Chapter 4. Collusion and horizontal agreements

Massimo Motta
European University Institute, Florence,
Universitat Pompeu Fabra, Barcelona, and CEPR, London
e-mail: massimo.motta@iue.it
INCOMPLETE DRAFT

Competition Policy: Theory and Practice
Cambridge University Press, 2003

This draft: 21 December 2002

Abstract

Contents

1 Introduction

Collusive agreements can take different forms: firms might agree on sales prices, allocate quotas among themselves, divide markets so that some firms decide not to be present in certain markets in exchange for being the sole seller in others, or coordinate their behaviour along some other dimensions. Institutional arrangements to sustain collusion might also differ: from a very well organised cartel-like structure where a central office (secret, if anti-trust laws exist) takes the main decisions, to situations where firms just find some form of communication to sustain the agreement. Further, a collusive outcome might be sustained even in a situation where firms never meet to discuss prices or never exchange sensitive information (but I shall argue that in such a case, labelled "tacit collusion", the law should not intervene.)

Collusive practices allow firms to exert market power they would not otherwise have, and artificially restrict competition and increase prices, thereby reducing welfare.[1] Accordingly, they are prohibited by any anti-trust law, and a large part of the antitrust authorities' efforts are devoted to fight such practices. However, if any serious anti-trust authority would certainly attack a cartel or an explicit agreement among competitors to set prices, or share markets, there might be divergences as to the standard of proof required in less blatant infringements of the law, and as to the treatment of cases where firms manage to keep prices high without overtly colluding.

The main purpose of this chapter is to identify the main mechanisms behind collusion, to study the factors which facilitate it, and to explain which behaviour should be treated as an infringement of the law and which one should not be. I shall also analyse what actions anti-trust authorities should take in order to deter and break collusion.

The chapter is structured in the following way. Section 1.1 briefly sketches the main features of collusion from an economic point of view. Section 2 investigates the industry features, contractual characteristics, and other factors that make collusion more likely to occur. These two sections are theoretical in nature, and draw on the existing and well-developed industrial organisation literature on collusion. Section 3 is also theoretical, but it deals with technical issues and is intended for advanced readers only. Section 4 deals instead with the 'practice' of collusion: I will indicate what should be the legal standards for finding firms guilty of collusion, and what actions to take

[1] See chapter 2 on the relationship between market power and welfare.

to deter and break collusion. The final section is devoted to the discussion of other horizontal agreements, that is, agreements among competing firms which do not concern the fixing of prices, outputs or markets. This is the case, for instance, when firms collaborate in research and development activities.

1.1 What is collusion?

...first, a warning In this section I will briefly characterise the concept of collusion from the point of view of industrial economics. Note that here I will not use the term "collusion" as synonymous of "collusive agreement that should be outlawed". Indeed, one of the main themes of this chapter is that whereas in economic theory collusion is defined as a market outcome (i.e., "high prices", in a sense to be specified further below), anti-trust authorities and judges should consider as illegal all practices where firms *explicitly* coordinate their actions to achieve a collusive outcome.

The reader should then be aware that in sections 1.1 and 2 I will use the term collusion in a way that adheres to economic theory but might be slightly misleading, as it comprehends both explicit and tacit collusion, whereas the law should punish only explicit agreements (as section 4 will argue).[2]

Defining collusion in economic theory In economics, collusion is a situation where firms' prices are higher than some competitive benchmark.[3] A slightly different definition would label collusion as a situation where firms set prices which are close enough to monopoly prices.[4] In any case, in economics collusion coincides with an outcome (high enough price), and not with the specific form through which that outcome is attained. Indeed, as I

[2]I will argue that antitrust authorities should also deal with tacit collusion, but in a preventive way, that is trying to eliminating business practices that might facilitate collusion, and through merger control.

[3]In technical terms, the benchmark is usually the equilibrium price of a game where firms meet only once in the marketplace (a situation where collusion would not arise). For instance, in a homogenous goods game where firms choose prices, a collusive outcome would exist whenever prices are higher than the one-shot Bertrand equilibrium price; where firms choose quantities, whenever they are lower than the one-shot Cournot equilibrium quantities.

[4]See Kühn (2001). Monopoly, or joint-maximisation prices, are the prices which would be set if all the firms in the industry were affiliates of the same company, or where managed by the same manager.

explain below, collusion can occur both when firms act through an organised cartel (explicit collusion), or when they act in a purely non-cooperative way (tacit collusion).

What are the main ingredients of collusion? It is not easy for firms to achieve a collusive outcome, even if they are free to agree on they prices they set. In particular, every firm would have the temptation to unilaterally deviate from a collusive action, as by doing so it would increase its profit.

Consider for instance an imaginary industry consisting of two fruit sellers in a street market. Imagine they both sell pears of identical quality, and that they each pay 1$ per kilo to their suppliers. Imagine also that each seller thinks that 2$ per kilo is the monopoly price, and believes the other thinks in the same way. When a seller arrives at his stall, he has to decide the sales price. Suppose that he thinks the rival is setting a price of 2$. If he charges 2$ for his pears, he will get roughly half of the buyers, as people who want to buy pears are indifferent between buying from him or from the other vendor. But he will have a strong temptation to *deviate*, that is to charge a lower price than his rival: if he sets a price of, say, 1.9$, consumers will all buy from him (why paying more for an identical product?). As a result, he will still enjoy a high unit margin but he will sell more units: in a word, he will make more profits than if he sold at the "collusive" price of 2$.[5]

The acknowledgment that any collusive situation naturally brings with it the temptation to *deviate* from it and therefore to break collusion, leads us to the identification of the two elements which must exist for collusion to arise. First, its participants must be able to *detect* in a timely way that a deviation (a firm setting lower price or producing higher output than the collusive levels agreed upon) has occurred.[6] Second, identifying the deviation is not enough: there must also be a *punishment*, which might take the form of rivals producing much higher quantities (or selling at much lower prices) in the periods after the deviation, thus depressing the profit of the deviator.[7]

[5] Of course, for this simple example to hold it is necessary that at the collusive price of 2$ the seller does not manage to sell all the pears he comes to the market with. Otherwise, he would not have an incentive to cut his price in order to increase sales.

[6] Detection of a deviation is not always easy: in many markets, firms' prices and outputs are not directly observable. Stigler (1964) was probably the first to underline this problem, and its consequences upon the likelihood of collusion.

[7] Note that a punishment should be thought of as a more aggressive market behaviour, and not as a direct monetary (or physical!) punishment. Note also that generally a

Only if a firm knows both that a deviation will be identified quickly and that it will be punished (i.e., it will have to forego enough profits because of the market reaction of the cartel members), might it refrain from deviating, so that the collusive outcome will arise.[8]

To continue our example, after having seen why a fruit seller has a temptation to cut prices below the collusive level of 2$, let us see under which conditions he will deviate. If the street market is small enough, and if the sellers post the prices of the fruit they sell, detection of the price cut will be immediate. After the price cut has been identified, one can bet that a seller who has so far sold at the price of 2$ will immediately retaliate, and likely will start to sell at a price lower than 1.9$ per kilo. The result will be a price war which will reduce the profit of both. A seller contemplating a deviation will certainly expect that the rival will retaliate. As a result, the prospect of selling for much of the day at very low prices will deter him from deviating in the first place. In other words, the awareness that a deviation will be easily detected and that a market punishment will ensue, will make each seller refrain from deviating and convince him to stick to the collusive price instead.

To summarise, for collusion to occur, firstly, there must be the possibility to *detect* deviations from a collusive action in a timely way. Secondly, there must be a credible *punishment* which follows a deviation.

It is important to stress that in the example, the two fruit vendors do not talk to each other, neither directly nor through intermediaries: collusive prices will arise through purely non-cooperative behaviour of the sellers. In other words, if detection of deviations is rapid, and if (market) punishments of deviations are likely and credible, then *tacit collusion* can arise: firms do not necessarily have to talk to each other, let alone agree on complicated schemes, for a collusive outcome to be sustainable. All is needed is the awareness that a deviation will be identified, and that a "punishment" will follow.

---

punishment also hits the punishing firms, and not just the deviating firm, precisely because it has to rely on market mechanisms (a low price affects all the firms' profits). It is therefore crucial that firms are willing to take part in the punishment.

[8] In turn, this implies that collusion can be sustained only if firms meet repeatedly in the marketplace. Otherwise, a punishment cannot take place. In technical terms, collusion will never arise in a one-shot game. This is why collusion should be modelled through dynamic (repeated) games.

Which collusive price? A difficulty in the example above is that it is not clear how the "collusive price" is chosen. Imagine that, for some reason, each seller thinks that the other would set a price of 1.5$, rather than a price of 2$. Then, again a collusive situation might occur in equilibrium, but this time with sellers setting a price lower than the monopoly price. In other words, the collusive mechanism I have described works for many different prices and results in firms getting quite different levels of profits.[9]

This result raises the important issue of *coordination*. Firms that are tacitly colluding might arrive at the fully collusive price, but this is just one of the many possible equilibrium outcomes (one of these also being the competitive outcome, i.e., the one-shot game equilibrium price). So, is there an outcome that is more likely than the other? And, since firms have an interest in coordinating on an outcome with the highest possible prices, how can they achieve that outcome?

Coordination: The difference between tacit and overt collusion Under *tacit collusion*, it is difficult for the firms to solve the coordination problem. If firms cannot communicate with each other, they can make mistakes, and select a price (or a quantity) which is not jointly optimal for the firms, and which might be difficult to change. Using the market to signal intentions to coordinate on a different price might be very costly. If a firm believes the "right" price for the industry is higher and increases its own price to signal it, it will lose market share in the adjustment period. If a firm decreases its own price to try and coordinate on a lower equilibrium price, this move might be understood as a deviation and trigger a costly price war. Therefore, experimenting with price changes to coordinate on another collusive equilibrium might be too costly.

Under *explicit collusion*, instead, firms can talk to each other and coordinate on their jointly preferred equilibrium without having to costly experiment with the market. Furthermore, if there are some shocks which modify market conditions, communication will allow the firms to change to a new collusive price without the risk of triggering a period of punishment.

Suppose for instance that, in the example above, one seller knows that demand for pears has decreased, so that he thinks the optimal price is now

[9] The 'folk theorem' (Friedman (1971)) says that in games with infinite horizon if the discount factor is large enough, firms can have any profit between zero and the fully collusive profit at the "collusive" equilibrium). See section 2.5.3 for a simple formal argument.

lower, say 1.8$. Absent communication with the other vendor, our seller faces a problem: if he reduces the price to 1.8$, as new market conditions suggest, collusion might break. Indeed, the rival vendor might have a different perception of market demand, and/or misinterpret the new low price as a "deviation", and start a price war as a punishment. However, if he sticks instead to the usual price of 2$, he will make lower profits, because demand is lower.

Explicit collusion avoids this problem: our vendor could simply tell his rival that he thinks it would be better to decrease the price, and communication will allow them to decide on a new price that suits them both, without risking any price war or a lengthy adjustment period.

Market allocation (or market-sharing) schemes - according to which a firm sells in a certain region (or serves customers of a certain type), whereas the rivals sell in other regions (or serve customers of a different type) - whether achieved by explicit collusion or historical accidents, have the advantage of allowing for prices to adjust to new demand or cost conditions without triggering possible price wars. A market allocation scheme avoids the possibility that, if a shock reduces production costs or market demand, a price reduction might trigger a price war. As long as each firm does not serve segments of demand (explicitly or tacitly) allocated to rivals, prices can change without the collusive outcome to break. This probably explains why such collusive schemes are often used.[10]

I shall come back to the question of communication and coordination among firms in section 2.2.2, and on why competition policy should focus on explicit collusive practices (that is, when some communication and coordination exist) in section 4. Before doing that, however, it is appropriate to study collusion more deeply, and in particular to study the factors which facilitate collusion. This is done in the following section 2.

2 Factors that facilitate collusion

The analysis of collusion in modern industrial economics is based on the so-called *incentive constraint* for collusion: each firm compares the immediate gain it makes from a deviation with the profit it gives up in the future, when rivals react. Only if the former is lower than the latter will the firm choose the collusive strategy. In general, collusion is more likely to arise the lower the

[10] I am grateful to Joe Harrington for stressing this point.

profit that a firm would obtain from deviating, the lower the expected profits it would make once the punishment starts, the more weight firms attach to the future (i.e., when the "loss from deviation" occurs). (See section 2.5.1 for a formal analysis.)

In this section, I briefly review the factors that are more often mentioned as those which foster collusive outcomes, and explain under which circumstances one should expect they have that role. The discussion (see section 2.5 for the formal analysis) is based on the framework delineated above (that is, the condition that says that a firm is better off colluding than deviating): if a given factor relaxes the incentive constraints of the firms, then it facilitates collusion; if it makes it more binding, it hinders it; if the effect is ambiguous, then the factor does not have a clear impact on collusion.

There are two main practical reasons behind the exercise made in this section. First, it is important to identify the factors that facilitate collusion so that anti-trust authorities can intervene so as to eliminate them whenever possible. Second, in some cases, especially in merger analysis, one has to evaluate whether a particular industry is prone to a collusive outcome or not. Studying the industry and assessing whether there are factors likely to lead to collusion becomes crucial.

Somehow arbitrarily, and mainly for the ease of presentation, I divide the study of facilitating factors in different categories.

2.1 Structural factors

Concentration Other things being equal, collusion is the more likely the smaller the number of firms in the industry. The comparison between gains and losses from deviations illustrates why this is the case. Imagine that there are many firms of identical size and of large capacity which co-exist in the industry. At a collusive situation, each of them will set a high price and get a (small) share of the total profits. However, if one of them deviates and sets a price lower than the rivals, it might get all the market for itself. Even if the punishment was harsh, so that a very small stream of expected profits would follow after a deviation, the gains from deviating would be so extraordinarily large in the deviation period that they would outweigh the collusive profits foregone during the punishment period. Compare this situation with the extreme one where there are only two firms in the industry. At a collusive equilibrium, they would get half the market, so that the gains from deviating are smaller relative to the lower profits due to the punishment which follows.

If firms are symmetric, a lower number of firms is equivalent to a higher degree of concentration, which is therefore associated - ceteris paribus - with more likely (tacit or explicit) collusion.

However, we shall see below that the more firms are asymmetric (in capacities, market shares, costs, or product range) the less likely collusion will be. This qualifies the finding that concentration facilitates collusion, in the following sense. If a measure of concentration rises with the asymmetric distribution of assets among the firms - as is the case with the Herfindahl-Hirschman Index (see chapter 3) - then one should expect an ambiguous relationship between concentration and collusion: such a measure confounds two factors - higher average market share and asymmetry - that affect collusion in opposite ways. If instead the measure of concentration used does not vary with asymmetry - like for the concentration ratios, C_k, that sum the market shares of the k largest firms in the industry - then an increase in measured concentration should correspond to higher likelihood of collusion.

Finally, note that the concentration also helps firms' coordination on a collusive outcome, not only its enforcement: the lower the number of players in the industry the easier for them to coordinate their behaviour.

Entry The easier entry into an industry (the lower entry barriers) the more difficult to sustain collusive prices. When prices and profits are high, new firms will be attracted into the industry, and this tends to disrupt the collusive outcome, by two possible mechanisms. Suppose first that the entrant does not want to pursue a collusive strategy and behave aggressively.[11] This will subtract market shares to the incumbent firms that will have to decrease prices to keep their customers, thereby breaking the collusive equilibrium. Anticipating that entry might occur, the incumbents will be forced to keep prices low.

Suppose instead that both the entrant and the incumbent firms follow an accommodating strategy, with the entrant taking part in the (explicit or tacit) collusive behaviour. Since the larger the number of firms the less likely that collusion can be sustained, entry might break the collusive outcome. All the more so since if a new firm does enter and takes its share of the industry

[11] Such a firm might be a so-called "maverick". A maverick might also be a firm that is substantially different from the rivals (for instance because it has a smaller discount factor, a different utilisation of capacity or a smaller portfolio of brands). See the discussion on symmetry and collusion.

collusive profits, more entrants will be induced to follow the same strategy, and sooner or later collusion will be unsustainable.

