Chapter 3
Collusion and tacit collusion

It was possible to effect an arrangement with a few of the outside firms as to selling prices with the object of stopping the rot, but the agreements entered into to this end were not enforceable . . . gentlemanly agreements—the passing of the "word of honour", were all that we had to rely upon and they were observed until the manufacturer saw his way to "making a bit" by ignoring them.

O'Hagen (1926, Volume II, p. 73), quoted in Cook (1958, p. 46).

3.1 Introduction

In Chapter 2 we examined the Cournot and Bertrand oligopoly models. These models are static, in the sense that they do not explicitly consider the passage of time; they are one-period models. In these and most other one-period models of noncooperative oligopoly, equilibrium output exceeds the monopoly level, with the result that firms can increase their profit if they are able to restrict output and raise price.

There is sometimes a tendency to identify collusion with output restriction (below the equilibrium output level of a one-period model), but this is imprecise. Whether output is or is not restricted below its one-period equilibrium level is an objective question, but collusion is a legal concept. At an elementary level, one can say that collusion occurs when independent firms reach joint decisions about their market conduct. If firms make independent decisions that move output to or toward the monopoly level, they have not colluded. If firms make joint decisions that have the same effects, they have colluded. The standards that must be satisfied to justify a finding that firms' decisions have been jointly arrived at depend on the details of the legal regime that applies to business behaviour.[1]

In this chapter, we first look at one way to formalize the factors that influence firms' output decisions when firms compete repeatedly over time, not just for a single period.

[1] In much the same way, whether one person has killed another is a reasonably objective question. Killing is only murder (a legal concept) if the applicable legal standards are satisfied.

We also examine some practices (basing point pricing and publicity of prices and quantities in individual transactions) that may facilitate output restriction and look at the impact of collusion on market performance. Finally, we study EC competition policy toward collusion.

3.2 Noncooperative collusion

3.2.1 Trigger strategy

Collusion is prohibited by Article 81 of the EC Treaty, which means (among other things) that firms cannot sign legally enforceable contracts to bind themselves to collusive understandings. Thus collusion, if it occurs, must be *noncooperative*, in the sense that each firm that is party to a collusive agreement, a member of a cartel, will keep the collusive agreement only so long as it is in its own interest to do so.

The basic elements of the economics of noncooperative collusion are straightforward. When firms independently maximize their own short-run profit, their combined profit is generally less than monopoly profit. We have seen this for Cournot oligopoly, and it is true for most other models of imperfectly competitive markets as well. It follows that firms can increase their profit, compared with noncooperative oligopoly equilibrium, if they can cooperate and restrict output.

If all firms restrict output, then any one firm can earn an even greater profit in the short run by cheating on than by keeping to the restrictive agreement—by producing more than its share of collusive output. However, if other firms realize cheating is taking place, it is reasonable to expect that the collusive agreement will break down, and that other firms would expand output as well. Thus a firm that defects from a collusive agreement for the extra short-run profit that defection brings would earn less profit in the more distant future than it would by keeping the agreement. It follows that the critical factors affecting whether or not noncooperative collusion will be successful are how firms trade off higher profit in the near future from cheating against lower profit in the more distant future and how quickly rival firms would become aware of cheating, if it should occur.

One way to formalize the idea of a tradeoff between profit received at different times is in terms of a so-called *trigger strategy* for noncooperative collusion (Friedman, 1971) in a market where firms repeatedly play a single-period Cournot game:

- each firm produces its share of the monopoly output in the first period, and produces its share of the monopoly output in each following period if all other firms produced their shares of monopoly output in the previous period;
- if in any period price is different from the monopoly price, all firms produce their Cournot outputs forever after.[2]

[2] The idea that one firm would believe that other firms would give up the possibility of colluding forever in response to one episode of defection may seem unrealistic. However, the results of the trigger strategy model carry over to strategies that involve less severe punishments.

Suppose, for example, that a market has the linear inverse demand curve

$$p = 100 - Q \tag{3.1}$$

in every time period and that all firms have constant marginal and average variable cost 10 per unit of output,

$$C(q) = 10q + F. \tag{3.2}$$

As explained in Section 1.3.1, fixed cost F does not vary with output.

Profit-maximizing monopoly output is 45,[3] and monopoly price 55. Monopoly profit per time period is

$$\pi_m = (55 - 10)(45) - F = 2025 - F. \tag{3.4}$$

If the market is a Cournot duopoly, equilibrium output per firm is 30, Cournot duopoly price is 40, and equilibrium profit per firm per time period is

$$\pi_1 = \pi_2 = (40 - 10)(30) - F = 900 - F. \tag{3.4}$$

Monopoly profit in this industry, before allowing for fixed cost, is 2025. The additional profit to be gained in any one time period if the firms collude and produce only the monopoly output is

$$2025 - 2(900) = 225. \tag{3.5}$$

If the two firms could agree to shut down one plant, they would also save one set of fixed cost, F, per period. Stepping outside the model, a firm that agreed to shut down its plant would most likely find its bargaining power sharply reduced in future communications about behaviour, making such an agreement unlikely.

Suppose the two firms agree that each will produce half the monopoly output; then each firm earns a profit that gives it half of the extra profit from colluding:

$$(55 - 10)(22.5) - F = 1012.5 - F. \tag{3.6}$$

If firm 2 produces 22.5 units of output, firm 1's residual demand curve is

$$p = 100 - 22.5 - q_1 = 77.5 - q_1. \tag{3.7}$$

Picking the output that makes marginal revenue equal to marginal cost along this residual demand curve gives firm 1's one-period profit-maximizing output if firm 2 is producing 22.5 units of output,

$$MR = 77.5 - 2q_1 = 10 \Rightarrow q_1 = 33.75 > 22.5. \tag{3.8}$$

If firm 2 produces half the monopoly output, firm 1's profit-maximizing output is greater than half the monopoly output.

If firm 1 defects from the cartel agreement in this way, price is

$$p = 100 - 22.5 - 33.75 = 43.75 \tag{3.9}$$

and firm 1's profit is

[3] This is found by setting marginal revenue, $100 - 2Q$, equal to marginal cost, and solving for output.

