector $\vec{\varphi}$ of marginal costs (not necessarily equal to the "true" vector $\vec{\theta}$), leading the regulator to specify a vector $\vec{q}$ of outputs and an associated vector $\vec{T}$ of transfer payments.

Each firm's cost can take only one of two possible values; these values are the same for all firms. To simplify notation in what follows, let us denote by $Eq_i(\theta_i)$ the amount that firm i expects to be asked to produce, and by $ET_i(\theta_i)$ the amount it expects to be paid, when it reports a marginal cost of θ_i:

$$Eq_i(\theta_i) = \frac{\int\limits_{\{\vec{\varphi}:\varphi_i=\theta_i\}} q_i(\vec{\varphi})f(\vec{\varphi})d\vec{\varphi}}{\int\limits_{\{\vec{\varphi}:\varphi_i=\theta_i\}} f(\vec{\varphi})d\vec{\varphi}} \tag{53}$$

and

$$ET_i(\theta_i) = \frac{\int\limits_{\{\vec{\varphi}:t_i=\theta_i\}} T_i(\vec{\varphi})f(\vec{\varphi})d\vec{\varphi}}{\int\limits_{\{\vec{\varphi}:\varphi_i=\theta_i\}} f(\vec{\varphi})d\vec{\varphi}} \tag{54}$$

The incentive compatibility (individual rationality) constraints bind for the low-cost (high-cost) type of each firm, leading to the following expressions:

$$ET_i(\theta^-) = \theta^- Eq_i(\theta^-) + (\overline{\theta} - \theta^-)Eq_i(\theta^+) \tag{55}$$

and

$$ET_i(\theta^+) = \theta^+ Eq_i(\theta^+) \tag{56}$$

where $Eq_i(\theta^+) < Eq_i(\theta^-)$. Note that in these expressions expectations are taken w.r.t. the types of the other firms. Since we are using a direct revelation mechanism, no firm knows exactly what output it will be asked to produce when reporting its cost - hence the expectation operator in the expression for output.

The precise form and implications of these conditions depend on the joint distribution of the firms' costs.

independent Costs We start by assuming that costs are drawn independently: the probability that firm i has high cost is denoted $F_i = prob(\theta_i = \theta^+)$. The expected payment to firm i is:

$$E_\theta T_i(\vec{\theta}) = [\theta^+ - (1-F_i)\theta^-]Eq_i(\theta^+) + (1-F_i)\theta^- Eq_i(\theta^-) \tag{57}$$

and the aggregate expected cost to the regulator is the sum of this expression over i:

$$\begin{aligned} E_\theta T(\vec{\theta}) &= E_\theta \sum_i T_i(\vec{\theta}) \\ &= \theta^+ \sum_i Eq_i(\theta^+) - \theta^- \left[\sum_i (1-F_i)[Eq_i(\theta^+) - Eq_i(\theta^-)]\right] \end{aligned} \tag{58}$$

The regulator must choose the optimal output levels $\vec{q}(\vec{\theta})$ to solve:

$$\max_{\vec{q}(\vec{\theta})} \left\{ E_\theta \widetilde{W}(\sum_i q_i(\vec{\theta})) - \theta^+ \sum_i Eq_i(\theta^+) - \theta^- \left[\sum_i (1-F_i)[Eq_i(\theta^+) - Eq_i(\theta^-)]\right]\right\} \tag{59}$$

We solve this problem by differentiation w.r.t. each $q_i(\vec{\theta})$.

After some tedious but straightforward algebra we obtain the following charaterisation of the optimal mechanism:

- If at least one firm claims low cost ($\theta_i = \theta^-$), the regulator will request a total output of $Q^o(\theta^-)$, divided in an arbitrary way between the firms reporting θ^-.
- If all firms claim to have high cost ($\theta_i = \theta^+ \quad \forall i$), then a firm i* with the *highest* prior probability of high cost ($F_{i*} = \max_j\{F_j\}$) produces $Q^M(\theta^+, F_{i*})$ and all others produce 0.
- The expected cost of the optimal mechanism is:

$$ET^{CN}(\vec{F}) = [1 - \prod_{i \neq i*} F_i]ET^M(0) + [\prod_{i \neq i*} F_i]ET^M(F_{i*}) \tag{60}$$