Overall, therefore, one should expect that the lower entry barriers (as determined by fixed entry costs that new firms would have to sunk into the industry) the more difficult to sustain collusion. Nevertheless, the existence of potential entrants might not always necessarily break collusion. Suppose for instance that an entrant expects the incumbent firms to react very aggressively to entry.[12] If the threat of such a strong reaction by the incumbents is credible, the entrants might decide not to enter in the first place.[13]

Cross-ownership and other links among competitors If a firm has a participation in a competitor, even without controlling it,[14] the scope for collusion will be enhanced. First and more obvious, if a representative of a firm is sitting in the board of directors of a rival firm, it will be easier to coordinate pricing and marketing policies. It might also be easier to exchange information on the marketing and pricing policies, which makes it easy monitor rival's behaviour and - as I discuss below - is an important facilitating factor for collusion. Second, even if a firm did not have any saying in the business policies of the other, but just owned a share of it without representation in the board, the incentives to compete in the marketplace might be reduced. This is because the profits of the rival firm would affect the firm's own financial performance, composed of market profits and financial returns: an aggressive market strategy (like a deviation from a collusive price) would be less profitable than if there were no stake in the rival firm, because

[12]There is no reason to believe, a priori, that incumbents would always accommodate, or always fight, entry. In a nice empirical study, Scott Morton (1999) analyses the British shipping cartels at the turn of the last century and finds that some entrants were accepted and joined the cartel, whereas others (the weaker entrants) were fought.

[13]Gilbert and Vives (1986) analyse the case where joint predation occurs without the need for predators to coordinate. In Harrington (1989b), entry triggers the same (strongest possible) punishment as a deviation from collusive actions by the incumbents, hence reacting strongly to entry without more coordination than needed to devise the punishment strategies. But if incumbents did coordinate, then they would breach both article 81 and 82 in the EU, and section 1 and 2 of the Sherman Act in the US. The EC's *CEWAL* decision provides an interesting recent case study where incumbents (members of a maritime shipping conference) put in place a sophisticated mechanism to fight entry.

[14]If there is control, then this will be a merger (see chapter 5). A minority shareholding of 10% or less is not subject to EU merger control.

it would decrease the returns on financial investments.[15]

Overall, it would therefore seem wise not to allow a firm to have minority (a fortiori, controlling) shareholding in a competitor.[16]

Regularity and frequency of the orders *Regular orders* facilitate collusion. Indeed, an unusually large order would give a very strong temptation to deviate: by deviating, a firm would make unusually large profits, and the perspective of losing collusive profits obtained under the typically small expected demand is not enough to deter the deviation. The high *frequency of the orders* also helps collusion because it allows for a timely punishment. If orders arrive only at very large distance between them, one has a higher incentive to deviate because the punishment will be started only much later in the future, and will accordingly be discounted.

Buyer power The ability to sustain collusive prices in a given industry also depends on the degree of concentration of the buyers. A strong buyer can make use of its bargaining power to stimulate competition among the sellers, either by threatening to withdraw orders from a current seller to give them to others or to potential entrants (whose entry prospects would be rosier if guaranteed demand by a large buyer), or by threatening to start upstream production itself.

By concentrating its orders, a powerful buyer can also manage to break collusion. Rather than keeping a steady flow of small orders, by grouping them into large and less frequent orders a large buyer can induce suppliers to deviate from the collusive strategy.[17]

Finally, strong buyers might design procurement auctions so as to minimise the risk of collusive behaviour among suppliers (see section 4.2 for a discussion of auction design).

[15]See exercise **??** to formalise this idea. Similar considerations might also apply to cases where competing firms have common interests - such as joint ventures - in certain markets or products. See also Martin (1995).

[16]Ritter et al. (2000: 553-4) say that it is not clear whether minority shareholdings are compatible with article 81 of the Treaty, although "the acquisition of 10% or less in a *competitor* which is not associated with a representation in the board or veto rights may be considered as a *purely passive* shareholding for investment purposes".

[17]See Snyder (1996).

Demand elasticity Although the elasticity of market demand is a factor that is sometimes mentioned as facilitating collusion, it is not clear why it should affect the *likelihood* of collusion. If demand is very elastic, then a given price cut will determine a large increase in the quantity demanded, but this is true both for the price cut in a deviation and for the price cut in the punishment period. In other words, elasticity of demand will in general affect both sides of the incentive constraint for collusion, and its net effect on sustainability of collusion is ambiguous.

It is true, however, that demand elasticity will affect the *level* of the maximum collusive price (the lower the elasticity of demand the higher the monopoly price), which implies that there will be less reason to worry about possible collusion if demand elasticity is high.

Evolution of demand The impact of demand evolution over time upon collusion depends on the nature of demand shocks.[18] Suppose that current demand conveys no information about future demand - that is, demand shocks are independently and identically distributed, as in Rotemberg and Saloner (1986). Then, it is as if a large order suddenly arrived, and the analysis would be as above: firms would break collusion to capture the profit of unusually large demand.[19] (Conversely, collusion will be more likely if a negative shock occurs.)

Suppose instead that demand movements are correlated over time, so that today's large order signals that demand will increase steadily. In this case, collusion is more likely: Why should a firm give up the prospect of large future collusive profits for a small gain today, when the market is still small? (Conversely, collusion will be less likely when firms face a future of declining market demand.)[20]

[18] It also depends on whether demand movements are ex-post observable or not, as discussed in section 2.2.1 below.

[19] A market characterised by frequent and drastic innovations will also be less prone to collusion, as less weight is attached to the future (as incumbents anticipate that they might not be competitive in the near future) and the incentive to deviate is therefore higher. See Rey (2002).

[20] There are a number of papers that have studied formally the problem of sustainability of collusion over the business cycle. The intuitive results reported here coincide with the results of Haltiwanger and Harrington (1991), who find that collusion is more likely to break when demand is declining. But Fabra (2001b) extends their analysis by also considering possible capacity constraints, and she shows that Haltiwanger and Harrington's results hold only when aggregate industry capacity is large enough, whereas when it is small

Demand stability might help sustain collusion, to the extent that it increases the degree of observability in the market (see also section 2.2.1). A market characterised by frequent demand shocks or large uncertainty might be one where it is difficult to understand whether poor sales are due to demand variability or to price undercutting of rivals. Accordingly, collusion might be more difficult to sustain. By contrast, in a mature stable market it would be easier to spot deviations and punish them, rendering collusion easier.

Product homogeneity Practitioners, anti-trust authorities and judges often maintain that it is easier to reach collusion with homogenous than differentiated products.[21] Theory is less clear about this point. Suppose that products are differentiated. In this case, it is harder to punish a deviant firm, since even a considerable reduction in prices by rivals would leave the deviant firm with a positive demand. This effect tends to discourage collusion, as only the fear of punishment makes firms refrain from deviating. However, for precisely the same reasons, under differentiated products a deviation is less profitable too: a deviant firm cannot expect to gain very large market shares from rivals unless it makes a very considerable cut in price, an effect that tends to facilitate collusion. Therefore, product homogeneity does not unambiguously raise the scope for collusion.[22] A priori, then, it is not clear that - other things being equal - collusion should be more likely in products like cement and gasoline than, say, cigarettes, colas or mineral waters, which are all consumer goods characterised by a high degree of consumer loyalty.[23]

Possibly, however, a reason why homogeneity might help collusion is that if firms sold not a single well-defined product but very many different product variants, then it would be more difficult for them to attain a collusive outcome. This makes sense, insofar as it reduces the visibility of deviations, and would thus diminish the possibility to resort to quick punishments.[24]

---

enough, collusion is less likely to be sustained when demand increases.

[21] For instance, product homogeneity is one of the recurrent features in European merger cases suspected to give rise to collusion.

[22] See for instance Ross (1992), for a formal analysis.

[23] The large advertising expenditures sustained in these markets explain why these products are perceived by consumers as very differentiated from each other. See Sutton (1991).

[24] An alternative explanation of why product homogeneity might help collusion has been offered by Raith (1996). He argues that different products might be subject to different demand shocks. If product homogeneity means that there is correlation in the demand

Symmetry Competition authorities and courts also regard symmetry among firms as a factor which facilitates collusion. Symmetry can concern different dimensions (such as market shares, number of varieties in the product portfolio, costs and technological knowledge, capacities), whose importance will clearly differ across industries. Many informal arguments exist which support the idea that symmetry helps collusion: for instance, it is intuitive that people who are in a similar position would find it easier to arrive at an agreement which suits all of them. Recently, there have been some formal contributions to the literature which give further support to this idea.[25]

Compte, Jenny and Rey (2002) analyse a model where firms produce homogenous goods and have identical costs, but differ in their capacities. In their model, the largest firm has the highest incentive to deviate from collusion (a firm whose capacity is filled when the price is at the collusive level does not have any incentive to cut the price), and the smaller firms are those with difficulties to punish (since they are capacity constrained, they cannot credibly threaten to punish a deviant firm). A more equal distribution of capacities would help collusion. If firms' positions were more similar, their incentives to deviate and to punish would be more aligned and collusion could be more easily sustained.

In Kühn and Motta (1998), there are multi-product firms that sell a different number of product varieties. The larger the firm (the higher the number of product varieties it sells) the stronger its interest in keeping prices high. Similarly to a firm having a large share of the market, for which a marginal reduction in the price would hurt all the infra-marginal units, here a price reduction in one variety would negatively affect all other varieties (the larger the firm the stronger pricing externalities). Instead, a small firm (that sells only one or few varieties) has a stronger incentive to deviate and free ride on high prices, as by decreasing prices would capture demand from all rivals' products. Therefore, in this model the large firm has more difficulties to punish and the small firms have the larger temptation to deviate from the collusive price.

Note that the mechanisms at work are quite different in Compte, Jenny and Rey (2002) and in Kühn and Motta (1998), but the results are the same: a more equal distribution of assets relaxes the incentive constraints of both

shocks, then this might help collusion.

[25] There also exist some evidence from empirical analysis: see Barla (2000) on firm-size inequality and collusion in the airline industry.

the small and the large firm, and would help collusion.[26]

Multi-market contacts For a long time, it has been argued that multi-market contacts - defined as the same firms meeting in more than one market - would help collusion. Indeed, there also exists some empirical evidence that this is the case. Evans and Kessides (1994) find that airline fares are significantly higher in routes where there exist carriers that have contacts on several routes; Parker and Röller (1997) find that prices tend to be higher in US mobile telephone markets characterised by multi-market contacts.

The suggested explanation used to be that when firms coexist in several markets, then it is more costly for them to deviate from a collusive outcome, since they would be punished in all the markets at the same time. This intuitive argument is incomplete though, and therefore faulty. Indeed, if a firm is present in many markets, it can also deviate in *all of them* at the same time, and this increases its incentive to deviate. A priori, then, the very fact that the same firms coexist in several markets rather than in just one, it might not be sufficient to explain why the latter should help collusion. Indeed, Bernheim and Whinston (1990) show that when firms and markets are perfectly symmetric multimarket contacts do not change the incentives for collusion. It is only when there are asymmetries that multi-market contacts help.

The intuition behind the result that multi-market contacts might help collusion is not straightforward, but an example might help. Suppose that there exist two markets that are asymmetric in the following sense. In market A there are only two firms, call them 1 and 2, that have a similar share of that market. In market B, there are n firms, two of them being 1 and 2 (our multi-market firms), that share equally the market. Suppose that if the two markets were taken in isolation (that is, if in market B the n firms were only single market firms), then collusion would arise in market A but not in market B (we know that collusion is harder to sustain where there are more

[26] Vasconcelos (2001b) confirms these results within a very elegant model where he analyses optimal punishment schemes. He has a homogenous good model where firms differ in their capacities, but contrary to Compte et al. higher capacity implies lower cost. In his paper, large firms have higher incentives to deviate along the punishment path and small firms along the collusive path. Again, symmetry would help collusion. Harrington (1989a) analyses collusion when firms have different discount factors. He shows that when firms are asymmetric they manage to reach a collusive outcome by redistributing market shares.

firms). This is equivalent to saying that, at the prevailing discount factor, the ICs of the multi-market firms in market A are lax, whereas the ICs in market B (taken in isolation) are not satisfied. But the ICs of the multi-market firms are given by the pooling of the two ICs, one for each market. To understand why multi-market contacts help sustain collusion, consider two points. First, if the other $n-2$ firms in market B had a larger share of the market, then their incentive to deviate would be reduced. (Recall that the larger the firm the lower its incentive to reduce its price.) Second, for each multi-market firm, the pooled IC will still hold if the firm sold a bit less in market B. Indeed, accepting a lower market share in market B would increase the incentive to deviate in market B, but this can be outweighed by the fact that the incentive to deviate in market A is low. By pooling the ICs, the slackness of the IC in market A can be used to enforce collusion on the other market.

To further clarify the concept, consider the following example as well. Suppose that in market A a firm has 70% of the market and the other has 30%, whereas exactly the opposite situation occurs in market B. Apart from one firm selling more in one market (and vice versa), the firms have similar characteristics, and the markets present similar features and sizes. Also suppose that, if the two markets were taken in isolation, collusion would not be sustainable (the asymmetry in market shares being too strong). But now the same two firms operate in the same two markets: when they analyse whether they prefer to collude or to deviate, they will look at both markets together. This implies that they will have a single IC where the gains and losses from deviations are pooled across both markets. Effectively, under pooling it is as if there were now two symmetric firms in one large market, each firm holding a 50% share of it. Multi-market contacts here smooth the market asymmetries and by making the ICs of the firms more symmetric they help collusion.[27]

Inventories and excess capacities Generally speaking, the role played by the presence of large levels of inventories and large excess capacity is ambiguous. Suppose for instance that all firms in an industry are endowed with excess capacity with respect to the expected levels of demand. Then this affects both sides of the incentive constraints in a similar way, rendering

[27] Another reason why multi-market contacts might help collusion is that they tend to increase the frequency of the firms' contacts.

ambiguous the final effect on the condition for collusion. On the one hand, large excess capacity implies that there is a stronger incentive to deviate (a price reduction would help fill capacity). On the other hand, if rivals are also endowed with large capacities, the punishment is more likely to be strong.[28]

Since both theory and empirical evidence on this point are ambiguous,[29] it is unclear whether excess capacity helps collusion.

2.2 Price transparency and exchange of information

In this section I first emphasise the role of observability of prices and quantities in sustaining collusion (section 2.2.1). This leads to the discussion of one practice which helps firms improve observability of their actions, namely agreements to exchange information on past and current individual data. I shall also discuss (section 2.2.2) the role of communication among firms. This also helps collusion in that it might help firms coordinate on a particular outcome.

2.2.1 Observability of firms' actions

Detection of deviations is a crucial ingredient for collusion, and Stigler (1964) argued that collusive agreements would break down because of *secret* price cuts. In fact, Green and Porter (1984) show that if actual prices (or price discounts) are not observable, collusion would be more difficult to sustain, but it could still arise at equilibrium. Their important contribution can be summarised in the following way. Imagine an industry where sellers cannot

[28] Of course, this would be different if excess capacities (or large inventories) were unequally distributed. In this case, the arguments made above about asymmetry would apply: the more unequal the distribution of capacities, the less likely that collusion can be sustained.

[29] The first to model the relationship between capacity and collusion were Brock and Scheinkman (1985). Benoit and Krishna (1987) find that excess capacity is needed to sustain collusive outcomes, and Davidson and Deneckere (1990) that the greater the degree of excess capacity (which is determined endogenously in their game, and depends on the cost of capital) the more collusive the prices that can be sustained at equilibrium. These two papers suggest that the punishment effect of excess capacity (the fact that it allows for punishment to occur) is stronger than the cheating effect (the fact that excess capacity makes it more tempting to deviate). However, in a different setting, Compte et al. (2002: 10) find that - for symmetric firms - the higher industry capacity the less likely that collusion can be sustained. Some empirical papers on this point are mentioned in Davidson and Deneckere (1990: 525-526).

observe the prices charged by rivals and where market demand levels are also unobservable. Then, a seller would not know if a lower number of customers served than expected is due to a negative shock in demand or to a price cut by a rival which has stolen some (or all) of his business. Green and Porter show that if the discount factor is high enough, there exists a set of collusive strategies that represent an equilibrium. The strategies are such that each firm sets a collusive price (which might be the price that maximises joint profits) as long as every firm faces a high level of demand. When a firm faces a low (or zero) demand, then the punishment is triggered and each firm sets the one shot equilibrium price for a finite number of periods. After this finite punishment phase, all firms revert to the collusive price.