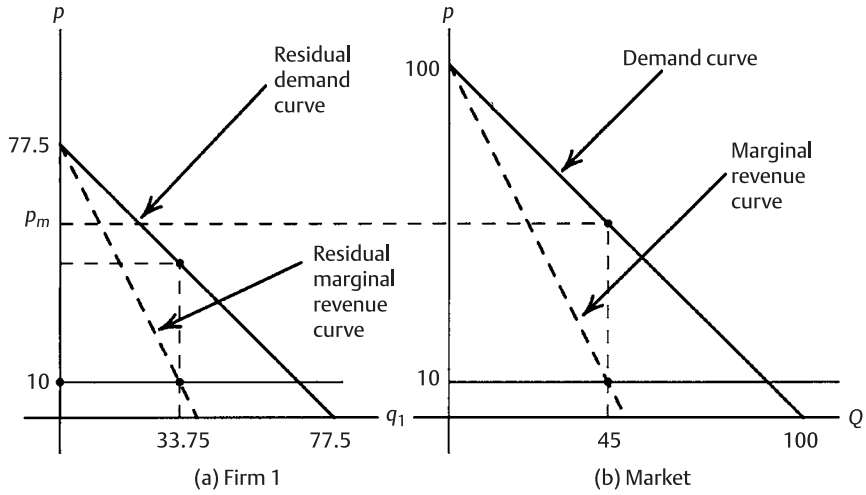

Fig. 3.1 Private profit maximization with a duopoly collusive agreement, $p = 100 - Q$, marginal cost = 10

$$(43.75 - 10)(33.75) - F = 1139.0625 - F, \tag{3.10}$$

which is greater than firm 1's share of the cartel profit (from (3.6), $1012.5 - F$).

The present-discounted value of the profit of a firm that follows the trigger strategy, assuming all other firms do so, is

$$V_{collude} = \frac{1}{1+r}\pi_{collude} + \frac{1}{(1+r)^2}\pi_{collude} + \ldots = \frac{1}{r}\pi_{collude}, \tag{3.11}$$

where r is the interest rate used to discount future income and $\pi_{collude}$ is the firm's collusive payoff in a single period.

The present-discounted value of the profit of a firm that defects from the trigger strategy, assuming all other firms follow the trigger strategy, is

$$V_{cheat} = \frac{1}{1+r}\pi_{cheat} + \frac{1}{(1+r)^2}\pi_{Cournot} + \frac{1}{(1+r)^3}\pi_{Cournot} + \ldots = \frac{1}{1+r}\pi_{cheat} + \frac{1}{1+r}\frac{1}{r}\pi_{Cournot}. \tag{3.12}$$

A firm would be willing to follow the trigger strategy if doing so gives it as least as great a value as cheating, that is, if

$$\frac{1}{r}\pi_{collude} \geq \frac{1}{1+r}\pi_{cheat} + \frac{1}{1+r}\frac{1}{r}\pi_{Cournot} \tag{3.13}$$

or (rearranging terms somewhat)

$$\frac{1}{r} \geq \frac{\pi_{cheat} - \pi_{collude}}{\pi_{collude} - \pi_{Cournot}}. \tag{3.14}$$

We know that $\pi_{cheat} > \pi_{collude}$; in a single period, a firm will earn a greater profit by cheating on a scheme to restrict output if no other firms cheat. We also know that $\pi_{collude} > \pi_{\text{Cournot}}$:

monopoly profit is greater than Cournot profit. Thus the right-hand side of (3.14) is positive. If the interest rate r that is used to discount future income is sufficiently small, each firm will have a greater present-discounted value if it does not cheat on the scheme to restrict output.

In the formulation presented here, if a firm cheats on a scheme to restrict output, the cheating is detected after one period of cheating, and other firms retaliate immediately. If it takes more than one period for retaliation to take place, then it is more profitable to cheat, and noncooperative collusion is less likely to be stable.[4] It follows that one way to increase the stability of noncooperative collusion is to put arrangements in place that make it easier for firms to detect and verify cheating, if it should occur.

Stepping back from the details of the trigger strategy, overall output restriction creates a tension between the profit that any single firm could gain by maximizing its short-run payoff and the future profit to be lost if all firms behave in the same way. Market characteristics that reduce the profit to be gained from short-run individualistic behaviour make output restriction more stable.

Noncooperative collusion: if the present-discounted value of the income to be gained over the long run by restricting output and raising price to maximize joint profit is greater than the present-discounted value of the income to be gained by maximizing short-run own profit and accepting tougher future behaviour from rivals as a result, it will be in the self-interest of each firm to restrict output and raise price.

3.2.2 Basing point pricing

The basing point pricing system is one arrangement that reduces the profit a colluding firm might gain by short-run individualistic behaviour.[5] The basing point system has been used in many markets for which the product is of relatively low value in relation to weight or volume and transportation cost is a large part of the total cost of getting the product to the consumer—cement, oil, plywood, and steel are examples.[6]

Under a single basing point system, one location is the basing point and the price a consumer pays for delivery from any location is the price at the basing point plus the cost of transportation from the basing point, no matter what the location of the plant from which delivery is actually made. With a multiple basing point system, the delivered price

[4] If it takes m time periods for other firms to detect cheating and then revert to Cournot behaviour, the value of a firm that cheats, (3.12) becomes

$$\left[\frac{1}{1+r}+\frac{1}{(1+r)^2}+\ldots+\frac{1}{(1+r)^m}\right]\pi_{cheat}+\frac{1}{(1+r)^m}\frac{1}{r}\pi_{Cournot}.$$

This would be substituted on the right-hand side of (3.13) to give a generalized form of the stability condition for noncooperative collusion with a trigger strategy.

[5] See Loescher (1959), Phlips (1983, 1993).

[6] The early steel-sector pricing policies of the European Coal and Steel Community compelled ECSC steel firms to use basing pricing, in effect enforcing a collusive scheme for the steel industry; see Section 10.3.1.

to any location is the base price plus transportation cost from the basing point that offers the lowest delivered price, no matter what the location of the plant from which delivery is actually made.

Figure 3.2 (which follows Phlips, 1983) shows a stylized version of a market with two firms, one located at basing point *I* and one plant located at basing point *II*.[7] mc_1 is marginal production cost at location *I*, mc_2 is marginal production cost at location *II*. The lower curves rising from location *I* show the marginal cost of delivery from location *I*, the lower curves rising from location *II* show marginal cost of delivery from location *II*. p_1 is the price to a consumer located at *I*, and the upper curves rising from *I* show the delivered price from plant *I* to consumers at various distances from the plant at *I*. p_2 is the price to a consumer located at *II*, and the upper curves rising from *II* show the delivered price from plant *II* to consumers at various distances from the plant at *II*.