- The expected net welfare is:

$$E(\tilde{W}-T)^{CN}(\vec{F}) = [1 - \prod_{i \neq i*} F_i]E(\tilde{W}-T)^M(0) + [\prod_{i \neq i*} F_i]E(\tilde{W}-T)^M(F_{i*}) \tag{61}$$

- If the number of firms becomes large in such a way that $\prod_{i=1}^{n} F_i \underset{n\to\infty}{\longrightarrow} 0$ (for example, if costs are i.i.d.), then:

$$ET^{CN}(\vec{F}) \longrightarrow ET^{M}(0) \tag{62}$$

and

$$E(\tilde{W}-T)^{CN}(\vec{F}) \longrightarrow E(\tilde{W}-T)^{M}(0) \tag{63}$$

In other words, as the number of firms increases the planners' expected cost and the net social welfare converge to the first-best outcomes in a world with a single, low-cost firm.

General Cost Distribution In this case, we must replace the independent probabilities F_i used in the previous section with a general distribution F on $\{\theta^-, \theta^+\}^n$. We denote the marginal and conditional distributions obtained from F by $F(\theta_i)$ and $F(\theta_{-i}|\theta_i)$. For tractability, we assume that $F(\theta_i) > 0 \quad \forall i, \quad \theta_i$. This is very close to the setup analysed by Cremer and McLean's analysis of a monopolist selling to customers with jointly-distributed demands. Following their approach, we shall asume that the conditional distributions are sufficiently informative: for each firm i there exists a type vector θ_{-i} for the other firms such that:

$$F(\theta_{-i}|\theta^+) \neq F(\theta_{-i}|\theta^-) \tag{64}$$

Under this condition, the optimal expected net welfare can be shown to be a convex combination of the two extreme monopoly expected net welfares:

$$\begin{aligned} E(\tilde{W}-T)^{CN}(F) &= [1-F(\theta^+,...,\theta^+)]E(\tilde{W}-T)^{M}(\theta^-) \\ &\quad +F(\theta^+,...,\theta^+)E(\tilde{W}-T)^{M}(\theta^+) \end{aligned} \tag{65}$$

As before, this converges to the first-best result $E(\tilde{W}-T)^{M}(\theta^-)$ as the set of firms gets large iff firms are added in such a way that $F(\theta^+,...,\theta^+) \to 0$.

Fixed Costs, Continuous Types, and other Speculations The analysis remains substantially unchanged if the firms have fixed costs, providing that the fixed costs are the same for all firms or an increasing monotone function of marginal cost. The main difference is that the regulator selects one of the firms reporting low cost (if any) at random (with equal probabilities in the iid case, and the firm with the highest prior probability of telling the truth in the independent case) and assigns it the entire output.

With a continuum of types, the incentive compatibility conditions will have to be modified in the standard way, the single-crossing condition must be verified, and other conditions (monotonicity) must be imposed or verified in order to deal with the tractable "relaxed program."

4 Appendix A: Existence and Uniqueness of Cournot Equilibrium

There are many conditions leading to existence of a unique Cournot-Nash equilibrium; see e.g. [Szidarovszky and Yakowitz [1982]], [Novshek (1985)], and[Cave and Salant (1987)] among other places. The simplest approach is the "backwards best response function" introduced in the first of these references: it asks what amount $q_i(X)$ each firm i would produce in an equilibrium where total output was X. Summing these functions gives an expression for total output Q as a function $F(X)$ of X, and one looks for conditions under which the equation $X = Q$ has a unique solution. Roughly, existence is ensured if: i) $F(0) > 0$; ii) $\lim_{X\to\infty} F(X)$ is finite; and iii) $F(X)$ never "jumps down." As [Cave and Salant (1987)] show, uniqueness is ensured if (for instance), conditions i) and ii) are satisfied, F is continuous, and has slope strictly less than one where it is differentiable.