Therefore, the model implies that collusion can be sustained at equilibrium, but unlike the standard model with perfect observability, collusive prices and profits will never be observed forever, even if no firm deviates. Indeed, the punishment is triggered whenever a low level of demand is observed, and will last for a certain number of periods, after which firms revert to the collusive prices.[30] The model has therefore an important implication. The observation of some periods with low prices is not sufficient to exclude that the industry is at a collusive equilibrium. Rather, price wars simply are the indispensable element of a collusive strategy when rivals' prices and market demand realisations are unobservable.[31]

Since observability of prices and quantities helps firms to reach the most collusive outcomes (under perfect observability, price wars that are costly for the firms would not occur), competition policy should pay special attention to practices that help firms monitor each other's behaviour. One example of such a practice is given by information exchange agreements, that is being discussed next. In section 2.3, I shall also address other pricing practices that increase observability of firms' actions, such as resale price maintenance and best price clauses.

[30] Playing price equal to marginal cost forever, that is an infinite punishment, would clearly be suboptimal here: since the punishment is triggered even if nobody has actually deviated, it would not make sense to condemn the industry to zero profit forever whenever a low level of demand is observed.

[31] On the other hand, as I discuss below, the alternance of high and low price levels is not proof either of a collusive outcome, since an industry at a non-collusive equilibrium might have lower prices under negative demand (or common input) shocks or increased capacities.

Information exchange, I: Data on past or current prices and quantities It is often the case that via trade associations or in other ways, firms in a given industry exchange data on prices, quantities, or other variables such as capacities, customer demand, cost and so on. In the light of the discussion above, it becomes important to identify the collusive potential of such communications among firms.[32]

First, we have seen above that exchange of information on past prices and quantities (or of verifiable information on prices and quantities set in the current period) of each individual firm facilitates collusion, as it allows to identify deviators and better target market punishments, that become then more effective and less costly for the punishing firms.

In the absence of disaggregate information on past prices and quantities, availability of more precise estimates of aggregate (market) demand would also help, as it allows firms to see whether a decrease in individual demand is due to cheating of rivals or to a negative shock in market demand. In turn, this implies that there would be no need of punishment phases which are triggered not by deviations but by a general decrease of market demand.[33]

Exchange of information about past (and current) prices and quantities helps firms sustain collusion, but it is possible that there might also be efficiency effects behind exchange of such data. For instance, better information about demand might allow firms to increase production in markets, times, and areas where demand is higher. The literature on information exchange has ambiguous findings.[34] Theoretically, it is possible in certain circumstances that exchanging information helps welfare. However, it is unlikely

[32] On collusion and exchange of information between competitors, see Kühn (2001).

[33] Porter (1983) shows that exchange of private information about market demand reduces demand uncertainty and allows more collusive outcomes to be sustained. In a similar vein, Kandori (1992) shows that as demand uncertainty decreases, firms can attain higher collusive outcomes (and punishment phases become more severe), and Kandori and Matsushima (1998) also find that communicating information about past realisations helps collusion. Technically, the last paper differs from Green and Porter (1984), Porter (1983) and Kandori (1992) in that it assumes that firms receive private rather than public signals, so that each firm might have a different belief of what has happened in the industry (has there been a demand shock, or has somebody deviated?). Other papers that deal with collusion under imperfect monitoring and private signals are Compte (1998) and Athey and Bagwell (2001).

[34] The incentives for firms to exchange private information, and more importantly the welfare effects of such exchange are not robust, as they crucially depend on whether the firms compete on prices or quantities, or whether the uncertainty concerns costs or demand. See Kühn and Vives (1995) or Raith (JET, 19xx) for surveys.

that firms need to exchange individual and disaggregate data in order to achieve whatever efficiency there might be. Kühn (2001) also argues that information about the industry might help firms devise incentive schemes for their personnel, based on relative productivity, but again, for such schemes to work firms do not need detailed data at a disaggregate level.[35]

Kühn (2001) convincingly concludes that while both types of information exchange help firms to collude, the observation of past and present quantities and prices of firms is a more effective collusive device than the exchange of private information about market demand. Further, if efficiency gains of information exchange exist, they would be reaped already with the exchange of aggregate data. This should lead competition policy to a more severe treatment of agreements concerning exchanges of information about individual prices and quantities (especially the more disaggregate and the more recent). Indeed, his conclusion that communication between firms about such individual firm data should be forbidden is compelling.

2.2.2 Coordination issues: The role of communication

When firms repeatedly meet in the marketplace, if the discount factor is large enough, any price between marginal cost and fully collusive price might be sustained. This raises the issue of which price is likely to arise as the market outcome. Habit, history, or particular events might provide firms with a *focal point* on which to coordinate.

Consider for instance a situation where two firms are told by a regulator that their prices cannot be higher than a certain level, say 100. In this case, this price will provide a clear benchmark (the focal point) for the firms, and one can bet that 100 will be the price that they will set.[36]

History might also provide hints. Many European markets have been protected from foreign competition for a long time, resulting in several national monopolies in many industries. Once tariff and non-tariff barriers started

[35] Some exceptions about detailed data might occur in particular sectors. In banking and insurance, for instance, markets are characterised by asymmetric information. If firms had information about clients' solvency history, this would be efficiency enhancing as it would lessen adverse selection problems and foster competition by helping customers to switch firms. See Padilla and Pagano (1999). Note, however, that although disaggregate, this is not information about prices set or quantities produced by firms.

[36] Schelling (1960) was the first to introduce the notion of focal points (or conventions) and show how they can help people to coordinate.

to fall, this created a potentially pan-European market. However, a situation where each firm stays in its own market without entering foreign ones would provide a good collusive equilibrium, which is just the continuation of something which has happened for a long time. Instead, starting to export might be considered a deviation and might trigger a retaliation in the home market, with rivals exporting in turn. Therefore, the *status quo* might be a focal point, and only when demand and technology conditions substantially change, might firms be tempted to break the current situation.

Whatever the reason, if firms have coordinated in the past on a certain collusive price or divided markets in a certain way, it might be too risky for them to experiment so as to change it. Firms might simply update such a price more or less mechanically with inflation or when raw materials commonly used in the industry become more expensive.

If firms were colluding explicitly they would simply communicate with each other and they could achieve higher collusive prices (provided that firms are symmetric enough, they would have similar preferences over prices), and/or more efficient market sharing rules.[37] But even if they did not overtly collude, they could still try to overcome coordination problems by transmitting information to each other, as I discuss in what follows.

Information exchange, II: Announcements of future prices Announcement of future prices (or production plans) might help collusion, in that it might allow firms to better coordinate on a particular equilibrium among all the possible ones.[38] Farrell (1987) was the first to show the role of non-binding and non-verifiable communication (known as "cheap talk") in achieving coordination among players in games with multiple equilibria.[39] Since then, both theory and experimental evidence seem to indicate that announcements about price intentions might help firms coordinate, although

[37] But explicit agreements might theoretically render collusion more difficult when it allows firms to renegotiate the punishment following a deviation, thereby undermining the collusive mechanism. See the technical section 3.2.

[38] Unilateral announcements help players to select a jointly optimal price, on which it would otherwise be difficult to coordinate if a focal price (that is, an obvious price to be chosen) does not exist.

[39] Farrell (1987) analysed a game with different features from supergames. He looked at a "battle of the sexes" situation, where there are two asymmetric equilibria, as in an industry where at equilibrium only one of two firms could profitably enter, whereas if both entered they would make losses.

not under all circumstances.[40]

However, not all announcements about future actions should be treated in the same way. One should distinguish two different situations, according to whom the announcements are directed: (1) *"private" announcements* directed to competitors only; (2) *"public" announcements* with commitment value to consumers.

"Private" announcements In the first case, announcements are directed to competitors only. To help fix ideas, think of a firm sending a fax to rivals where it is stated that from next month it intends to set a certain price. As Kühn (2001) remarks, it is hard to imagine any efficiency reason behind such announcements. Most likely, they just help rivals to coordinate on a particular collusive price, and therefore helps them collude by avoiding costly periods of price wars and price instability.

Advance notice of price changes, as long as it does not fully commit the firm to the price announced, might also be a tool to avoid costly experimentation with the market.[41] A firm might announce a price increase effective, say, in 60 days, but then revert to the current price if the other firms did not follow suit with similar announcement of price changes.[42] This way, firms might arrive at a commonly agreed price without incurring the risk of losing market shares or triggering price wars during the period of adjustment to the new prices.

Particularly instructive in this respect is the *Airline Tariff Publishers* (ATP) case in the US.[43] ATP is a company owned by the major US airlines whose main purpose is to disseminate price information to airlines and operators, like travel agents, using computer reservation systems. Such information is fed to the ATP by each company, and it contains several elements, such as the fare and the route to which the fare is applied, the possible restric-

[40] See Farrell and Rabin (1996) for a non-technical discussion of the possible role of cheap talk in different games, and of the conditions one should expect it to affect equilibrium outcomes or equilibrium selection. A number of experiments have been performed on this issue, see for instance Cooper at al. (1992). See Kühn (2001) for other references on experiments on the collusive effects of information.

[41] However, advance notice of *effective* price changes could be in the interest of consumers, who might want to know in advance the prices they will have to pay, to reduce uncertainty.

[42] See Hay (1999) for the *Ethyl* case, where this was one of the allegedly anti-competitive prices used by the firms.

[43] See Borenstein (1999) for a detailed account of this intriguing case.

tions to this fare (for instance, which type of consumer can buy it, if advance payment is required, if a minimum number of days of stay is required and so on), *first and last ticket dates* (that indicate the period in which the fare can be sold), and *first and last travel dates* (that indicate when the travel for which the fare applies should take place).

The DOJ alleged that airlines used this information to coordinate price increases without any explicit collusion taking place. For instance, airline A could announce today a price increase on route from city 1 to city 2 and put a first ticket date in thirty days time, so that nobody could actually sell a ticket for that route at the new fare. Since this information was public to all other airlines, airline A could then wait and see the reaction to this price announcement. If its competitor on the same route, say airline B, matched the price increase, then it would be left unchanged and later become effective. But if airline B did not match the price increase, then airline A could still revise the fare. The process could then continue, with airlines adjusting their fares until a convergence is reached, but without consumers ever having the possibility to buy tickets at the announced future fares.[44]

Communication in auctions An interesting example of how firms manage to achieve collusive outcomes through communication is given by their bidding behaviour in *simultaneous ascending auctions*. These are first-price auctions where several objects are for sale at the same time,[45] and the auction ends only when no new higher offer is made on any of the objects. These auctions have good efficiency properties (the value of one object might depend on which other objects are also obtained; information about the value of an object is increased by observing the bids made by rivals), but they are not immune to collusion, as players might use their bids to signal a way to share the objects.[46] For instance, if the government was auctioning off an asset in Milan (whose area code is 02) and other assets elsewhere, I might indicate my interest in the Milan asset by submitting one or more bids for, say, 1,002 euro each. Other bidders will then understand that I am proposing to share objects in such a way that I obtain that particular asset.

[44] The case ended by a settlement, with the firms involved agreeing to discontinue these practices.

[45] The most standard auction is the first-price auction on a single object. Think for instance of the way works of art are often sold: people in a room make higher and higher bids for a particular object, until a bid will not be matched by anybody.

[46] See Klemperer (2001) and Cramton and Schwartz (2001) for a more detailed analysis.

A number of such signaling examples are described in Cramton and Schwartz (2001), who discuss collusion in the FCC auctions in the US, where licenses in different markets were simultaneously for sale. For instance, after a number of rounds where it had not bidden for Amarillo, Texas, the firm Mercury PCS made the higher bid for that market, bumping a rival firm, High Plains Wireless, from it. High Plains had been an aggressive bidder in another market, Lubbock, Texas, and Mercury's bid was intended as a signal to its rival that had it not stopped bidding for Lubbock, it would have retaliated by increasing the price in Amarillo. To make sure that the signal was clear, Mercury's bid for Lubbock had "013" , the code number of Amarillo, as its last three digits, and its bid for Amarillo ended with "264", the code number of Lubbock.[47]

Another case is given by a 1999 auction of ten identical blocks of spectrum in Germany.[48] This was again a simultaneous ascending auction, with the rule that any new bid had to be at least *10 per cent higher* than the previous one. Mannesmann and T-Mobil, the strongest incumbent firms, were the only serious players. In the first round, Mannesmann bid 18.18 million DM per megahertz on blocks 1-5 and 20 million DM on blocks 6-10, and T-Mobil had even lower bids. By making different bids for otherwise identical objects, Mannesmann was signaling the intention to share the market. Further, by choosing 18.18 and 20 it was signaling the fact that it would have been happy with its rival increasing its bid on blocks 1-5 by 10 per cent (the minimum admissible bid increase), so that both would have ended up with five blocks each paid at 20 million DM. Sure enough, T-Mobil's managers got the message (as they admitted in public declarations, see Klemperer (2001: 3)), and in the second round they bid 20 million DM on blocks 1-5, after which the auction ended.

These examples clearly show that communication among players help them reach collusive outcomes in auctions. Therefore, public statements about bidding intentions should not be allowed,[49] and measures might be

[47] See Cramton and Schwartz (2001: 8-9). Such examples of communication of intentions is not restricted to auctions. In the ATP case discussed in the text above, airlines included "footnote designators" attached to a particular announced fare so as to indicate their intentions to punish a competing airline on a particular route: see Borenstein (1999: 314-316) for details.

[48] See Klemperer (2001).

[49] Klemperer (2001) reports episodes where bidders disclosed their objectives to the press, thereby affecting the behaviour of rivals in a straightforward way.

taken to avoid signaling through bids: "For example, bidders can be forced to bid "round" numbers, the exact increments can be pre-specified, and bids can be made anonymous." (Klemperer, 2001: 12)

"Public" announcements In the second case, price announcements are public, and therefore seen by rival firms as well as consumers. Think for instance of a firm advertising the prices of its products on newspapers. On the one hand, it might be argued that transparency of prices still helps collusion, for the reasons indicated above. On the other hand, though, market transparency is good for consumers, as it allows them to "shop around" for the best offer. The latter positive effect is generally considered stronger than the collusive effects of the announcements. Both theoretical arguments and empirical evidence suggest that price advertising in this sense is generally beneficial and brings prices down.[50] Therefore, when prices are "transparent" for both consumers and firms, this should not be considered as an anti-competitive practice.

In some cases, however, prices are made public mainly to increase transparency among rival firms, rather than to the benefit of customers. The *Ethyl* case provides an interesting example.[51] In that case, the four producers of antiknock gasoline additives allegedly used facilitating practices to decrease competition. One such practice was that firms issued press notices on price increases (the other practices were advance notice of price changes, uniform delivered pricing and most-favoured nation clauses). In that industry, there were few large buyers and prices were set by negotiation, and it seems reasonable to think that the press notice did not have the purpose of informing buyers, but of letting rivals know intentions about price changes.[52]

To conclude, whereas announcements directed to rivals only should be forbidden, announcements about current and future prices which carry commitment value vis-a-vis consumers should be regarded as welfare enhancing.

[50] For a survey of both the theoretical and the empirical literature on price advertising, see Fumagalli and Motta (2001).

[51] See Hay (1999).

[52] Note, however, that the FTC was overturned by an appeal court, that found there was insufficient evidence to prove the existence of a cartel.