Over the interval from *I* to *a*, the lowest delivered price is from *I* and this price is below *II*'s marginal cost of delivery. Thus it would not be profitable for the plant located at *II* to make sales to consumers located between *I* and *a*. Between *a* and *b*, the delivered price from *I* is lower than the delivered price from *II*; it is the delivered price from *I* that is the basing point price. Between *b* and *c*, the delivered price from *II* is lower than the delivered price from *I*; it is the delivered price from *II* that is the basing point price. Between *a* and *c*,

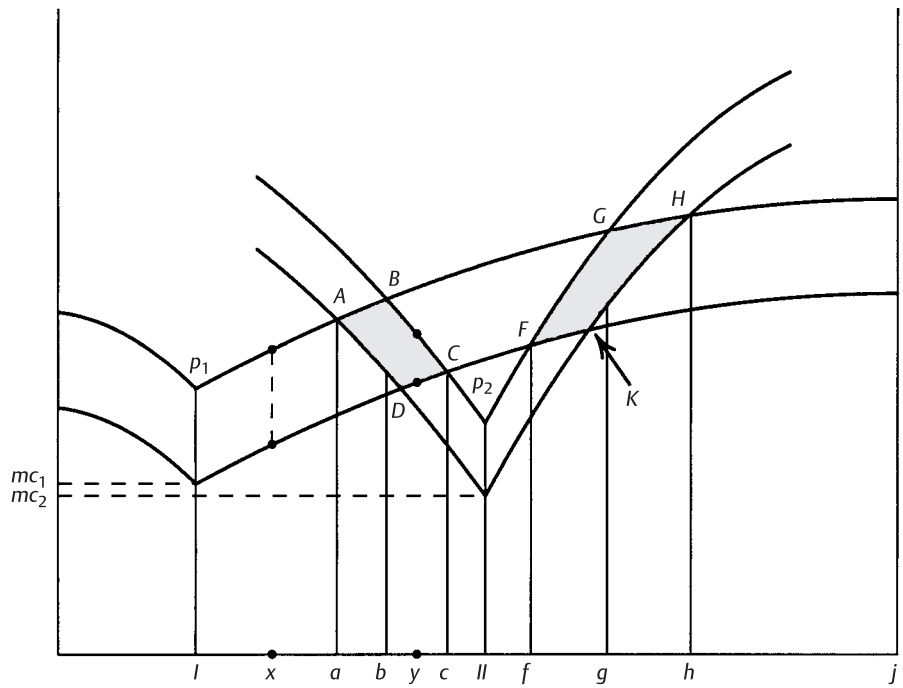

Fig. 3.2 Multiple basing point pricing

[7] More generally, there might be plants not located at a basing point: the delivered price from such a plant would be the delivered price from the basing point that offers the lowest delivered price.

the basing point price is greater than marginal delivery cost for both firms, so both firms would find it profitable to sell to consumers located between *a* and *c*.

Between *c* and *f*, the delivered price from *II* is lower than the delivered price from *I*; it is the delivered price from *II* that is the basing point price. This price is less than the marginal cost of delivery from *I*, so the plant located at *I* would not find it profitable to make sales to consumers located between *c* and *f*. By similar arguments, both firms would find it profitable to sell to consumers located between *f* and *h*, and only the plant located at *I* would find it profitable to sell to consumers located to the right of *h*.

Basing point pricing creates an appearance of rivalry, in the sense that there are some regions (*a* to *c*, *f* to *h*) where customers have a choice of more than one supplier. Basing point pricing also helps firms avoid tough price competition. With basing point pricing, it is easy for any firm to calculate what the delivered price to any location is supposed to be: since businesses make their price lists available to consumers, price lists will become available to rivals as well, without any need for direct communication between firms.[8] Transportation (for what are typically bulky, heavy products) will often be by rail or barge, and the rates for these modes of transportation will normally be publicly available. Since it is easy to compute the basing point price, if any firm were to charge a lower price, the act of price discounting would be obvious to any rival that learned of it, and in regions where both firms would be willing to make sales, consumers seeking similar discounts could be counted on to pass along information about discount prices offered to themselves or to other customers. Thus any one firm considering the possibility of cutting price (possibly only to a few select customers) would expect other firms to quickly detect and match the price cuts, reducing the extra profit that might be gained by cutting price.

With basing point pricing, all firms offer the same price for delivery to any point, but firms discriminate between consumers in terms of net price—price minus marginal cost. In Figure 3.2, for example, firm *I*'s marginal profit supplying a customer located at *x* is $p_1 - mc_1$, while its marginal profit supplying a customer located at *y* is something less than that: the difference between the delivered price from *II*, p_2 plus the cost of shipment from *II*, and mc_1 plus the cost of shipment from *I*. Sales at *x* and *y* are both profitable for firm *I*, but firm *I* makes a greater net profit on sales to customers located at *x* than to customers located at *y*. When firms have different costs of supplying different locations, identical delivered prices mean discriminatory net prices.

3.2.3 Publicity

The theory of noncooperative collusion highlights the rapid detection of price cuts as a factor reducing the expected profitability of price cuts and, therefore, making it more likely that noncooperative collusion will be stable. It may seem surprising, therefore, that an early approach to fighting collusion was to rely on a policy of price transparency price (Interparliamentary Union, 1930):

> cartels, trusts and other analogous combines are natural phenomena of economic life towards which it is impossible to adopt an entirely negative attitude. Seeing, however, that those combines

[8] In modern times, price lists might very well be posted on the internet.

may have a harmful effect both as regards public interests and those of the State, it is necessary that they should be controlled. This control . . . should simply seek to establish a supervision over possible abuses and to prevent those abuses.

An efficacious means of fighting such abuses and a basic condition for eventual control is to be found in publicity, which implies an obligation for cartels and similar combines to announce their existence and to register in the books of the state.

One rationale for the belief that publicity would bring better market performance seems to have been the idea that publicity would shame managers of collusive firms and lead them to alter their behaviour (Marshall, 1923, p. 442):

upright men are often half-way converted towards removing such just grounds as there may be for complaint against their conduct, by reading a well-informed and well-balanced statement of those grounds; and knowing that an impartial public is forming its judgment on them. In this and many other ways a careful authoritative inquiry, with publication of the evidence taken, goes a long way towards removing sources of social harm . . .