These conditions are simplified by the assumption of constant marginal cost, but in the current setting they should be understood as joint constraints on the demand function and the transfer function. We can state sufficient conditions precisely as follows:

1. Demand P is at least twice differentiable;
2. $nP(0) - \sum_i \theta_i + nT'(P(0))P'(0) > 0$;
3. $\lim_{Q\to\infty} \{nP(Q) - \sum_i \theta_i + nT'(P(Q))P'(Q)\} < \infty$;
4. $\forall Q, \forall q \leq Q$,

$$(n+1)P'(Q) + [q + nT'(P(Q))]P''(Q) + nT''(P(Q))(P'(Q))^2 < 0$$

This is certainly satisfied if P satisfies the "Novshek conditions" and T is a concave (composite) function of Q. However, one should also note that "deficiencies" in either P or T can be compensated for by the properties of the other function.

This establishes conditions under which a unique Cournot equilibrium total output $Q*$ exists. The uniqueness of the distribution of this output across firms follows from the observation that each firm i will produce

$$q_i = \frac{P(Q*) - \theta_i}{-P'(Q*)} + T'(P(Q*))$$

by inspection.

5 References

References

Armstrong and Vickers (1993) Armstrong, M. and J. Vickers, 1993, "Price Discrimination, Competition, and Regulation," *Journal of Industrial Economics*, 4, pp. 335-359.

Armstrong, Cowan and Vickers (1994) Armstrong, M., M. Cowan, and J. Vickers, *Regulatory Reform: Economic Analysis and the British Experience*, MIT Press.

Baron and Myerson (1982) Baron, D. and R. Myerson, 1982, "Regulating a Monopolist with Unknown Costs," *Econometrica* **50**, pp. 553-568.

Bergstrom and Varian (1985) Bergstrom, T. and H. Varian, 1985, "When are Nash Equilibria Independent of the Distribution of Agents' Characteristics," *Review of Economic Studies* **52**, pp. 715-718.

Cave and Salant (1987) Cave, J and S. Salant, 1987, "Cartels that Vote: Agricultural Marketing Boards and Induced Voting Behavior," in *Public Regulation: New Perspectives on Institutions and Policies*, E. Bailey, ed., MIT Press.

Cave and Salant (1995) Cave, J. and S. Salant, 1995, "Cartel Quotas under Majority Rule, *American Economic Review*, March 1995, pp. 88-102.

Cremer and McLean (1985) Cremer, J. and R. McLean, 1985, "Optimal Selling Strategies under Uncertainty for a Discriminating Monopolist when Demands are Interdependent," *Econometrica*, **53**, pp. 345-361.

Ireland (1991) Ireland, N., 1991, "Welfare and Non-Linear Pricing in a Cournot Oligopoly," *Economic Journal*, **101**, pp. 949-957.

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

Loury (1986) Loury, G., 1986, "A Theory of 'Oil'-igopoly: Cournot Equilibrium in Exhaustible Resource markets with Fixed Supplies," *International Economic Review* **27**, pp. 285-301.

Novshek (1985) Novshek, W., 1985, "On the Existence of Cournot Equilibrium," Review of Economic Studies 52 pp. 85-98.

Samuelson (1947) Samuelson, P., 1947, *Foundations of Economic Analysis*, Harvard Univ. Press.

Slade (1994) Slade, M., 1994, "What does an Oligopoly Maximize?," *Journal of Industrial Economics*, **42**, pp. 45-61.

Szidarovszky and Yakowitz [1982] Szidarovszky, F. and S. Yakowitz, 1982, "Contributions to Cournot Oligopoly Theory," *Journal of Economic Theory* **28**, pp. 51-70.

Vagliasindi (1995) Vagliasindi, M., 1995, "The Cream-Skimming Paradigm," mimeo, University of Warwick.

Wang (1994) Wang, G. H., 1994, "Regulating the Oligopoly with Unknown Costs," Economics Department Research Report 9406, Univ. of Western Ontario.

Wolinsky (1993) Wolinsky, Asher, 1993, "Regulation of Duopoly under Asymmetric Information," CMSEMS Discussion Paper 1061, Northwestern U.