2.3 Pricing rules and contracts

Most-favoured nation and meeting competition clauses Some clauses in long-term contracts between a seller and a buyer might condition the price paid by the buyer either to the price offered by the seller itself to its other buyers, or to the price offered by other sellers to the same buyer. I look at these two types of clauses in turn, and argue that the latter is more likely to have anti-competitive effects.[53]

A *most-favoured nation* (or most-favoured customer) clause (MFN) engages a seller to apply to a buyer the same conditions offered (by the same seller) to other buyers. The clause can be of two types. A *retroactive MFN* states that the buyer will be offered a price reduction if future buyers will get a lower price for the same good (the reduction being equal to the difference between present and future price); a *contemporaneous MFN* engages a seller to offer a buyer the same price as its other buyers (usually in the same area), and effectively amounts to an engagement not to price discriminate.

These clauses are often said to be anti-competitive, because they make it more costly for firms to give price discounts: a firm that wants to attract new customers must also reduce its price and margins on existing customers. However, it is not clear that a MFN clause would facilitate (tacit or overt) collusion.[54] On the one hand, the clause makes it harder to deviate from a collusive outcome (the additional profit to be gained is smaller because a price reduction would apply to existing customers); on the other hand, it also makes it more costly to carry out a punishment, for precisely the same reason (the punishment cannot be selective, and target only rivals' customers).

Meeting competition clauses are different. They state that if the buyer receives a better price offer from *another* seller, the current seller will match that price.[55] In this case, the potential for collusion is higher, and twofold. First, the clause works as a device to exchange information: whenever a buyer

[53] Salop (1986) was among the first to point out the possible anticompetitive effects of these clauses.

[54] There is a small literature that shows that if a MFN clause is adopted, firms might enjoy higher prices and profits than without. See for instance Cooper (1986), Holt and Scheffman (1987) and Schnitzer (1994). But these are all models where firms play a finite horizon game, and collusion as we have defined it would not arise. In other words, I know of no model where MFN is found to increase sustainability of collusion in an infinite horizon game.

[55] A *meet-or-release* clause gives the seller the possibility to match the price or free the customer from the contract.

is offered a better price, it will have an incentive to report that information to the current seller. This will make firms immediately aware of a deviation from a collusive outcome in the industry, and we know that timely detection of deviations is a crucial element for collusion. Second, the clause reduces the incentives to deviate in the first place: if rivals can retain their current customers due to a meeting competition clause, the price decrease can only attract new buyers, but cannot steal existing buyers from other firms.

Both MFN and meeting-competition clauses might have efficiency explanations, too.[56] First, if buyers are risk-averse, these clauses might have some insurance properties: a retroactive MFN from the possibility that only future buyers will benefit from future shocks; a contemporaneous MFN from the possibility that rival buyers will get the same input at lower prices; a meeting competition clause from the possibility that a price reduction will be offered by other sellers (although, as pointed out above, the very existence of the clause might push other sellers not to offer lower prices!). Second, if gathering information about prices is a costly process, these clauses might speed up purchase since they insure the early buyer that it is not missing better deals. Third, they introduce some price flexibility in long-term contracts, by ensuring that shocks that affect outside options are internalised in the contracts.

There is certainly need of more research on the possible anti-competitive effects of the MFN and meeting-competition clauses. Overall, however, the pro-collusive impact of meeting competition clauses seems so strong that antitrust agencies should probably put them under a per se prohibition rule.

Resale price maintenance Resale (or retail) price maintenance (RPM) is a vertical agreement whereby a manufacturer imposes to its retailer(s) the price at which the good should be sold in the final market. As I will show in chapter 6, there are a number of reasons why RPM can be efficient, and therefore pro-competitive. However, RPM might also facilitate collusion among manufacturers. The intuition is clearly conveyed in the following quote:

> "With a competitive retail market and stable retail cost conditions, manufacturers could assume agreed-upon retail prices by fixing their wholesale prices appropriately. In reality, however, variation over time in the costs of retailing would lead to fluctuating retail prices. If wholesale prices are not

[56] See Salop (1986: 283-284) and especially Crocker and Lyon (1994).

> easily observed by each cartel member, cartel stability would suffer because members would have difficulty distinguishing changes in retail prices that were caused by cost changes from cheating on the cartel. RPM can enhance cartel stability by eliminating the retail price variation."[57]

This story has been recently formalised by Jullien and Rey (2001). In their paper, manufacturers can only sell through local retailers, and there exist local shocks on retail cost or demand. Only the retailer will observe the local shock on the demand for the product it sells, or on its distribution cost. If the retailer was free to choose the price of the final good, the price would be more flexible and would respond to the shock. RPM would therefore be less efficient because it would not allow for adjustment to local conditions. However, RPM implies that prices would be uniformly set by the manufacturers, and therefore allows to better identify deviations from a collusive action, as the quote above suggested. Jullien and Rey (2001) show that whenever the manufacturers will find it optimal to adopt RPM, collusion can be sustained and the result will be welfare detrimental.

Other pricing practices that increase observability When producers are located in different geographic areas, and serve consumers that are also spread out on the territory, it might be difficult for firms to compare prices and to detect price changes, since prices vary with transportation costs. The practice of setting *uniform delivered prices* would then facilitate price observability among rivals. A firm would set the same price inclusive of transportation cost everywhere on the territory, and independent of the customers' locations. Somebody located next to a firm's plant would pay exactly the same as somebody located hundreds of kilometres away.

A similar effect is achieved by *basing point pricing*, a system whereby each producer sets the final price as the mill price at the common basing point (which might be the seat of plants of one or more firms or it might be completely arbitrary) plus transport cost from that point to the final destination. Again, this allows to increase transparency on the producers' side, in that it allows to better compare prices.[58]

[57] Mathewson and Winter (1998: 65).

[58] While delivered pricing and basing point pricing have been claimed to facilitate collusion for a long time, I am not aware of any model that shows rigorously their pro-collusive effect in a supergame context. However, see Thisse and Vives (1992: 257-8) for some considerations on this issue.

2.4 Analysis of collusion: conclusions

The above analysis has identified the main factors which affect the likelihood of collusion. For the purpose of competition policy, this should be useful in two respects. First, it indicates which factors an anti-trust authority should pay attention to if it wanted to try and prevent (explicit or tacit) collusion. Therefore, I shall rely on this analysis in section 4, where I focus on the legal treatment of collusion. Second, in a merger case which raises doubts of coordinated effects (joint dominance), it is important to assess correctly the role played by each variable in determining the likelihood of future collusion in the industry. The analysis will often be very complex. Apart from the cases where all factors point towards the same direction, in general one should expect that the analysis of these factors will leave some space for discretion, as it is difficult to understand how such factors interact and whether collusion is likely to arise from the merger or not.[59]

2.5 Factors that facilitate collusion*

In this section I first formalise the basic model of collusion based on the so-called incentive constraints, then I analyse the factors that facilitate collusion, relying on the standard industrial organisation textbook model.[60]

2.5.1 Conditions for collusion to arise*

Consider an industry where n firms play an infinite horizon game.[61] Call π_i^c and V_i^c respectively the current profits and the present discounted value of profits that firm i receives if it chooses a certain collusive action, given that all firms also collude. Call π_i^d the current profit of firm i if it deviates when all other firms take the collusive action, and V_i^p the present discounted value of firm i's profits in the punishment phase, that is in all periods that follow the deviation period. Denote with $\delta \in (0,1)$ the discount factor, assumed identical for all firms in the industry. Note that the discount factor can be expressed as $\delta = 1/(1+r)$, where r is the interest rate between two periods

[59] See also the section on joint dominance of chapter 5.

[60] See Tirole (1988: ch. 6).

[61] If firms played a repeated game with a finite (and certain) horizon, the collusive outcome would never be attained at equilibrium. At the last stage of the game, all firms would play the Nash equilibrium action, since it is as if they played a one shot game. By backward induction, the Nash equilibrium would be played in each period. See chapter 8.

of time, and therefore indicates how much is worth in today's terms 1 euro that one receives in the following period. Therefore, $\delta = 0$ corresponds to the case where $r \to \infty$: one euro earned in the future is not worth anything in today's terms (people are infinitely impatient, and do not attach any value to the future); at the other extreme, $\delta \to 1$ corresponds to the case where $r = 0$: one euro earned in any future period has the same value as one euro earned today (people are infinitely patient, and they attach equal value to the future as to the present).

Collusion can arise only if each firm will prefer to play the collusive action rather than deviating from it (and be punished). Therefore, it must be that the following *Incentive Constraints* (ICs) hold, one for each firm in the agreement:

$$\pi_i^c + \delta V_i^c \geq \pi_i^d + \delta V_i^p \qquad i = 1, .., n. \tag{1}$$

Clearly, the lower the deviation profit one makes relative to the collusive profit, and the lower the profit in the punishment phase, the more likely collusion will be sustained. (The harsher the punishment the stronger the deterrent to cheating on the collusive agreement.) The n incentive constraints can also be written as:

$$\pi_i^d - \pi_i^c \leq \delta(V_i^c - V_i^p) \qquad i = 1, .., n \tag{2}$$

which states that the gains from deviating obtained today must be lower than the losses from deviating from the collusive strategy, incurred from tomorrow onwards. Again, this condition must be satisfied for all firms, otherwise one or more deviations will occur and collusion cannot be sustained.

Finally, another way to express the same incentive constraint is:

$$\delta \geq \frac{\pi_i^d - \pi_i^c}{V_i^c - V_i^p} \equiv \overline{\delta}_i \qquad i = 1, .., n \tag{3}$$

This is the form in which the condition for sustainability of collusion is most often expressed. Collusion arises at equilibrium only if the discount factor is large enough, that is, if it is larger than a certain "critical discount factor", $\overline{\delta}$. Only if firms are patient enough will the collusive agreement be sustained. This is very intuitive: if the discount factor is very low, firms do not give importance to what will happen in the future, and they will prefer to cheat so as to reap all the benefit they can today. Hence, collusion will not arise.

Note that the conditions identified above hold for both tacit and explicit collusion.

2.5.2 The basic model to analyse facilitating factors*

Consider an industry where n identical firms play a repeated game with infinite horizon (or, which is equivalent, with an uncertain final date). They produce the same homogenous good at the same unit cost c. In each period t of the game, firms set prices simultaneously and non-cooperatively.[62] They all have the same discount factor δ. (If one prefers to consider the game as one where the market exists in the following period with a probability $\phi \in (0, 1)$, take the discount factor to be equal for all and label it d. By setting $\delta = d\phi$, the analysis is equivalent.) Assume also that there are no capacity constraints, and that each firm wants to maximise its present discounted value of profit.

Finally, specify firm i's demand in the following way (which is the usual one for a Bertrand game with homogenous products). If all the firms charge the same price $p_i = p_j = p$, then they all share equally demand, so that $D_i = D(p)/n$ and $\pi_i = \pi(p)/n$, where π_i denotes firm i's profit and $\pi(p)$ is the aggregate profit when all firms charge price p. If firm i sets a price $p_i < p_j$, for any $j \neq i$, then $D_i = D(p_i)$ and $\pi_i = \pi(p_i)$. Finally, if there exists a k such that firm i sets a price $p_i > p_k$, then $D_i = 0$ and $\pi_i = 0$.

Consider the following *trigger strategies*. At the initial period $t = 0$, each firm sets the collusive price p_m (that is, the price that maximises joint profits).[63] At time t, it sets the price p_m if all firms have set p_m in every period before t. Otherwise, each firm sets $p = c$ forever. In words, each firm behaves in a collusive way as long as all others do, but if one of them deviates from the collusive action, then the punishment is triggered and they all revert to the one-shot Bertrand equilibrium for the rest of the game.[64]

[62] Exercise **??** analyses the case of quantity competition.

[63] As I shall discuss below, this is not the only price that can be sustained as an equilibrium of the game, under the same conditions on the discount factor.

[64] This is the harshest punishment: after a deviation occurs, all firms earn zero profit forever: no credible punishment can be harsher than that. Therefore, with homogenous

Collusion arises at equilibrium if no firm has an incentive to deviate from the behaviour indicated by the trigger strategies. Since all firms are identical, we just need to consider the incentive constraint for a firm:

$$\frac{\pi(p_m)}{n}(1+\delta+\delta^2+\delta^3+...) \geq \pi(p_m). \tag{4}$$

The LHS gives the total payoff a firm receives if it colludes (i.e., if it follows the trigger strategy when all other firms do). In each period, and for all periods, the firm receives its share $1/n$ of the aggregate monopoly profit. Profits earned at a time t are discounted by a factor δ^t. The RHS gives the profit under the optimal deviation. If a firm decides to "cheat" when all others collude, the best payoff is obtained by slightly undercutting p_m. By setting $p_m - \varepsilon$, all consumers will buy from the deviating firm, that will thus earn a profit $\pi(p_m - \varepsilon)$. For ε small enough, the firm will therefore get profit very close to $\pi(p_m)$ in the period it deviates. In the following period, however, the punishment occurs as all firms will revert to the Nash equilibrium forever. Therefore, the deviating firm (like all others) will make zero profit in all periods of the game.

Note that $\sum_{t=0}^{\infty}\delta^t = 1/(1-\delta)$. Hence, after simple algebra, the incentive constraint above becomes:

$$\delta \geq 1 - \frac{1}{n}. \tag{5}$$

For $n = 2$ we have the standard textbook (duopoly) case: collusion is sustainable as long as $\delta \geq 1/2$; for $n \to \infty$, collusion is impossible, as a discount factor higher than one would be necessary for the incentive constraint to hold, whereas $\delta \in (0, 1)$.

Concentration From the inequality above it is easy to see that the larger the number of firms n the tighter the incentive constraint, that is the less likely that collusion will be sustained at equilibrium. In a model where firms are symmetric, this is tantamount to saying that an increase in concentration ($1/n$ being an index of concentration) makes collusion more likely.[65]

products and price decisions, no other set of strategies can improve the chances of collusion upon these trigger strategies.

[65] But recall that asymmetries - that might increase concentration measures - hinders concentration: see section 2.1.

Regular orders facilitate collusion To see how regular orders facilitate collusion, it is convenient to show how a *large* order would break collusion. Suppose that we have the same game as above but that at $t = 0$ market demand and market profit are respectively $kD(p)$ and $k\pi(p)$, with $k > 1$. In the following periods, $t = 1, 2, ...$, demand and profit are back to the usual levels $D(p)$ and $\pi(p)$. This is equivalent to saying that there is an unusually large order in one period. The incentive constraint would then become:

$$\frac{\pi(p_m)}{n}(k + \delta + \delta^2 + \delta^3 + ...) \geq k\pi(p_m). \tag{6}$$

This can be rewritten: $\delta \geq (n-1)k/[1 + (n-1)k]$. The IC is the more binding the larger is k, as k increases the RHS more than the LHS. In the limit, if $k \to \infty$, there would be no value of the discount factor that can satisfy the condition for collusion.

Exercise **??** shows that high frequency of market contacts also helps collusion.

Demand elasticity Since demand elasticity enters the IC for collusion only through the expressions of profits, $\pi(p_m)$, and these cancel out, this variable does not seem to have any obvious impact on the likelihood of collusion.[66]

Collusion and the evolution of demand One of the first papers to analyse the effects of the evolution of demand on the likelihood for collusion is Rotemberg and Saloner (1986). In their model, demand can be either in a low (d_L) or in a high (d_H) demand state, and each state occurs with a probability 1/2. The crucial assumption in their model is that shocks are independently and identically distributed: there is no correlation between today's and tomorrow's state of demand. Rotemberg and Saloner (1986) show that collusion is less likely to hold in the good states of demand: firms' temptation to deviate is higher then. Full collusion, defined as firms setting the monopoly price in both states of demand, can be achieved only for very high discount factors. For intermediate discount factors a less than fully

[66] To see it better, write the collusive profit as $(p_m - c)D(p_m)$. The IC becomes: $(p_m - c)D(p_m)(1 + \delta + \delta^2 + ...)/n \geq (p_m - c)D(p_m)$, which simplifies to the usual condition $\delta \geq 1 - 1/n$, where the demand function does not enter.

collusive outcome can be achieved: firms set the monopoly price in low state of demand, and lower prices in the high state of demand.[67]

To understand the mechanism behind Rotemberg and Saloner's paper, note that in their paper the occurrence of a high state of demand is equivalent to a situation where there is an unexpectedly large order today, while tomorrow the market goes down to the usual demand levels. The expected level of demand in their paper is $(d_L + d_H)/2$, so if there is a high demand shock a firm faces a demand higher than the expected: $d_H > (d_L + d_H)/2$. Everything is the same as in the previous section when the usual level of demand is D and there is an unexpected large order such that demand in the current period is $kD > D$.