This would not now be found to be a convincing argument.

A rationale for a publicity policy against collusion rooted in a less optimistic but perhaps more realistic view of human nature can be found in the idea that publicity would alert rivals to profit possibilities, inducing entry and improved market performance (Jenks, 1900, pp. 223–4):

publicity . . . which should show with a reasonable degree of detail the profits of the larger combinations would, in the case of the abuse of their power, so stimulate competition against them, either actually or potentially, that consumers would to a great degree be protected against excessive increase in prices.

A publicity policy was practiced in many European nations in the early and middle 20th century, and gradually abandoned.[9] It survived in the Danish Competition Act of 1990 (since amended), which provided that:

The purpose of this Act is to promote competition and, thus, strengthen the efficiency of production and distribution of goods, services, etc., through the largest possible transparency of competitive conditions and through measures against restraints on the freedom of trade and other harmful aspects of anti-competitive practices.

In 1993, the Danish Competition Council suspected that collusion was taking place in the market for ready-mixed concrete, a product that is heavy in relation to value and which is normally shipped no more than 20 to 30 kilometres from the plant.[10] At this time, there were 115 production sites in Denmark. The largest firm had a national market share in 1987 of 37 per cent; the second largest firm had a national market share of 11 per cent. Since geographic markets were local, many local markets were very close to being monopolized.

Implementing a transparency policy against the suspected anticompetitive practices, the Danish Competition Council began to collect price data—transaction prices, taken

[9] A publicity policy was also pursued in the US, before passage of the 1914 Federal Trade Commission Act; see Scherer (1990), Martin (1998a).

[10] This discussion is based on Albæk et al. (1997).

from invoices—and publish them on a quarterly basis. Prices rose 15 per cent to 20 per cent in the first year after publication of the transaction prices, and the variability in prices over time fell.

During this period, there was no particular boom in the construction industry, the major user of ready-mix concrete. Thus it does not seem possible to explain the price increase in terms of demand factors. Average concrete prices increased more, in percentage terms, than the price of cement, a major ingredient of concrete. The most likely explanation for the price increase is that by publishing actual transaction prices, the Danish Competition Council reduced incentives for firms to secretly cut prices. With published transaction prices, price cuts would quickly be revealed, inviting retaliation. In effect, the Danish Competition Council made it easier for the ready-mix concrete industry to avoid price competition. In a perfectly competitive market, no transparency policy is needed. In an imperfectly competitive market, the effects of a transparency policy are perverse: it makes noncooperative collusion easier, because it makes it easier for firms to detect rivals' output expansion.

Price transparency and market performance: in imperfectly competitive markets, business or government policies that make it easier for firms to detect sporadic price cuts make such price cuts less profitable and facilitate tacit collusion.

3.3 Welfare consequences of collusion

If firms collude on the monopoly price and the number of firms supplying the market is held constant, collusion means that market performance changes from that of oligopoly to that of monopoly: the price rises, output falls, and deadweight welfare loss increases. This follows from the fact that monopoly output is less than equilibrium oligopoly output.

Additional effects arise if colluding firms succeed in raising price but cannot prevent new firms from coming into the market. Suppose, for example, that in a market with inverse demand curve (3.1) and cost function (3.2), fixed cost F is 550.

Then Cournot duopoly profit per firm per period is positive:

$$\pi_1 = \pi_2 = (40 - 10)(30) - F = 900 - 550 = 350 > 0\,. \tag{3.15}$$

With three firms in the market, Cournot triopoly profit would be negative:

$$\left(100 - 3\,\frac{100 - 10}{4} - 10\right)\left(\frac{100 - 10}{4}\right) - 550 = -43.75 < 0\,. \tag{3.16}$$

With Cournot behaviour, the equilibrium number of firms is two.

If three firms collude on monopoly output, each firm produces 15 units. The monopoly price is 55; hence each of the three firms earns positive profit:

$$(55 - 10)(15) - 550 = 675 - 550 = 125 > 0\,. \tag{3.17}$$

In this particular example, collusion on the monopoly output raises the equilibrium number of firms from two to three. More generally, if firms collude and raise the price, the higher price may make additional entry profitable, depending on the size of fixed costs.

In the example, if two incumbent firms collude on the monopoly price and as a result a third firm comes into the market, the first two firms earn less profit than they would as Cournot duopolists: monopoly profit split among three firms is 125 per period, Cournot duopoly profit per firm is 350. It might very well be, however, that firms could find ways to restrict output and raise price without violating anticollusion laws (we take up this topic in greater detail later) but that actively trying to keep a third firm out of the market would leave evidence of other kinds of anticompetitive behaviour (abuse of a dominant position) against which it would be more difficult to defend. It might simply be, in the oft-quoted words of Hicks (1935, p. 8), that "The best of all monopoly profits is a quiet life".

If raising price to the monopoly level would make entry profitable, incumbent firms might noncooperatively collude to raise price above the Cournot level, but not so high as the monopoly level. Noncooperative collusion would still worsen market performance, but not so much as if entry were not possible. Thus we have the possibility that potential entry—the threat of competition, rather than actual competition—might improve market performance.[11]

A kind of natural experiment generated by UK competition policy suggests that collusion may in fact result in excess entry (Symeonidis, 2000). Before 1956, agreements to restrain trade were not illegal under UK competition law, but they could not be enforced in court. Trade associations in a number of industries administered elaborate schemes to exchange information about prices and quantities sold, schemes which appear to have had the effect of making it easier for rivals to detect output expansion or price cutting.[12] As suggested by our discussion of noncooperative collusion, making it easier for rivals to detect cheating should discourage defection from a collusive scheme.

In 1956 the UK Restrictive Trade Practices Act, made restrictive agreements illegal, and required that existing restrictive agreements be registered with public authorities. The Act left open the possibility that a Restrictive Practices Court could permit an agreement if it found benefits that outweighed the negative effects of a restrictive agreement.

Immediately after the law went into effect, it was not known what approach the Restrictive Practices Court would take, and it appears that most existing restrictive agreements were in fact registered. As it developed, the Court took a tough approach to applying the law, and did not permit many restrictive agreements. Information exchange agreements ceased to function, and between 1958 and 1977 the combined market share

[11] We also have the possibility that one equilibrium oligopoly price is the highest price that will not induce entry, in contrast to the long-run equilibrium price of a perfectly competitive industry, which is the lowest price that will not induce exit.