The introduction of a positive correlation among the demand shocks changes completely the scenario. If the boom is likely to continue in the future, then the incentives to deviate are reduced: the punishment reduces profit in periods of high demand. If instead a decline in demand is expected to be persistent, then collusion is less likely to be sustainable, as one would prefer to deviate now that demand is high, since the cost of the punishment is lower tomorrow that demand is lower. The following example shows these effects.

At time $t = 0$, demand and profit are given by $D(p)$ and $\pi(p)$. At time t, they are given respectively by $\theta^t D(p)$ and $\theta^t \pi(p)$, with $\theta > 0$. The incentive constraint can then be rewritten:

$$\frac{\pi(p_m)}{n}(1 + \delta\theta + \delta^2\theta^2 + \delta^3\theta^3 + ...) \geq \pi(p_m), \tag{7}$$

or: $\delta \geq (1/\theta)(1 - 1/n)$. We now have to distinguish two cases.

(1) (Continuous) demand growth. If $\theta > 1$, it is easy to see that collusion is easier (the IC is relaxed): the expected rise in future demand increases the future cost of a deviation.

(2) (Continuous) demand decline. If $\theta < 1$, collusion is less likely to be sustained (the IC is tightened): the temptation to deviate is stronger because the future cost of deviating, that is the punishment, is lower.

[67] In this sense, there are "price wars during booms". However, note that the actual (lower than monopoly) price set in the high state of demand might be higher than the monopoly price set in the lower state of demand. The term price war just means that firms are not able to set the fully collusive price in the high demand state.

In the real world demand is likely to fluctuate in cycles, with periods of growth followed by periods of decline. However, Harrington and Haltiwanger (1991)'s results on collusion over the cycle are consistent with the picture given by this simpler model: collusion is more likely to break when demand is falling.[68]

Symmetry helps collusion Consider a market A where two firms 1 and 2 operate. Firm 1 has a market share λ and firm 2 has a market share $1-\lambda$.[69] Assume $\lambda > 1/2$: firm 1 is "large" and firm 2 is "small". Firms have the same technology, summarised by a constant marginal cost c, and the same discount factor δ. Firms play the price game for an infinite number of times. Denote the fully collusive price by p_m, and consider simple trigger strategies with Nash reversal forever. The incentive constraint for firm $i = 1, 2$ is given by:

$$\frac{s_i\left(p_m - c\right) Q(p_m)}{1-\delta} - \left(p_m - c\right) Q(p_m) \geq 0, \qquad (8)$$

where $s_1 = \lambda$, and $s_2 = 1-\lambda$ are respectively firm 1 and 2's market shares. The IC of the two firms in market A are therefore given by IC_1^A : $\lambda/(1-\delta) - 1 \geq 0$, and IC_2^A : $(1-\lambda)/(1-\delta) - 1 \geq 0$. The former can be simplified as $\delta \geq 1-\lambda$ and the latter as $\delta \geq \lambda$. Clearly, the binding constraint for collusion in market A is $\delta \geq \lambda$, that of the small firm. The intuition is simple: along the collusive path, the large firm has a larger share of the market, while by deviating each firm obtains (for one period) all the market for itself. Clearly, the incentive to deviate is higher for the small firm, which can capture a higher additional share by decreasing prices. Collusion is limited by the extent to which the market shares are asymmetric. In the symmetric case, where firms have the same market share, we fall back to the standard condition for collusion $\delta \geq 1/2$. The higher asymmetry the more stringent the IC of the smallest firm.

[68] But see Fabra (2001b) for the case of capacity constraints.

[69] The assumption that market shares are exogenously given is simplistic: typically, market shares should be determined endogenously. However, one could interpret this example as the reduced form of a richer game where firms differ in some assets - capacities as in Compte et al. (2002) and Vasconcelos (2001), or number of brands as in Kühn and Motta (1998) - and market shares will mirror the distribution of those assets.

Multi-market contacts Consider now another market B with identical features as market A above, but where the same two firms have reversed market positions. Firm 1 has a market share λ (respectively $1-\lambda$) in market A (resp. B) and firm 2 has a market share $1-\lambda$ (resp. λ) in market A (resp. B), with $\lambda > 1/2$, so that firm 1 is large in market A but small in market B, and vice versa for firm 2.

Consider the ICs for collusion, first for each market in isolation and then for both markets together. We shall see that multi-market contacts help facilitate collusion in this example.

The IC for firm $i = 1, 2$ in market $j = A, B$ *considered in isolation* is:

$$\frac{s_i^j\,(p_m - c)\,Q(p_m)}{1-\delta} - (p_m - c)\,Q(p_m) \geq 0, \qquad (9)$$

where s_i^j is the market share held by firm $i = 1, 2$ in market $j = A, B$. We have seen that the ICs in market A are given by $IC_1^A : \delta \geq 1 - \lambda$, and $IC_2^A : \delta \geq \lambda$. In market B the two ICs are $IC_1^B : (1-\lambda)/(1-\delta) - 1 \geq 0$ and $IC_2^B : \lambda/(1-\delta) - 1 \geq 0$, that reduce to $\delta \geq \lambda$ and $\delta \geq 1 - \lambda$. Firm 1 is the small firm in market B so its IC, $\delta \geq \lambda$, is the binding one.

We can now conclude that by looking at the two markets *in isolation* (or, equivalently, by assuming that firms 1 and 2 operating in the two markets are different) collusion will arise in each market if $\delta \geq \lambda$, where $\lambda > 1/2$.

Multi-market contacts

So far, we have assumed that each firm takes its decision about collusion in each market separately. However, this is not correct: each firm $i = 1, 2$ is selling in two markets, and it will therefore take into account both markets when taking its decisions (recall that if a firm deviates it will deviate in both markets, since after a deviation collusion would be broken in both markets), so its IC is:

$$\frac{s_i^A\,(p_m - c)\,Q(p_m)}{1-\delta} + \frac{s_i^B\,(p_m - c)\,Q(p_m)}{1-\delta} - 2\,(p_m - c)\,Q(p_m) \geq 0, \qquad (10)$$

Both incentive constraints simplify to $\lambda + 1 - \lambda \geq 2(1-\delta)$, whence $\delta \geq 1/2$. In the case where firms 1 and 2 were operating in one market only, collusion would arise if $\delta \geq \lambda$. Multi-market contacts help collusion, since the critical discount factor is lower: $1/2 < \lambda$.

These results follow from the fact that under multi-market contacts firms pool their incentive constraints and can use slackness of the constraint in one market to enforce more collusion in the other.[70] More intuitively, in the example presented here multi-market contacts restore symmetry in markets which are asymmetric. Under multi-market contacts, collusion arises if $\delta \geq 1/2$, which is precisely the same condition faced by symmetric firms in a given (taken in isolation) market.[71]

2.5.3 A problem with supergames: multiple equilibria*

One problem with games where firms play over an infinite horizon is that they admit a continuum of equilibrium solutions. Consider for instance precisely the same basic model as the one analysed in the previous section 2.5 but assume that firms have the following trigger strategies. Each firm sets a price $p \in [c, p_m]$, where c is the marginal cost and p_m the joint profit maximisation price, at $t = 0$. It sets p at period t if all firms have set p in every period before t; otherwise, it sets $p = c$ forever. These strategies, that differ from those studied in the previous section only in that those restricted the collusive price to be $p = p_m$, represent an equilibrium of the game if the following IC holds:

$$\frac{\pi(p)}{n}(1 + \delta + \delta^2 + \delta^3 + ...) \geq \pi(p). \tag{11}$$

It is straightforward to rewrite the IC and check it amounts to $\delta \geq 1 - 1/n$. Therefore, the same value of the discount factor allows for a continuum of equilibrium solutions to hold: any price between the Bertrand price and the monopoly price can be sustained at equilibrium![72]

This result is not entirely satisfactory, since a number of outcomes are possible, and a priori it cannot be said which ones are more likely than others.

[70] Of course, the example given here is extreme and very simplified for illustrative purposes. See Bernheim and Whinston's paper for more general examples.

[71] Spagnolo (1999) assumes firms have a strictly concave objective function (for instance, instead of maximising the sum of profits $\pi_A + \pi_B$, they would maximise $\ln(1 + \pi_A + \pi_B)$) and finds that multimarket contacts facilitate collusion also under market symmetry.

[72] The result that the critical discount factor is independent of the collusive price is an artefact of the Bertrand model. In general, one should expect that the lower the collusive price the easier to sustain collusion (that is, the lower the critical discount factor). This is showed in exercise 11.

More work is needed to understand if some particular equilibria are more likely to be selected. In turn, this calls for more theoretical on the refinements of the equilibrium in supergames, but also for experimental evidence which helps understand how agents in the real world are coordinating on certain equilibria rather than others. See also section 2.2.1 for a discussion on the practical implications of this multiplicity of equilibria.

3 Advanced material**

3.1 Credibility of punishment and optimal penal codes**

I have so far not discussed the credibility of punishments once a deviation takes place, but this is a key issue for both tacit and explicit collusion. If firms were not be willing to carry out the punishment, that is entering a phase of low prices, once a deviation is actually observed, then the threat of the punishment would not exist, and collusion could not be sustained. Formally, this implies that for collusion to be sustained other incentive constraints (one for each firm) should hold, that is that a firm prefers to engage in the punishment action (i.e., retaliate) rather than to deviate from it (i.e., not retaliate).

There are (at least) two different technical solutions for the credibility of the punishment. In "trigger strategies" models, after a deviation from the collusive action is observed the punishment consists in all firms reverting to the one-shot Nash equilibrium forever. Clearly, here participating in the punishment is rational: given that all other firms play the Nash equilibrium actions, a firm has no incentive to deviate from its own Nash equilibrium action. Accordingly, I have so far not explicitly introduced incentive constraints along the punishment path, because they are always satisfied.

In another class of models, due to Abreu (1986, 1988) and Abreu, Pearce and Stacchetti (1986), firms carry out optimal penal codes that involve *stick and carrot* strategies. Abreu notes that playing Nash equilibrium forever is not necessarily the optimal punishment, whenever V_i^p is positive. If a harsher punishment can be imposed, this will improve the likelihood of collusion.[73]

Abreu's proposed punishment involves a very strong punishment (the

[73] Note that playing the Bertrand equilibrium forever is the optimal penal code when firms choose prices and they have homogenous products, since it gives firms zero payoff along the punishment path, $V^p = 0$.

stick) immediately after a deviation, resulting in firms making negative profits during the punishment period. To make participation in such a strong market punishment credible, the penal code also establishes that firms will immediately revert to the collusive actions if they do take part in the punishment (the *carrot*), whereas the punishment would continue unless all firms take part in it. If the present discounted value of profit after a deviation is zero ($V_i^p = 0$), the sustainability of collusion will be maximal, since a harsher punishment cannot be built (a firm would never be willing to take part in a punishment which leaves it with a negative present discounted value of profits, because it could be better off producing zero forever). The following section illustrates a simple model with symmetric two-phase punishments, and shows how they can improve upon Nash reversal trigger strategies.

An example of optimal punishments Consider n firms producing homogenous goods and choosing *quantities* in each period for an infinite number of periods. Demand in the industry is given by $p = \max\{0, 1 - Q\}$, Q being the sum of individual outputs. All firms in the industry are identical: they have the same constant marginal costs $c < 1$, and the same discount factor δ.

Nash reversal trigger strategies Let me first give the conditions for collusion under trigger strategies with Nash reversal. Exercise **??** shows that in this industry individual outputs and profits under collusion are given by $q^m = (1-c)/(2n)$ and $\pi^m = (1-c)^2/(4n)$; the optimal outputs and profits a firm can make by deviating under collusion are given by $q^d = (n+1)(1-c)/(4n)$, and $\pi^d = (1-c)^2(n+1)^2/(16n^2)$; finally, the Cournot quantities and profits are given by: $q^{cn} = (1-c)/(n+1)$ and $\pi^{cn} = (1-c)^2/(n+1)^2$.

The ICs for collusion are given by $\pi^m/(1-\delta) \geq \pi^d + \delta\pi^{cn}/(1-\delta)$, which after substitution and rearranging becomes:

$$\delta \geq \frac{(1+n)^2}{1+6n+n^2} \equiv \delta^{cn}. \tag{12}$$

Optimal punishment strategies[74] Under the above trigger strategies with Nash reversal forever, firms make positive profits along the punishment

[74]There are two differences between my treatment here and Abreu (1986), who also analyses the Cournot model. First, he focuses only on the conditions under which the

path: $V^p = \delta\pi^{cn}/(1-\delta) > 0$. Abreu (1986) stresses that by strengthening the punishment, that is reducing the present discounted value of the profits made after the deviation, call it V^p, collusion might be achieved under milder conditions on the discount factor. He focuses on *stick and carrot* (or two-phases) strategies, and shows conditions under which there exist a symmetric optimal punishment where each firm produces the same quantity q^p and earns negative profit π^p for the period immediately after the deviation, to then revert to the collusive quantity in the following period: $V^p(q^p) = \pi^p(q^p) + \delta\pi^m/(1-\delta)$. The punishment is optimal in the sense that the quantity q^p is chosen in such a way that $V^p = 0$, which is the worst possible punishment that firms can resort to.

Of course, the punishment must be credible, that is firms should not have an incentive to deviate from the punishment path. Denoting with $\pi^{dp}(q^p)$ the profit made by deviating from the punishment path (that is, the one-shot game best response when all other firms set quantity q^p), it must be: $V^p(q^p) \geq \pi^{dp}(q^p)+\delta V^p(q^p)$, or $\pi^p(q^p)+\delta\pi^m/(1-\delta) \geq \pi^{dp}(q^p)+\delta\left(\pi^p(q^p) + \delta\pi^m/(1-\delta)\right)$. (Note that after a deviation the punishment would be restarted.)[75]

The conditions for collusion are then determined jointly by a pair of ICs (since the firms are symmetric and we focus on symmetric optimal punishments, the two ICs are identical for all the firms), which slightly rearranged are:

$$\delta \geq \frac{\pi^d - \pi^m}{\pi^m - \pi^p(q^p)} \equiv \delta^c(q^p) \quad \text{(ICcollusion)} \tag{13}$$

$$\delta \geq \frac{\pi^{dp}(q^p) - \pi^p(q^p)}{\pi^m - \pi^p(q^p)} \equiv \delta^p(q^p) \quad \text{(ICpunishment).} \tag{14}$$

First of all, note that the two ICs depend on the harshness of the punishments. The harsher the punishment (the higher the output q^p and the lower $\pi^p(q^p)$ after the deviation) the more likely that the IC for collusion is satisfied: $d\delta^c(q^p)/dq^p < 0$. However, a harsher punishment also tightens the IC along the punishment path: ceteris paribus, the punishment is sustainable

optimal punishment is sustained (but V^p is not necessarily the best firms can do, as shown below). Second, he does not restrict to the highest collusive payoff π^m as I do. As a result, he finds optimal symmetric punishments but for varying levels of collusion.

[75] Note also that the punishments are "history-independent": the punishment scheme is the same when a firm deviates from the collusive path and when it deviates from the punishment path.

only for a higher discount factor. Indeed, the higher the output q^p the lower $\pi^p(q^p)$; further, one can check that the higher q^p the lower $\pi^{dp}(q^p)$ (we are dealing with strategic substitutes: the higher the aggregate output produced by the $(n-1)$ rivals, the lower the quantity the n-firm wants to produce at its best response): $d\delta^p(q^p)/dq^p > 0$.