[12] See Chandler (1990, Part III) for case studies documenting the inclination of British industrialists to avoid tough competition.

of the five largest firms in industries affected by the Restrictive Practices Act rose by an average of 15 percentage points. The most likely explanation is that in the UK, trade association activities supported noncooperative collusion but did not restrict entry, thus permitting the survival of relatively inefficient firms. When trade association activities were limited by a more vigorous competition policy, rivalry increased and less efficient firms went out of business.

Collusion or tacit collusion and market performance: successful tacit collusion or tacit collusion raises price, leaving consumers worse off than would otherwise be the case, and may allow excess entry and the continued operation of inefficient firms, leaving producers worse off as well.

3.4 Article 81

3.4.1 Provisions

It is Article 81 (see box, below) of the EC Treaty that deals with collusion and cooperative arrangements. Article 81, Paragraph 1 prohibits agreements and decisions by firms or associations of firms and concerted practices that affect trade between the Member States and have the object or effect of distorting competition, on the ground that they are incompatible with the common market.[13] Paragraph 2 declares that such agreements are void. Paragraph 3 allows the European Commission to permit restrictive agreements, if the agreements have certain types of specified beneficial effects (improving static efficiency or technological progress) and if a fair share of those benefits are passed along to consumers. Under Article 81(3), the Commission has, for example, sometimes permitted agreements of rival firms to specialize in the production of different varieties of a product, to set up joint marketing arrangements, or to cooperate in research and development.

3.4.2 Applications[14]

As we have noted above, Article 81 refers to "agreements and decisions by firms or associations of firms and concerted practices". Vital practical questions are:

[13] To be precise, Article 81 refers to undertakings rather than firms; "undertaking" is a classification that includes private firms but also other organizations that supply goods and services to markets (for example, public enterprise). A completely different meaning of the word "undertaking" in the context of EC competition policy is an agreement made by firms to close an investigation before a negative decision is formalized.

[14] See, generally, Neven et al. (1998).

Article 81 of the EC Treaty (ex Article 85)

1. The following shall be prohibited as incompatible with the common market: all agreements between undertakings, decisions by associations of undertakings and concerted practices which may affect trade between Member States and which have as their object or effect the prevention, restriction or distortion of competition within the common market, and in particular those which:

(a) directly or indirectly fix purchase or selling prices or any other trading conditions;

(b) limit or control production, markets, technical development, or investment;

(c) share markets or sources of supply;

(d) apply dissimilar conditions to equivalent transactions with other trading parties, thereby placing them at a competitive disadvantage;

(e) make the conclusion of contracts subject to acceptance by the other parties of supplementary obligations which, by their nature or according to commercial usage, have no connection with the subject of such contracts.

2. Any agreements or decisions prohibited pursuant to this Article shall be automatically void.

3. The provisions of paragraph 1 may, however, be declared inapplicable in the case of:

— any agreement or category of agreements between undertakings;

— any decision or category of decisions by associations of undertakings;

— any concerted practice or category of concerted practices, which contributes to improving the production or distribution of goods or to promoting technical or economic progress, while allowing consumers a fair share of the resulting benefit, and which does not:

(a) impose on the undertakings concerned restrictions which are not indispensable to the attainment of these objectives;

(b) afford such undertakings the possibility of eliminating competition in respect of a substantial part of the products in question.

- what sort of evidence must be produced to establish, in a legal sense, the existence of an agreement, decision, or concerted practice?;[15] and
- what is the distinction, in terms of standards of proof required, between an agreement and a concerted practice?

Franco-Japanese ball-bearings

Given the historically charitable attitude of EC Member States to collusion, it is not surprising that the most clear-cut examples of collusion are found early in the history

[15] To find a violation of Article 81, it must also be shown that the behaviour complained of "may affect trade between Member States" and has the object or effect of preventing, restricting, or distorting competition within the common market.

of EC competition policy.[16] One such example is the Franco-Japanese ball-bearings case.[17] Representatives of Japanese and French trade associations and ball-bearing producers met in Paris in 1972 and agreed that Japanese firms would raise their prices in Europe from 15 per cent below the prices of European firms to not more than 10 per cent below, with a later increase to only 5 per cent to 8 per cent below the prices of European firms. The French participants in the meetings prepared careful minutes, which were obtained by the European Commission during the course of its investigation. Based on this and other evidence, the Commission concluded that the agreements violated Article 81(1) of the EC Treaty.

Suiker Unie

Experience quickly teaches business executives that it is not wise to leave such unambiguous evidence of cartel meetings.[18] The difficulty this creates for competition policy is that in oligopoly industries, the kinds of market performance that might be obtained by naked collusion can also be obtained by genuinely independent behaviour. If prudent managers learn not to engage in obvious collusion, but the standards of proof required to find existence of a concerted practice are not as strict as those required to find existence of collusion, one might think that the Article 81(1) prohibition of concerted practices would make it possible to effectively promote independent market behaviour. Thus the question of the distinction between the evidence needed to show that firms have reached an agreement and the evidence needed to show that firms have engaged in a concerted practice moves to the fore.

The Suiker Unie case,[19] which involved business conduct that had the effect of splitting the EC sugar market along national lines, sheds light on the concepts of agreement and concerted practice. The decision records ample evidence of meetings and direct communications among the firms involved, so that it most likely would have been possible to find a violation of Article 81 on the basis of a restrictive agreement alone. But the case also found the presence of concerted practices, which the Court defined as ([1975] ECR 1663 at 1696) "a form of coordination which knowingly substitutes practical cooperation for the risks of competition". The Court also said that the notion of a concerted practice ([1975] ECR 1663 at 1942; emphasis added):

> must be understood in the light of the concept inherent in the provisions of the Treaty relating to competition that *each economic operator must determine independently the policy which he intends to adopt on the common market* including the choice of the persons and undertakings to which he makes offers or sells.

Although it is correct to say that this requirement of independence does not deprive economic

[16] During the interwar and indeed early post-World War II periods, as well as in the initial phases of the European Coal and Steel Community, cartels were seen as an effective way to organize sectors of the economy which (it was believed) would otherwise have been subject to politically unacceptable fluctuations. See Gillingham (1995), Morelli (1997).

[17] OJ L 343 21 September 1974, p. 19.

[18] Such episodes nonetheless continue to surface from time to time (the Archer Daniels Midland case under US antitrust law being one example), suggesting (as if evidence were needed) that greed and stupidity are not mutually exclusive human characteristics.

[19] *Suiker Unie and others* v. *Commission*, Joined Cases 40–48/73, 50/73, 54–56/73, 111/73, 113–114/73 Judgment of 16 December 1975 [1975] ECR 1663; [1976] 1 CMLR 295.

operators of the right to adapt themselves intelligently to the existing and anticipated conduct of their competitors, *it does however strictly preclude any direct or indirect contact between such operators, the object or effect whereof is either to influence the conduct on the market of an actual or potential competitor or to disclose to such a competitor the course of conduct which they themselves have decided to adopt or contemplate adopting on the market.*

The first part of this discussion affirms that independent behaviour does not violate Article 81. The second emphasizes the importance of direct or indirect contacts that alter the incentives a competitor faces when it makes its own decisions. Such contacts are critical to a finding that there is a concerted practice. They might also, however, be thought to justify a finding that firms had acted in agreement, raising the question of the distinction between an agreement and a concerted practice.

Italian flat glass

Early decisions of the European Commission suggested that the distinction between the two concepts lay in the nature of the proof required to establish a violation of Article 81. In the Italian flat glass decision,[20] for example, the Commission relied heavily on evidence that Italian producers of flat glass (used in the windows of buildings and automobiles) set essentially identical price schedules, changed those price schedules at essentially the same times, and granted discounts from price schedules in essentially the same way to justify its finding that there was a concerted practice (OJ L33, p. 61):

the fact remains that the publication of identical price lists over a long period, the existence of the same discount scales and the application of uniform terms of sale to the same customers could only be the result of concerted practices either directly between the three producers or through the intermediary of the spokesman for the wholesalers.

It must be said that viewed through the lens of economic theory, this analysis is exactly wrong. In theory, firms that operate in the same market over long periods of time could all independently decide to engage in the kind of parallel behaviour that is described (sustained by the kinds of independent reactions modelled by a trigger strategy). On a more practical level, it is precisely over long periods of time—during which firms have an opportunity to study rivals and acquire experience that allows them to anticipate rivals' conduct—that parallel patterns of behaviour might be expected to emerge in oligopolistic markets as a result of entirely independent decisions.

Woodpulp and the oligopoly problem

In its Woodpulp decision,[21] the European Court of Justice moved away from accepting parallel conduct as sufficient evidence of a concerted practice. Wood pulp is an input in the production of paper. During the period involved in this case, the European Community was supplied with wood pulp by firms located in North America and Northern Europe. The European Commission relied on several factors to justify its conclusion that the firms involved in the case had engaged in a concerted practice in violation of Article

[20] Commission Decision of 7 December 1988 OJ L 33 4 February 1989, p. 44.

[21] *A. Ahlström OY and others* v. *EC Commission* [1988] 4 CMLR 901; [1993] 4 CMLR 407; see also Commission Decision of 19 December 1984 OJ L 85 26 March 1985, p. 1.

81(1). Some US producers were members of an export cartel.[22] Some firms were members of a trade association, based in Switzerland, that hosted regular meetings at which firms exchanged information about prices and formulated price policies. The Commission also relied on evidence of parallel behaviour: the prices set by different firms changed by more or less the same amount at more or less the same time, even though the firms involved were based in many different countries and kept their accounts in many different currencies.[23]

The Advocate General's arguments when the case was appealed, later accepted by the European Court of Justice, rejected the parallel pricing approach ([1993] 4 CMLR 407 at 470):

> Each company is entitled to align itself independently on the conduct of its competitors, if knowledge of such conduct is obtained solely by monitoring the market.

and ([1993] 4 CMLR 407 at 478):

> In any event, if there is a plausible explanation for the conduct found to exist which is consistent with an independent choice by the undertakings concerned, concertation remains unproven.

The European Court of Justice found that while parallel price movements might have been the result of collusion, they might also have been the result of a combination of a high degree of price transparency—prices widely known and news of price changes circulating rapidly—and the oligopoly structure of the market. For the Court, evidence of parallel pricing in combination with other factors might justify a finding of concertation, but in this instance the other evidence assembled by the Commission was not sufficient.

This is the heart of the problem that oligopoly presents for the enforcement of competition law: if the number of firms is small, independent action can lead to results that closely approximate those of collusion. But competition law in free-market economies does not condemn the results of decisions that are independently arrived at, and long experience suggests that neither regulation nor public ownership are likely to result in improved (or even good) market performance. Thus even if firms in an oligopolistic market really are colluding in a legal sense, it may be difficult to obtain legally acceptable evidence of that collusion.

There is also the problem of remedy. If firms are colluding, they can be ordered not to collude, and they may be subject to large enough fines so that they decide it is in their own best interest not to collude. But firms in oligopolistic industries will often be able to avoid tough competition without colluding; equilibrium market performance in oligopoly may be much the same with or without collusion. But if such an outcome is produced by independent business decisions, there is not much competition authorities can do about it, taking market structure as given.

A clear implication of the oligopoly problem for competition policy is that if there are

[22] The US Sherman Act of 1890 prohibits collusion with respect to the US market (and thus corresponds, in a general way, to Article 81). But the 1918 Webb-Pomerene Act makes it legal for US firms to collude with respect to export markets (see Section 9.2.2). The EC Treaty does not cover export cartels that do not affect trade between the Member States. However, the European Commission has been sceptical toward the possibility that firms could collude on export markets without some spillover effect on the EC internal market.

[23] OJ L 85 26 March 1985, p. 16.

practical difficulties in influencing business conduct in oligopoly markets, competition authorities should take great care administering policy toward market structure. Proposed mergers should be carefully evaluated to determine if they are privately profitable because they increase efficiency or because they increase market power.

Trade association activities

Trade associations may serve legitimate purposes (development of guidelines for product standardization, for example, or dissemination of aggregate information about demand trends, which helps firms make effective investment decisions) but they frequently serve as a forum for activities that restrict competition. When they do, they will run foul of Article 81.

One example is found in the Commission's FEG/TU decision.[24] FEG was an association of Dutch wholesale distributors of electrotechnical equipment (cables, plugs, switches, sockets) used in construction. FEG members supplied 96 per cent of the Dutch market for electrotechnical equipment. TU was the largest single member of FEG.

FEG administered informal exclusive dealing arrangements between manufacturers, its members, and the construction firms that were the main direct customers for the products distributed by FEG members. Dutch and foreign manufacturers for the most part agreed to distribute only through FEG members; construction firms for the most part agreed to obtain supplies only through FEG members. This informal agreement was the successor to a formal agreement that had been in place from 1928 to 1959.