More specifically, in the linear demand Cournot example, we have:

$$\pi^p(q^p) = \begin{cases} (1 - nq^p - c)q^p, & \text{for } \frac{1-c}{n+1} < q^p < \frac{1}{n} \\ -cq^p, & \text{for } q^p \geq \frac{1}{n}. \end{cases} \tag{15}$$

Note first that I focus on $q^p > (1-c)/(n+1)$, which is the Cournot-Nash output; second, that the punishment profit function changes when $q \geq 1/n$, because the market price becomes zero (I assume p is non-negative).

$$\pi^{dp}(q^p) = \begin{cases} \left(1 - (n-1)\, q^p - c\right)^2 /4, & \text{for } \frac{1-c}{n+1} < q^p < \frac{1-c}{n-1} \\ 0, & \text{for } q^p \geq \frac{1-c}{n-1}. \end{cases} \tag{16}$$

Note that $\pi^{dp} = 0$ if a firm cannot make positive profits even if it reduces unilaterally its output given all the other firms set output q^p. In the present setting, this translates into the condition $p = 1 - (n-1)q^p \leq c$, or $q^p \geq (1-c)/(n-1)$. (Only in this case is the punishment optimal, in the sense that $V^p = 0$. To see this, note that the harshest punishment is credible only if $0 = V^p \geq \pi^{dp} + \delta V^p$ which implies $\pi^{dp} = 0$.)

Accordingly, one can now substitute and write the ICs as functions of q^p:

$$\delta^c(q^p) = \begin{cases} \frac{(1-c)^2(n-1)^2}{4n(1-c-2nq^p)}, & \text{for } \frac{1-c}{n+1} < q^p < \frac{1}{n} \\ \frac{(1-c)^2(n-1)^2}{4n(1-2c+c^2+4ncq^p)}, & \text{for } q^p \geq \frac{1}{n}, \end{cases} \tag{17}$$

and:

$$\delta^p(q^p) = \begin{cases} \frac{n(1-c-q-nq^p)^2}{(1-c-2nq^p)}, & \text{for } \frac{1-c}{n+1} < q^p < \frac{1-c}{n-1} \\ \frac{4nq^p(-1+c+nq^p)}{4n(1-c+2nq^p)}, & \text{for } \frac{1-c}{n-1} \leq q^p < \frac{1}{n} \\ \frac{4ncq^p}{1-2c+c^2+4ncq^p)}, & \text{for } q^p \geq \frac{1}{n}. \end{cases} \tag{18}$$

Figure 4.1 illustrates the problem: by making a stronger punishment (that is, by increasing q^p), one relaxes the IC on the collusive path, but tightens the IC on the punishment path). The point of intersection between the two ICs determines the output level $\widetilde{q}^p$ which most relaxes the conditions for both collusive and punishment path to be sustainable (that is, the lowest possible δ).

INSERT Figure 4.1 Incentive constraints along collusive and punishment paths.

It turns out that there are two possible points of intersections of these two curves. In Figure 4.1(a) the output level $\widetilde{q}^p$ that makes collusion more likely is lower than $(1-c)/(n-1)$, that is, $V^p = 0$ cannot be enforced. In Figure 4.1(b), instead, $\widetilde{q}^p$ is such that $V^p = 0$.[76]

Note that only in the second case $\widetilde{q}^p$ is such that $V^p = 0$. In other words, the best possible strategies (in the sense that they allow firms to enforce collusion for the largest possible range of discount factors) are not necessarily those which require the strongest possible punishment $V^p = 0$. The harsher the punishment, the higher the discount factor needed for making the return to collusion desirable enough to participate in the punishment itself. Imposing $V^p = 0$ might make the IC along the punishment path even tighter than the IC along the collusive path. This is precisely what happens when a small number of firms operates in the industry, as in figure 1a: setting a punishment $q^p \geq (1-c)/(n-1)$ (which implies $V^p = 0$) requires too high a discount factor for the punishment to be enforced. Hence, better to resort to a milder punishment under which $V^p > 0$.

One can then check that the lowest possible critical discount factor $\underline{\delta}$ for which the highest level of collusion can be sustained is given by:

$$\underline{\delta} = \begin{cases} \frac{(n+1)^2}{16n}, \text{ for } n < 3+2\sqrt{2} \\ \frac{(n-1)^2}{(n+1)^2}, \text{ for } n \geq 3+2\sqrt{2}. \end{cases} \qquad (19)$$

INSERT Figure 4.2 Conditions for collusion: Nash reversal (δ^{nc}) v. two-phase ($\underline{\delta}$) punishment strategies.

[76]Figure 4.1 is drawn for $c = 1/2$ and (a) $n = 4$; (b) $n = 8$. Formally, in Figure 4.1(a) $\widetilde{q}^p = (3n-1)(1-c)/[2n(n+1)] \equiv \widetilde{q}^p_1$. In Figure 4.1(b), $\widetilde{q}^p = (1+\sqrt{n})^2(1-c)/[4n\sqrt{n}] \equiv \widetilde{q}^p_2$. It turns out that $\widetilde{q}^p_1 < (1-c)/(n-1)$ for $n < 3+2\sqrt{2} \simeq 5.8$, whereas $\widetilde{q}^p_2 > (1-c)/(n-1)$ for $n > 3+2\sqrt{2}$. Recall that n is the number of firms, a discrete variable which is treated as continuous in the text just for simplicity.

Figure 4.2 illustrates the critical discount $\underline{\delta}$ obtained under these carrot and stick strategies, and compares it with the critical discount factor δ^{cn} obtained under Nash reversal strategies.

3.2 Cartels and renegotiation**

As we have said, a collusive outcome can be reached without actual meetings ever taking place. But suppose now that we consider explicit agreements (rather than tacit collusion), with firms meeting to coordinate on prices and punishment strategies. McCutcheon (1997) shows that the possibility that firms meet again and re-negotiate the agreement might actually break down the cartel, a result which has probably limited practical applications but is nevertheless very interesting and deserves attention.

Consider again our simple model of collusion with two homogenous products firms choosing *prices* in every period and reverting to the one-shot Nash equilibrium after a deviation. This model calls for both firms to have zero profit forever along the punishment path. Indeed, if they cannot meet and renegotiate their behaviour (or if for some reasons renegotiation costs were prohibitively high), neither firm would have an incentive to deviate from the punishment path: given that the rival charges at marginal cost, a firm cannot obtain higher profits by adopting a different price.

But consider now the possibility that the two firms could meet again (and assume that it is costless for them to do so, for the time being) and renegotiate their strategies. After a deviation has taken place, they would have an incentive to agree not to punish each other and restart colluding instead. Indeed, it is in their mutual interest to try and avoid a situation where they would both make zero profits forever. In other words, the enforcement of such a punishment would not be renegotiation-proof. This has a strong implication on the existence of collusion: since firms anticipate that the punishment would be renegotiated and therefore would not take place, there is nothing that will prevent them from cheating in the first place. Only if firms could commit not to meet again or they found it very costly to meet, would they be able to sustain collusion at equilibrium.[77]

This conclusion still holds if one considers a different set of strategies than simple grim strategies, for instance one where asymmetric punishments exist

[77] Note again that this argument does not apply to tacit collusion, where no meeting ever takes place.

and where the punishment lasts for a finite number of periods.[78] It makes sense to consider asymmetric treatment of players because one way to break the incentive to renegotiate is to give different payoffs to the firms along the punishment path: this way, one might want to renegotiate while the other does not. In particular, the non-deviant must get at least as much as it would obtain by accepting to renegotiate. Therefore, in this game, in order not to be induced to forgive the deviant, it should get at least $\pi(p^m)/2$.

Consider therefore a set of strategies which, for T periods after a deviation, calls for the non-deviant and the deviant firm to charge respectively prices p^p and $p > p^p$, where p^p is chosen so that the non-deviant gets at least $\pi(p^m)/2$.[79] After T periods, firms would revert to monopoly pricing.[80] Of course, T should be chosen appropriately, so as to give a low enough payoff along the punishment that a firm is not tempted to deviate. In other words, T should satisfy the incentive constraint along the collusive path:

$$\frac{\pi(p^m)}{2(1-\delta)} \geq \pi(p^m) + \frac{\delta^{T+1}\pi(p^m)}{2(1-\delta)}, \tag{20}$$

which could be rewritten $\delta(2-\delta^T) \geq 1$.

However, for collusion to hold the deviant must be induced to punish itself. The incentive constraint along the punishment path described by the strategies above would be:

$$\frac{\delta^T\pi(p^m)}{2(1-\delta)} \geq \frac{\pi(p^m)}{2} + \frac{\delta^{T+1}\pi(p^m)}{2(1-\delta)}. \tag{21}$$

[78] Only equilibria that use finite punishments are renegotiation-proof.

[79] The non-deviant firm might also be rewarded and get a profit as high as the monopoly profit along the punishment path.

[80] Note that this would not be an optimal punishment strategy, since it would leave the deviant with a strictly positive payoff along the punishment path, unlike the grim strategy. However, an optimal punishment strategy that gives a zero present discounted value of profits to the deviant along the punishment path would not eliminate the incentives of the non-deviant to renegotiate. Consider for instance a strategy which calls for the deviant firm to make negative profits for one or more periods. This requires the deviant to supply all demand below marginal cost for one or more periods. But then the non-deviant firm would not be able to make positive profits during the punishment period (either it also supplies below cost or it does not sell anything). Hence, it would have an incentive to renegotiate and restart a collusive phase.

The LHS denotes the payoff the deviant would get if it accepts to get zero payoff for T periods. The RHS denotes the payoff it would get if it deviated and charge a price slightly lower than p^p, thereby getting all demand and a profit slightly lower than $\pi(p^m)/2$ in the deviation (from the punishment path) period, but then the punishment would be restarted the next period.[81]

It is immediate, however, to check that the previous condition can be rewritten as $\delta^T \geq 1$, that is false. Therefore, an alternative set of collusive strategies which call for asymmetric payoffs during the punishment cannot be sustained at equilibrium. Hence, we are left with the conclusion that a collusive outcome cannot be sustained at equilibrium when renegotiation is allowed.[82]

Costly renegotiation McCutcheon (1997) uses this framework to argue that intermediate fines might even be counterproductive and promote collusion. Suppose that every meeting entails a probability θ of being found out, leading the firm to receive a fine F. The expected cost of a meeting, whether the first one or a renegotiation one, is therefore θF. The benefit of the initial meeting is given by the difference between the stream of collusive profits and the one-shot equilibrium payoff (equalling zero): $\pi(p^m)/(2(1-\delta))$. This meeting would therefore take place if $\theta F < \pi(p^m)/(2(1-\delta))$.

Consider now the incentive to have another meeting, after a deviation takes place. Under grim strategies that involve zero profit forever, the expected gain from renegotiation would again be $\pi(p^m)/(2(1-\delta))$. Therefore, a renegotiation meeting would take place again if $\theta F < \pi(p^m)/(2(1-\delta))$. Hence, collusion could never be sustained, since renegotiation would always take place. It makes sense, therefore, to focus on strategies that reduce the benefit from renegotiation after a deviation, that is, strategies that reduce the cost of the punishment. A strategy which calls for a punishment to take place the minimum possible number of periods T, rather than forever, fulfils this task. Such number T is the one that satisfies the incentive constraint along the collusive path (20), $\delta(2-\delta^T) \geq 1$, with equality.

[81] I am assuming that the non-deviant firm gets $\pi(p^m)/2$ along the punishment path. It could get any profit up to $\pi(p^m)$ (in which case, the deviation from the punishment path would give a profit $\pi(p^m)$ too) without changing the conclusion that collusion cannot be sustained.

[82] Note that the same result holds for any possible collusive price (just replace p^m with any $p > c$ in the discussion above to verify it). Given that firms are organised in a cartel, however, it makes sense to focus on the jointly profit maximising price p^m.

Accordingly, the cost of the punishment or, which is equivalent, the benefit from renegotiating, amounts to foregoing the collusive profits $\pi(p^m)/2$ for T periods:

$$\sum_{t=0}^{T-1} \delta^t \frac{\pi(p^m)}{2} = \frac{\pi(p^m)}{2} \left(\frac{1-\delta^T}{1-\delta} \right). \tag{22}$$

Therefore, a renegotiation meeting would take place after the punishment if $\theta F < \pi(p^m)(1-\delta^T)/\left(2(1-\delta)\right)$. Otherwise, renegotiation has no bite: firms would prefer to go through the punishment and collusion will arise.

We can then summarise the analysis as follows, according to the expected cost of a meeting, θF.

1. $\theta F \geq \pi(p^m)/\left(2(1-\delta)\right)$. In this case, each meeting is very costly because it would involve a high expected fine: firms would never start to collude in the first place: no collusion arises at equilibrium.

2. $\pi(p^m)/\left(2(1-\delta)\right) > \theta F \geq \pi(p^m)(1-\delta^T)/\left(2(1-\delta)\right)$. The fines and the probability of being found out are small enough for the initial meeting to take place, but high enough for a subsequent renegotiation meeting not to take place. This leads to the worst outcome, where collusion arises at equilibrium since the punishment will not be renegotiated.

3. $\pi(p^m)(1-\delta^T)/\left(2(1-\delta)\right) > \theta F$. Here the expected cost of a meeting is so small that renegotiation will take place, bringing the initial agreement to collapse: no collusion can be sustained at equilibrium.

Discussion McCutcheon's analysis is provocative, but rather than its policy implications (if anything, I would stress that laws should make sure that fines are large enough, since no fines might be better than small ones). Its implications are also limited by the fact that the model applies only to cartels, were there are explicit meeting, and not to tacit collusion, where no meeting takes place.[83]

I think that it underlines the importance of bargaining and negotiation among potential colluders. However, this work emphasises the negative (from

[83] Some of the assumptions she makes are questionable too: the initial meeting - which has to define all the collusive strategies - might well be more costly than the following ones.

the point of view of prospective cartel members) aspects that renegotiation of a collusive agreement might have. Yet, in different contexts, such further meetings might have a different role. For instance, after some shocks occur, such meetings might be indispensable for collusion to continue without having the industry entering costly punishment phases.

We know little about the actual working of cartels, but Genesove and Mullin (2000), one of the few rich description we have (indicate that renegotiation meetings were important to face new unforeseeable circumstances intervened in the sector, and that punishments did not occur nearly as frequently as theory would suggest, despite actual deviations did take place and perhaps due to such meetings. And this, without leading to the collapse of the collusive agreement. It seems an area where further research is needed.

3.3 The Green-Porter (1983) model**

Consider an industry with n identical firms selling a homogenous good and simultaneously playing the Bertrand game for an infinite number of times.[84] Demand realisations are stochastic and independently and identically distributed. In each period, demand is either in a low state ($D = 0$) with probability α, or in a high state ($D > 0$) with probability $1 - \alpha$, and firms do not know the state of demand at the moment they set their price, nor can they know it afterwards. Each firm does not observe the prices set by rivals either. Hence, a firm facing zero demand does not know if this is due to some rivals having undercut its price, or rather to a low demand state. (Keep the usual assumptions of the Bertrand game: if demand is positive, then all demand is shared among the firms setting the lowest price, whereas firms with higher prices receive zero demand.)

Clearly, reverting to Nash equilibrium (with zero profit) forever after observing zero demand cannot be an optimal strategy any longer, as a punishment might be triggered even if no firm has actually deviated from the collusive path. Consider the following collusive strategy instead. Each firm sets the collusive price p^m at the outset of the game and continues with it as long as all firms have positive demand. When at least one firm observes zero demand,[85] the industry enters a punishment phase which lasts for T

[84] In their paper, Green and Porter (1984) assumed firms to choose quantities. I follow here Tirole's (1988) version of the model, that is much simpler but equally insightful.

[85] Note that the event "at least one firm has zero demand" is common knowledge in this model. A firm can always infer if all firms have positive demand or not. If it receives

periods and in which each firm sets price equal to marginal cost. After the punishment phase ends, all firms revert to collusive behaviour.

As usual, to see whether these strategies represent an equilibrium, one has to derive the IC, by comparing the profit from collusive strategies and from deviating from them (given all other firms follow these strategies). To do so, it is convenient to define the following two variables, V^+ and V^-, that represent respectively the present discounted value of a firm's profit in a period which belongs to the collusive phase and in a period where a punishment phase has just started.

$$V^+ = (1-\alpha)\left(\frac{\pi(p^m)}{n} + \delta V^+\right) + \alpha\delta V^-, \tag{23}$$

where the first term is the payoff when all firms collude and demand is positive (which occurs with probability $1-\alpha$), therefore resulting the following period to be in the collusive phase again; and the second term is the payoff if a negative demand shock realises (which occurs with probability α), thereby triggering a punishment phase. Next, write V^- as:

$$V^- = \delta^T V^+; \tag{24}$$

note that firms do not get any profit during all the T periods a punishment phase lasts. The two equations above form a system in the two unknowns V^+ and V^-, whose solution is given by:

$$V^+ = \frac{(1-\alpha)\frac{\pi(p^m)}{n}}{1-(1-\alpha)\delta - \alpha\delta^{T+1}}, \quad V^- = \frac{\delta^T(1-\alpha)\frac{\pi(p^m)}{n}}{1-(1-\alpha)\delta - \alpha\delta^{T+1}}. \tag{25}$$

We now have to write the incentive constraint of a firm. If it abides to the collusive strategy, a firm gets V^+, whereas by (optimally) deviating it gets $V^d = (1-\alpha)\left(\pi(p^m) + \delta V^-\right) + \alpha\delta V^- = (1-\alpha)\pi(p^m) + \delta V^-$. Indeed, if after having slightly undercut the other firms a high demand state realises (which happens with probability $1-\alpha$), the deviating firm gets all the industry profit by itself in the current period, but it then triggers a punishment phase, which

positive demand and it did not undercut, it knows that all others have positive demand too. If it receives negative demand or if it has undercut, it knows that at least one has zero demand.

means that the present discounted value of all its profit at the beginning of the next period is V^-. If however a low demand state realises (which occurs with probability α), then the deviating firm (like all other firms) will get zero in the current period, followed in the next period by V^- as total value of all the future profit.

One can now write the IC as $V^+ \geq (1-\alpha)\pi(p^m)+\delta V^-$. After substitution, and some simple algebra, the IC becomes:

$$[\delta n(1-\alpha)-(n-1)]+[\delta^{T+1}(\alpha n-1)] \geq 0. \tag{26}$$

The first term is non-negative if $\alpha \leq (1-n+n\delta)/(n\delta) \equiv \alpha_1$; the second term is non-negative if $\alpha \geq 1/n \equiv \alpha_2$. Since $\alpha_1 < \alpha_2$, there are three possible cases. If $\alpha \leq \alpha_1$, (26) may hold; if $\alpha_1 < \alpha \leq \alpha_2$, both terms are negative and it will not hold; finally, if $\alpha > \alpha_2$, (26) does not hold because the second term (which is positive) is lower in absolute value than the first term (which is negative).

Therefore, two *necessary* conditions for the IC to hold are as follows. First, $\alpha < 1/n$: if there is a high probability to be in a low demand state, collusion cannot be sustained. Intuitively, this is because there is a high probability that demand will be zero anyhow, so the loss from deviating is low, which tightens the IC. Second, $\delta \geq (n-1)/(n(1-\alpha))$, that corresponds to $\alpha \leq \alpha_1$.[86] Note also that, even if these necessary conditions are satisfied, if $T=0$ the IC will never be satisfied,[87] whereas if $T \to \infty$ (which makes the second term equal zero) it will always be satisfied if $\delta \geq (n-1)/(n(1-\alpha))$ holds.

For further use, (26) can be rewritten as:

$$\delta^T \leq \frac{\delta n(1-\alpha)-(n-1)}{\delta(1-\alpha n)}. \tag{27}$$

The IC implicitly defines the values of the punishment length T for which collusion can be sustained. To find the optimal punishment length (that corresponds to the highest sustainable collusive outcome), note from (25)

[86]Note that for $\alpha = 0$, this condition is $\delta \geq 1-(1/n)$, the usual condition for collusion in the demand certainty case.

[87]If $T=0$, expression (26) simplifies to $\delta > 1$, which is false. The punishment phase cannot be of negligible duration.

that V^+ decreases with T, so firms would like to have the lowest possible punishment phase. Formally, the optimal T can be found by solving the problem $\max_T V^+$, subject to (27). The solution is given by the IC holding with equality. To find the optimal T explicitly, take the logarithm of both sides of expression (27), and readjust. It then becomes:

$$T^* = (\ln\delta)^{-1} \ln \frac{\delta n(1-\alpha) - (n-1)}{\delta(1-\alpha n)}. \tag{28}$$

One can check that $\partial T^*/\partial\alpha > 0$: the higher the probability to be in a low demand state the longer the punishment phase will have to be.

4 Practice: What should be legal, and what illegal?

In this section, I first deal with the standards of proof for collusion (section 4.1), then with the possible measures to deter collusion (section 4.2), and finally with those that might break ongoing collusive practices (section 4.3).

4.1 Standards of proof: market data v. hard evidence

Although industrial economics has been successful in identifying the mechanisms through which collusion acts and the factors that facilitate it, its practical implications for legal purposes are less straightforward and require some discussion. Consider for instance a literal use of the economic definition of collusion. Since a collusive outcome is defined as a situation where prices are "high enough", one could think that to verify the existence of collusion in the legal sense (that is, of anti-competitive behaviour), one has to analyse price data in a given industry, and infer if they are above some threshold levels above which they should be considered collusive.

However, it would be very difficult in practice to look at market outcomes to decide whether there has been an infringement of anti-trust law, for several reasons.[88] First of all, in many circumstances price data might not be available, and when they are they might refer to list prices rather than effective

[88] Of course, looking at market outcomes is important to the extent that it helps identify those sectors where there might be collusion, and that therefore should be subject to a more thorough investigation.

prices (in many industries, actual prices are negotiated between buyers and sellers and in others they might differ from list prices because of discounts, that might also differ across customers).

Second, even when reliable data existed, there would probably be disagreement about what the monopoly price in an industry is. Sellers might have very different views of what that price would be, and an outside observer could have yet different perceptions. It is also well known that estimates of costs might differ widely, sometimes even within the management of the same firm.

Third, suppose that there is agreement on what the monopoly price in the industry would be: how close to the theoretical monopoly price sales prices should be for them to be judged "too high" and therefore collusive?

Fourth, the very principle that firms could be convicted solely on the grounds that they charge "too high" prices is a dangerous one, and it might open the way for anti-trust interventions whenever firms are successful enough to find consumers willing to pay high prices for their products. As I have repeatedly argued in chapter 2, it is not market power by itself that anti-trust intervention should be after.

Rather than looking at the *level* of prices in the industry, one might then be tempted to infer the existence of collusion (as an infringement of the law) by analysing the *evolution* of industry prices over time. For instance, courts and anti-trust authorities have sometimes been tempted to infer the existence of collusive (illegal) behaviour from the fact that sellers charge similar prices over time, the so-called "price parallelism" (or "conscious parallelism").[89] But to observe that prices move in a similar way is not enough to establish that firms are guilty of collusion. Common exogenous shocks such as the increase in input prices of all the suppliers, or an increase in inflation, or an increase in property prices would probably lead all the sellers to increase prices proportionally, without implying that they are colluding.

Further, we have already noticed that a collusive outcome might arise without firms agreeing or communicating to coordinate their behaviour. Suppose for instance that - even without common shocks on demand or input prices - one day a seller increases prices by 10%, and that the next day a rival follows suit. Is this price parallelism enough evidence to convict firms? Surely not. It is of course possible that the two firms have talked to each

[89] See Scherer and Ross (1990: 339-346) for a discussion of the conscious parallelism doctrine and its evolution in the US.

other and previously agreed about changing prices. But it is also possible that they have taken their decisions without communication. The first firm might have simply raised prices in the expectation that the rival would follow its price rise, and the rival might have simply decided to follow, either happy to raise prices or in the expectation that a failure to do so would trigger a price war that would reduce profits. In the absence of hard evidence, a court would have to prove infringement of the law by second-guessing the firms' intentions and their motivations.

An example where history provides a focal point for tacit collusion, rendering explicit collusion superfluous, is given by the *Soda-Ash* case. Soda-ash is a commodity used as a raw material in the production of glass. ICI, a British company, and Solvay, a Belgian company, are the main producers in the industry. The two firms had a long history of explicit market-sharing agreements (at times when cartels were not illegal), started in the 1870s and renewed immediately after the war with a so-called 'Page 1000' agreement, which divided Europe (and some overseas markets) in spheres of influence: for instance, ICI was to sell in the United Kingdom and Solvay in Continental Europe.

The agreement (that the defendants indicated as out of date since 1962) was terminated from 31 December 1972, when the UK entered the European Community (to comply with the anti-trust rules of the Treaty), but:

> "The alleged desuetude of the 'Page 1000' arrangement did not however manifest itself in any significant change in the commercial policy of Solvay or ICI in the soda-ash sector, either in 1962 or at any later stage. Neither ever competed with the other in their respective home markets in the Community. Similarly in overseas export markets each continued to respect the other's sphere of influence." (*Soda-Ash*: 27)

What is noticeable is that each firm admitted that it had no intention of invading the other's home market, but simply because it feared retaliation if it had done so (*Soda-Ash*: 42-44). They therefore justified a collusive outcome as the result of independent decisions that made sense from a business viewpoint. In this case, continuing to share markets was an easy way to reach tacit collusion.

The other interesting point here is whether tacit collusion is an infringement of article 81 (ex-85). In this case, the Commission decided it was, and that the term 'concerted practice' mentioned in article 81 among the prohibited practices covered also tacit collusion:

> "The Commission fully accepts that there is no direct evidence of an express agreement between Solvay and ICI to continue to respect the 'Page 1000' cartel in practice. However, there is no need for an express agreement in order for article 85 to apply. A tacit agreement would also fall under Community competition law." (*Soda-Ash*: 55)

Note, however, that in some cases price parallelism can only be credibly explained by coordination among firms, even if there is no proof of the latter. For instance, in the *Dyestuffs* case, price rises were so simultaneous that it is impossible that they had not been previously agreed:

> "In Italy, apart from Ciba who had already ordered its Italian subsidiary to increase prices, all other producers, with the exception of ACNA, sent by telex or fax - from their headquarters, seated in places very distant from each other - instructions to their respective agents in the afternoon of 9 January: Sandoz at 17.05, Hoechst at 17.09, Bayer at 17.38, Francolor at 17.57, BASF at 18.55, Geigy at 19.45, and ICI at an undetermined time, since instructions were given by phone." (Dyestuffs: 2. My translation.)

Likewise, a simple rule such as the "parallelism plus" rule, which consists of finding illegal behaviour whenever price parallelism is accompanied by the presence of a facilitating factor (such as for instance resale price maintenance, best price clauses, or exchange of information) is not more convincing either, unless it can be proved that firms have coordinated in order to introduce or keep the facilitating practice at stake. In other words, if there is no proof that they have agreed on a particular practice, the very fact that they have followed it should not be proof of collusion. For instance, in the *Woodpulp* case, the EC saw a facilitating practice in the fact that producers of pulp for paper production quoted all prices in dollars and announced their price changes with the same advance. Together with parallel price movements, this was seen as evidence of collusion. However, the ECJ found that the woodpulp sellers introduced such practices at the request of their customers, the paper producers, who wanted to have more price transparency as well as to be informed in time about input price changes, so as to take the appropriate steps to face them. Accordingly, the ECJ quashed the EC decision.[90]

[90] Similarly, in the *Ethyl* case in the US, a Court of Appeal ruled that the pricing practices adopted by Ethyl had been introduced when it was still a monopolist, and that other firms had later independently adopted them. See Scherer and Ross (1990: 345).

The presence of periods of price wars is no proof of collusion either. We have seen that under imperfect price observability (and demand uncertainty), full collusion cannot be sustained and price wars are an integral part of collusion. However, observing that there are episodes in an industry where prices fall considerably might be consistent with either Green and Porter's type of collusion or with some other events, such as new capacity appearing in the market, occasional competitors temporarily in the market (e.g., imports), a reduction in demand and so on. Certainly, repeated episodes of this type would raise suspicion and deserve a careful scrutiny of the industry, but they should not be seen as the ultimate proof that collusion exists. Again, it is possible that firms trigger price wars so as to sustain a Green and Porter's type of collusion, but in the absence of communication among them, I do not see why they should be convicted. Firms would probably argue that lower prices in some particular periods were due to particular capacity or demand conditions, or that they simply behaved so as to match prices set by one of them, or again that price fell but not to the competitive (or sub-competitive) levels predicted by theory, and it would be very difficult for a court to rule out such arguments, unless there exists evidence of communication among firms to coordinate their behaviour.

The use of econometric techniques, however sophisticated, does not change the nature of the arguments above. It is instructive to look at the estimates carried out on the railroad cartel in the US, that pre-dated the Sherman Act (and therefore for which public data exist). This cartel was studied first by Porter (1983b), whose econometric analysis found that the data were not inconsistent with the firms in the industry behaving à la Cournot (and therefore not colluding). However, Ellison (1994) studied the same data but used a slightly different econometric specification, and he found that firms' behaviour was close to full collusion.[91]

The lesson from these papers is that even if one thought that collusion might be proved on the basis of market outcomes alone, and if good and reliable data existed, econometric techniques might not always provide unambiguous answers as to the existence of collusive prices in a given industry. Perhaps in the future there will be more consensus on how to design and assess econometric exercises of this type, but for the time being econometrics is more likely to give complementary evidence, rather than conclusive proof,

[91] In the same vein, and using the same data on the railroad cartel, Porter (1985) and Vasconcelos (2001a) find a different impact of entry upon collusion.

of collusion.

For all these reasons, inferring *illegal* collusive behaviour - i.e., inferring conspiracy in the US or infringement of article 81 in the EU - from market data (that is, only looking at the outcomes) would not be desirable, and the legal approach which requests some hard evidence as proof of collusion is sensible practice. Firms should be convicted for anti-competitive behaviour only insofar as there is proof that they have communicated with each other to sustain collusion. Such communication can be of very different sort. First of all, and more obvious, minutes of meetings, e-mail messages, memos and other written (or recorded) evidence concerning agreements on prices and quantities are still the most likely proof of collusion. But firms might also sustain collusion without discussing openly prices or quantities, but coordinating so as to establish the environment that facilitates collusion. For instance, they might decide to exchange detailed price and quantity information via their trade association, or they might set up a forum where they can announce future prices to each other (as in the *Airline Tariff Publishing* case), or agree on a resale price maintenance scheme or other practices that make more uniform or transparent their prices.[92] In all such cases, if there is evidence that firms have not acted unilaterally, firms should be found guilty of collusion.[93]

This approach has the advantage that it is based on observable elements verifiable in court: if there is any evidence of communication or coordination (on prices, quantities, or on facilitating practices in the sense described above) among the firms, then courts and competition authorities should fine them for anti-competitive (collusive) behaviour. Otherwise, they should not.

Too lenient with the firms? It might seem that this approach is particularly generous with firms. After all, I have explained above that they might be able to sustain high prices even without communicating to each other. Therefore, why should they take actions that could leave hard evidence? And if they do not need to communicate, is there any hope to punish (and deter) supra-competitive pricing?

Two considerations are in order here. First, it is true that tacit collusion

[92] See the *Sugar Institute* case, described in Genesove and Mullin (2001).