Given the large share of the market covered by FEG members, the exclusive dealing arrangements made it difficult for firms based outside the Netherlands to make significant sales in the Netherlands. Combined with the fact that one requirement for FEG membership was annual sales revenue in the Netherlands of 5 million NLG for three years in a row, the effect of the exclusive dealing arrangements was to create artificial barriers to entry around the Dutch wholesale market for electrotechnical equipment.

FEG subgroups, organized along product lines, carried out regular meetings at which information about price lists, discounts, and prices actually paid by customers. Some of the product subgroups took active measures to keep supplies out of the hands of price-cutting wholesalers based in the EC but outside the Netherlands. The result was that the wholesale price level for the products in question in the Netherlands tended to be higher than elsewhere in the EC.

In its decision under Article 81, the European Commission concluded that FEG and TU had violated Article 81, ordered them to end the restrictive behaviour, and fined FEG €4.4 million and fined TU €2.15 million.

[24] OJ L 39 14 February 2000, p. 1. FEG is an acronym for Nederlandse Federatieve Vereniging voor de Groothandel op Elektrotechnisch Gebied; TU is an acronym for Technische Unie. See Ferdinandusse (2000).

3.5 Enforcement

3.5.1 Leniency

The Commission's *Leniency Notice*[25] creates incentives for members of collusive agreements to repent and return to the path of righteousness. A participant in a cartel may benefit from a reduction of from 10 per cent to 100 per cent of the fine that would otherwise be imposed, depending on the nature of its cooperation and the extent to which its evidence contributes to Commission handling of the matter.

3.5.2 Cooperation

The European Commission's *Draft Guidelines* (27 April 2000b)[26] take an economic approach to the analysis of inter-firm cooperation. Leaving naked collusion to restrict output and raise price aside, inter-firm cooperation may have some efficiency effects that improve market performance and some anticompetitive effects that worsen market performance. The Commission's view is that if the combined market share of cooperating firms is relatively small (less than 20 per cent for specialization agreements, less than 25 per cent for R&D cooperation), beneficial effects of cooperation are likely to outweigh anticompetitive effects, since effective competition from other suppliers will restrict the exercise of market power by cooperating firms. In such cases, the cooperation qualifies for an automatic block exemption[27] under Article 81(3).

If the cooperating firms have a combined market share that is greater than the threshold for a block exemption, the Commission may nonetheless grant an exemption under Article 81(3). In this case, factors to be considered in evaluating the likely impact of the proposed cooperation on market performance, in addition to the market shares of the firms involved, include the degree of seller concentration, the nature of entry conditions, whether suppliers or buyers are likely to exercise countervailing power, and other elements of market structure. If the conclusion of the analysis is that the net effect of the cooperative agreement is positive and that a fair share of the benefit will go to consumers, the Commission will grant an exemption. In evaluating the tradeoff between increased efficiency and market performance, priority is given to market performance: even great efficiency gains do not justify the complete elimination of effective competition (Draft Guidelines, p. 37).

3.5.3 Direct applicability

The Commission (1999a, 2000c) has also proposed a fundamental change in the way Article 81(3) is enforced. At the beginning of the European Community, the Commission

[25] Commission Notice on the non-imposition or reduction of fines in cartel cases OJ C 207/4 18 July 1996. Peña Castellot (2001) reviews early applications of the Leniency Notice.

[26] See also Lücking (2000), Lücking and Woods (2001).

[27] Under Regulation 19/65.

had to introduce a prohibition-based competition policy to a Community of six Member States, only one of which (West Germany) had a similar national competition policy. The other Member States either had no explicit competition policy or followed an abuse control approach that viewed collusion as a natural and inevitable outcome of market processes.[28] The theory of an abuse control policy was that cartels would register with the government, which would keep tabs on their behaviour and take steps to prevent excesses. In practice, abuse control often turned into "industry capture'', with cartels using government policy to promote collusive outcomes.

Faced with the need to introduce a culture of competition in the fledgling European Community, the Commission established a system under which cooperating businesses were required to notify the Commission of their agreement, and only the Commission could decide whether or not an exemption would be granted.[29]

Particularly in the early years of EC competition policy, some agreements were not notified to the Commission, and these were condemned if and when the Commission discovered their existence. But the Commission was faced with a flood of notifications that required some response and overwhelmed the Commission's resources. One reaction of the Commission was to develop the block exemption system. Block exemptions are formal policy statements outlining broad classes of agreements, typically based on market share thresholds and other conditions, that automatically qualify for an Article 81(3) exemption. Another response was to issue so-called comfort letters to notifying firms. A comfort letter is a statement that on the basis of the information before the Commission, the notified arrangement either does or does not appear to meet the conditions for an exemption under Article 81(3). The Commission issued comfort letters on an informal basis, and the European Court of Justice held that they did not constitute legally binding decisions.

Today the foundations of a culture of competition in the EC seem firmly established, and a good deal of the supporting superstructure as well. All EC Member States have national competition authorities; most have adopted national competition policies with provisions that parallel those of the EC Treaty (Martin, 1998). The number of Member States is expected to double in the foreseeable future. The advent of the euro and monetary integration is likely to broaden and deepen integration and market integration. The increase in the number of Member States and progress in market integration will both increase the workload of EC competition authorities.

The Commission has therefore proposed to eliminate the requirement that it be notified of cooperative arrangements. Article 81(3) would be enforced by national authorities and before national courts, as is now the case for Article 81(1).[30] The Commission would then be able to concentrate its attention and resources on the most serious competition policy matters, leaving routine enforcement to national authorities.

A necessary condition for the success of this reform is that Article 81(3) be enforced in a

[28] Historically, the predominant European approach to market behaviour has what Gillingham (1995, p. 163) calls "organized capitalism. Its characteristic features are the cartel, the producer association, and close ties to the state." See Lister (1960, pp. 126–9), and also Chamberlin (1954).

[29] Regulation No. 17/62 JO 13/204 21 September 1962 [1959–62] OJ spec. Ed. 57.

[30] In some Member States, national law would need to be amended to implement this approach. The Commission would have the power to intervene in matters decided at the Member State level if it feels that such intervention is appropriate.

consistent and vigorous way in the different Member States. Competition policy continues to be an essential force in the ongoing process of EC integration, and it will not fulfil its role if weak enforcement in some Member States undermines overall confidence in its effectiveness.

3.6 Market definition for the application of competition policy

If cooperating firms have a combined market share that falls below the limits specified in the Draft Guidelines, they will be able to go forward with their cooperation without having to seek and obtain explicit permission. This approach reduces compliance costs for business and enforcement costs for the Commission. It also brings the process of deciding what the market is to centre stage: businesses need to know how the Commission thinks markets should be defined, for the purposes of competition policy, to decide if they need to seek explicit exemption for a cooperation agreement. As we shall see, the process of market definition is central to the application of almost all elements of competition policy.

For the Commission (1997b, p. 1):

> The main purpose of market definition is to identify . . . the competitive constraints that the undertakings involved face. The objective of defining a market in both its product and geographic dimension is to identify those actual competitors of the undertakings involved that are capable of constraining their behaviour and of preventing them from behaving independently of an effective competitive pressure.

In principle, the Commission emphasizes the demand side when it defines markets (1997b, p. 2): "A relevant product market comprises all those products and/or services which are regarded as interchangeable or substitutable by the consumer" and (p. 3) "the exercise of market definition consists in identifying the effective alternative sources of supply for customers of the undertakings involved, both in terms of products/services and geographic location of suppliers". The primary theoretical test that the Commission applies is one based upon how consumers would respond to price changes (1997b, p. 4):

> The question to be answered is whether the parties' customers would switch to readily available substitutes or to suppliers located elsewhere in response to an hypothetical small (in the range 5%–10%), permanent relative price increase in the products and areas being considered. If substitution would be enough to make the price increase unprofitable because of the resulting loss of sales, additional substitutes and areas are included in the relevant market. This would be done until the set of products and geographic areas is such that small, permanent increases in relative prices would be profitable.

Implicitly, this methodology takes it for granted that products are differentiated; otherwise, there would be only one price, and it would not make sense to speak of "relative prices" or "prices in products and areas being considered". In the real world, of course,

most markets do involve differentiated products. The existence of product differentiation has implications for how market share should be measured for the purpose of assessing market power; see Problem 3–1. If products are differentiated but market size and market shares are calculated as if products were standardized, the result will be to underestimate market shares.

Although the Commission clearly gives priority to demand-side substitutability in the principles for defining markets, it also indicates that it will take supply-side substitutability into account if suppliers are (1997b, p. 4) "able to switch production to the relevant products and market them in the short term without incurring significant additional costs or risks in response to small and permanent changes in relative prices". In such cases, for the Commission (1997b, p. 5) "the relevant product market will encompass all products that are substitutable in demand and supply, and the current sales of those products will be summed to calculate the total value or volume of the market".

To the extent that it is motivated by economic models of oligopoly, such a procedure is incorrect. Economic models clearly indicate that the total value of the market (adjusted for product differentiation, if it is present) is the sum of sales of products that are sold in the market, and does not include the sales of products that could be sold in the market if the firms producing those products found it profitable to switch their activities from one market to another (Problem 3–1). It is demand-side substitutability that determines where market boundaries are. Supply-side substitutability (entry conditions) determines the extent to which a firm with a large market share will find it profitable to exercise market power.[31]

It is possible to carry out sophisticated econometric analyses of demand patterns to reach conclusions about the nature of the market. The time and data requirements of such tests often mean that they cannot be applied by the Commission when it defines a market. The Commission often looks at practical information about the nature of demand for a class of products (1997b, pp. 8–10): historical substitution patterns, the views of customers and competitors, evidence of costs of switching from one brand or supplier to another, whether there are distinctive national preferences or national (for example, tax or environmental) policies.

3.7 Summary

Equilibrium output in an imperfectly competitive market generally exceeds the monopoly level, which gives firms in such markets an incentive (the potential for increased profit) to restrict output. It is possible (depending on the tradeoff between long-term

[31] Consider the following stylized example: a single local firm supplies a standardized product in a national area. The firm's average cost is €10 per unit of output, and if there were no possibility of entry, the monopoly price would be €12. But if price is higher than €11 per unit of output, foreign suppliers with capacity sufficient to supply the entire local area would find it profitable to enter the national market. If the local firm sets a price of €11 or less, it is the only supplier and has a 100 per cent share of the national market: but the threat of entry makes holding price below the unconstrained monopoly price the most profitable option. For a price of €11 or less, including the sales that foreign suppliers could make at higher prices would give a misleadingly lower evaluation of the local firm's market share. See Gaskins (1971) for a formal model.

gains from output restriction and short-term gains from own-profit maximization) that it will be a noncooperative equilibrium for individual firms to restrict output. If firms reach such an equilibrium as a result of meetings or of exchanges of information, trade association activities, and some types of contracts with distributors, they can be found to have agreed or engaged in a concerted practice in violation of Article 81 of the EC Treaty. But if firms reach such an equilibrium in a genuinely noncooperative way, as a result of independent decisions independently arrived at, then they will not have colluded (made agreements in restraint of trade) in the sense of EC competition law.

Study points

- trigger strategy (page 50)
- basing point pricing (pages 53–5)
- price transparency and market performance in imperfectly competitive markets (pages 55–7)
- collusion or tacit collusion and the survival of inefficient firms (page 58)
- EC Treaty, Article 81 (page 59)
- competition policy and tacit collusion in oligopoly (page 63)
- enforcement of Article 81 (pages 65–7)
- market definition in applications of competition policy (pages 67–8)

Problem

3–1 (Measuring market share with differentiated products) Show that if (for example, for duopoly) inverse demand curves have equations (2.41) and (2.42),

$$p_1 = 100 - (q_1 + \theta q_2), \tag{3.18}$$

$$p_2 = 100 - (\theta q_2 + q_1), \tag{3.19}$$

the expression for the Lerner index of market power that corresponds to (2.34) is

$$\frac{p_1 - c'(q_1)}{p_1} = \frac{s_1}{\varepsilon_{Q_1 p_1}}, \tag{3.20}$$

where

$$s_1 = \frac{q_1}{q_1 + \theta q_2} \equiv \frac{q_1}{Q_1} \tag{3.21}$$

is firm 1's market share, taking account of the imperfect substitutability of variety 2 for variety 1, and

$$\varepsilon_{Q_1 p_1} \equiv -\frac{Q_1}{p_1}\frac{dp_1}{dQ_1}. \tag{3.22}$$