[93] If they have acted unilaterally, or on the basis of habit, but the investigation determines that the practices in place favour collusion but do not entail any efficiency gain, the court should not fine them but impose a cessation of such practices.

might be sustained by firms. However, we have also seen that there are very good reasons why firms would like to communicate and/or to coordinate their actions. They might want to avoid unnecessary and costly experimenting with the market and choose instead the best (for the firms) prices, or they might want to create facilitating practices and more generally an environment which improves observability of firms' actions to favour collusion. This will lead firms to try and communicate among them so as to coordinate their actions, thereby leaving traces of hard evidence behind them. Firms have known for a long time that they will be found guilty if there is any written proof of their coordination, and yet anti-trust authorities keep on uncovering such hard evidence in cartel cases.[94]

Second, there is no alternative to such an approach. Any other rule - such as for instance inference from market data - which is not based on observables could not be easily enforced in courts. It would also be detrimental to legal certainty, as firms would not know whether their pricing policies might be accepted or fined. At the same time, competition authorities would have to decide on a case-by-case basis, rather than follow a clear rule.

Still, it might appear that this approach does not do enough to deter and punish collusion. What else can be done, then? There are two types of competition policy interventions which might help deter tacit or explicit collusion or break explicit collusion (i.e., cartels). I divide them into ex-ante and ex-post measures.

4.2 Ex-ante competition policies against collusion

Collusive agreements are possibly the most serious infringement of competition law in any jurisdiction, and they are accordingly heavily fined. Firms might be subject to different penalties if they are caught guilty of collusion. First, they will generally pay a fine (which usually is a transfer to the country's general budget). Second, they might have to pay damages to private parties (in the US, treble damages). Third, the firms' managers might be given prison sentences. Recently, in the UK and in the EU there have been discussions about the opportunity of making collusion a criminal offence. If this was the case, executives caught guilty of collusive agreements could be

[94]Noteworty are a stream of high profile international cartels prosecuted by both US and EU authorities in the late 1990s, among which *Citric Acid*, *Lysine*, *Vitamins*, and *Graphite Electrodes*.

imprisoned, like in the US. Arguably, this provides a stronger deterrent of collusion as risk averse managers would find it very dangerous to collude.

However, what matters for deterrence purpose is not the size of the fine if caught guilty of collusion, but the expected fine, namely the amount of the fine multiplied by the probability of being caught and found guilty. In turn, this raises the issue, so far little studied, of how anti-trust authorities should design their policies and organise their investigations so as to efficiently deter collusion.[95]

Apart from making firms face tougher penalties, I argue in what follows that other ex-ante measures might have an important role in deterring collusion.

Black list of facilitating practices Since collusion is facilitated by certain practices, competition authorities should identify business practices that should be forbidden and others which might be tolerated apart from specific cases. Some practices should therefore be on a black list and be per se prohibited, and others should be under a rule of reason. In terms of European competition law, for instance, the EC might issue guidelines indicating black list practices that would represent an infringement of article 81(1) - but firms might ask for an individual exemption under article 81(3) provided that they prove that they achieve efficiency gains.

Such practices should include announcements about future price and quantity conduct; exchange of disaggregate information about individual prices and outputs; any *coordination* among firms aimed at harmonising business practices that increase price observability among sellers (without increasing transparency for buyers), such as resale price maintenance, basing point pricing, and best price clauses. Minority shareholdings among competitors also appear as a pro-collusive practice that should be authorised only if efficiency effects can be showed.

Auction design to avoid bid-rigging In section 2.2.1 above I have discussed the possibility of collusion in auctions, and more particularly in

[95]For instance, it might make sense to tolerate prices that are perceived to be above some competitive benchmark by not much, but intervene as soon as prices are above a certain threshold, as argued in Besanko and Spulber (1989). See also section 4.3 below on how anti-trust authoritiess might optimally use their limited resources. The issue of how colluding firms change their conduct in the presence of an anti-trust authority is also analysed in LaCasse (1995).

simultaneous ascending auctions. One problem there was that players might use their bids to signal their collusive intentions. Since it is difficult to intervene ex-post to contest the legality of certain bids, Klemperer (2001) suggests to intervene ex-ante by choosing an auction design that can minimise such problems. This suggestion is in the same spirit of what I propose here: it is better to try and create an environment that discourages collusion in the first place, than trying to prove unlawful behaviour afterwards. A clear advantage of auction markets is that the environment can be affected directly, since the rules of the game are specified at the beginning by the auctioneer.

In simultaneous ascending auctions signaling can be avoided by authorising round number and/or anonymous bidding, and objects can be aggregated in larger lots so as to make it harder for players to divide them, and/or a final round with sealed bid offers among the two remaining players can be introduced.[96]

More generally, collusion in auctions can be made more difficult by appropriate auction design. Consider for instance the school milk cartels in Florida and Texas studied by Pesendorfer (2000, especially page 389). The market was characterised by many small contracts (on average, 239 per year in Florida, and 136 in the Dallas-Forth Worth area), that allowed the cartels to better divide the spoils. The Boards of Education in every school district awarded contracts independently of each other, and at different dates. They also publicly announced bids and identities of all bidders. This made it possible for the bidding firms not only to spot immediately deviations from a collusive agreement but also to punish a deviation in the following auctions. A different procurement strategy might have helped. For instance, the school districts might have coordinated and acted as a single buyer (recall that large orders break collusion), or they could have fixed the auctions in the same day, or again they could have not revealed information about bids and bidders.

Merger analysis Another ex-ante instrument to prevent collusion from arising is given by merger control. Indeed, we know that a reduction in the

[96] This is the so-called Anglo-Dutch auction. See Klemperer (2001) for details and discussion. The UK generation market is also a case where auction design matters. Before the 2001 reform, there was a (sealed-bid) *uniform price* auction, where generators submit a bid that states the different quantities they are ready to supply and at which price. All generators will then be paid the highest accepted bid. Fabra (2001a) formally shows that this type of auction has a higher collusive potential than a *pay-as-bid* auction, where the generator will be paid at its own bid. See her paper for a discussion.

number of firms in the industry, or a more symmetric distribution of their assets would favour collusion. Therefore, by increasing concentration, and to the extent that they increase symmetry, mergers might create favourable conditions for collusion to be sustained in a given industry. It is crucial, therefore, that competition authorities should be vigilant on mergers (see chapter 5).

4.3 Ex-post competition policies against collusion

Surprise inspections ("Dawn raids") Next to ex-ante measures aimed at preventing collusion, competition authorities should also intervene to try and break existing cartels. From the discussion above, which attributes great importance to the discovery of hard evidence as a proof of collusion, it is obvious that one such a measure is given by dawn raids. Police search in the headquarters of firms suspected of collusive conduct (or trade associations, or even executives' homes) often results in crucial evidence being uncovered. Accordingly, competition authorities should be given extensive search powers so that they can, in collaboration with police forces, seize documents which might help prove collusive agreements. But competition authorities might also resort to more clever ways to break collusion, and provide incentives for firms to withdraw from collusive agreements and reveal hard information needed to prove them in courts. It is to such measures that I turn next.

Leniency programmes In recent years, competition authorities have devoted a lot of attention to more sophisticated fine schemes. Such schemes are called "leniency programmes" and grant total or partial immunity from fines to firms that collaborate with the authorities. They work on the principle that people who break the law might report their crimes or illegal activities if given proper incentives.[97] In competition law, the Antitrust Division of the Department of Justice (DOJ) in the US have been the first to introduce such a law, in 1978, granting immunity from criminal sanctions if

[97] Similar schemes are routinely used in several fields other than antitrust, such as fiscal law and environmental law. In Italy, the so-called "turncoat laws" ("*leggi sui pentiti*") have been successfully used to fight organised crime such as the mafia and terrorist organisation such as the Red Brigades. Of course, there are ethical issues involved because punishment is abandoned in exchange of deterrence of further crimes: criminals might be set free (and sometimes even rewarded) in exchange for information that allows to imprison other criminals.

certain conditions occurred.[98] In August 1993, this scheme - that had not been particularly successful - was thoroughly redesigned by the DOJ, and it now works as follows.

There is *automatic leniency* for firms that report evidence of a cartel *before* an investigation has begun, provided that: the firm is the first to come in; it terminates participation in the illegal activity; it fully and continuously collaborates with the DOJ; it makes restitution to injured parties; it did not coerce another party in the activity, nor was it its leader or originator. *Discretionary leniency* exists for firms that report evidence *after* an investigation has started, provided that the DOJ does not yet have evidence against the company that is likely to result in a sustainable conviction (plus similar accessory conditions as above).

The 1993 reform has improved the leniency programme in two major ways. First, it has extended the possibility of leniency to firms that cooperate *after* an investigation is already under way (that is precisely what theory would suggest, see below). Second, it is now *clear and certain* in that - at least for the case where cooperation unveils a cartel before an investigation started - amnesty is automatic rather than discretionary.[99] The changes have been extremely fruitful: while under the old policy on average only one corporation per year applied for amnesty, under the revised policy applications for amnesty have come in at the rate of approximately two per month.[100]

The EU introduced a leniency policy in 1996.[101] It established that a fine might have been very substantially (75-100%) reduced if a company informed the European Commission before an investigation started; and substantially (50-75%) reduced if cooperation started after an investigation had started but before the EC had obtained sufficient grounds for initiating the procedure; in both cases, the company had to be the first to report, had terminated all cartel activities and it had not been the instigator of the cartel; it might have been significantly (10-50%) reduced if the company cooperated with the EC in the investigations (for instance by not challenging the EC findings and allegations) without the previous conditions for more generous reduction of

[98] Recall that in the US, unlike Europe, managers risk prison sentences up to three years and criminal fines if their firms are caught guilty of price-fixing.

[99] An additional novelty is that all officers, directors, employees who cooperate are protected from criminal prosecution.

[100] Spratling (1998: 2).

[101] See "Commission Notice on the non-imposition or reduction of fines in cartel cases", published in OJ C207 of 18 July 1996.

fines being met.

However, this policy did not give the results the EC hoped for, mainly for two reasons.[102] First, leniency was given in a discretionary way by the EC (rather than being automatic like in the US), and firms did not know what fines they would get until the final decision is adopted. This clearly reduced the benefit from disclosing evidence. Second, firms did not receive immunity if an investigation had already begun.

In February 2002, the EC has adopted a new leniency policy.[103] It improves on the first point since it introduces transparency and certainty: complete immunity from fines is given to the firm first reporting a cartel and, after providing evidence, the firm will receive (conditional) immunity in writing from the EC. Further, the new rules specify that any firm can apply for immunity as long as it had not coerced other firms to participate in the cartel (the previous condition, requiring a firm not to be an "instigator" of the cartel, left room for interpretation).

It also improves on the second point, since immunity is given to a firm that provides evidence that enables the EC to establish an infringement even when the EC is already in possession of enough information to launch an inspection (but not to establish an infringement).[104]

The use of leniency programmes in anti-trust has been studied first by Motta and Polo (1999, 2001).[105] They show that such programmes might have an important role in the prosecution of cartels provided that firms can apply for leniency *after* an investigation has started. When a firm decides whether to join a cartel, it takes into account the risk of being caught. In other words, its decision of colluding or not weighs the benefit of collusive profits against the expected cost of it, namely the probability μ of being caught times the fine F it expects to pay in such a case. If, after having decided to take part in a cartel, a firm receives the opportunity to give itself in (that is if a leniency programme becomes available), but the expected

[102] Although a reduction of fines was given in several cases, full immunity was granted only three times under the old leniency policy: to Rhône-Poulenc (in two vitamins cartels), Interbrew (Luxembourg brewers cartel), and Sappi (carbonless paper cartel). All three decisions were taken in November and December 2001, when the change in policy was already under way. See European Commission Press Release IP/02/247 of 13/02/2002.

[103] "Commission Notice on immunity from fines and reduction of fines in cartel cases", OJ C45/3, 19.2.2002.

[104] A reduction of fines is granted to firms that do not fulfill the previous conditions, but provide evidence that has *significant value added* for the investigation.

[105] See also Spagnolo (2000) and Rey (2000).

cost μF of it does not change, then there is no reason why the firm should report the cartel. If the benefit of collusion was already higher than its expected cost, and nothing changes in the information available to the firm, then the firm will continue to collude even if a lower fine is available in case of reporting.[106]

But consider now a leniency programme that is open to firms even after an investigation has already begun. Decompose the probability μ of being caught as the probability α that the cartel is investigated times the probability p that the competition authority gathers enough hard evidence for the investigated cartel to be proved guilty. (Therefore, $\mu = \alpha p$.) In this case, a decision to start a cartel is taken when the expected cost of being caught is αpF, but *after* an investigation of the industry has started (if any), the expected cost is pF, which is higher than αpF (since $\alpha < 1$): the trade-off is changed, as the expected cost of continuing to collude has become higher, whereas the collusive profit is the same. If given the possibility to apply for leniency, the firm might well decide to give up its participation in the cartel.

Leniency also helps in that it saves resources of the authority: building up a convincing enough case to be defendable in courts is very costly, but the cost of this prosecution stage can be avoided or greatly reduced by leniency, since the firms would bring themselves enough evidence to the authority.

Leniency programmes** This section presents a simple model, due to Motta and Polo (1999), to study the effects of reduced fines for firms cooperating with the Anti-trust Authority (AA from now on) can be useful.

The timing of the (infinite horizon) game is as follows:

At time $t = 0$, the AA can commit to a Leniency Programme (LP) which allows for reduced fines $0 \leq R \leq F$ to firms which reveal information useful to prove collusion,[107] F being fixed by law as the maximum fine that firms can receive if found guilty of collusion.[108] This policy is observed by all firms, that also know the probability α that the AA will open an investigation and

[106] Only if new managers with a different perception of the risk of being caught have meanwhile taken over the firm, would the leniency programme have some effect.

[107] I assume that information given by a single firm is enough to prove that all the firms which have taken part in the collusion are guilty. This might be the case when each firm has access to the minutes of the meetings, or has other written evidence of communication that the firms have used to coordinate on the collusive outcome.

[108] It is always optimal for the AA to set the fine for non-cooperating firms at the maximum level.

the probability p that it will be able to prove colluding firms guilty.[109] Note that the reduced fine R will be granted to any firm cooperating even *after* the investigation is opened.[110]

At time $t = 1$, n identical firms decide whether to collude or deviate and realize the per-period associated payoff, respectively Π_M and Π_D (with $\Pi_N < \Pi_M < \Pi_D$). Consider grim strategies: a deviation triggers the punishment of the other firms, which will play the one-shot non-cooperative equilibrium action forever afterwards, giving a payoff Π_N to each firm.

Assume that the existence of a collusive outcome in the industry is perfectly observed by the AA, but this is not enough for collusion to be proved in courts. To build a case against the firms, the AA needs to find some "hard" proof of coordination.[111] Perfect observability of collusive prices also implies that the AA will never open an investigation on firms which do not collude at equilibrium.

At time $t = 2$, the AA opens an investigation with probability $\alpha \in [0, 1]$. If the inquiry is not opened, each firm realizes the profit Π_M. If it is opened, firms simultaneously decide whether to reveal information to the AA; if at least one firm reveals, the AA is able to prove them guilty. A firm which cooperates with the AA pays $R \leq F$ whereas one that does not pays the full fine F. If no firm reveals, the AA is able to prove them guilty with probability $p \in [0, 1]$. If the AA has not been able to prove the firms guilty of collusion at the end of this inquiry, the firms will never be investigated again in the future. If proved guilty, they will behave non-cooperatively forever in the future. After the investigation has ended, payoffs are realised. Figure **??**

[109] The former refers to the preliminary activities (general monitoring) necessary to open an investigation such as collecting information about the firms in the industry, interviewing firms, suppliers and customers, collecting data from the different sources; the latter (prosecution) involves collecting more focused information on the case, ordering surprise "raids" in the firms' headquarters, processing the information collected and preparing the case against the firms according to the existing laws. Here α and p are exogenously given for simplicity. See Motta and Polo (1999) where they are determined by the AA.

[110] Spagnolo (2000) shows that it might be optimal not only to give fine discounts to collaborating firms, but also to reward them ($R < 0$). However, giving extra benefit to firms reporting evidence of a cartel is unlikely to be used in practice because of moral issues. Further, it might give firms incentives to fabricate evidence in order to get an extra benefit. The authorities might then have to invest resources to check the claims of the reporting firms, thereby losing much of the appeal of the leniency programme itself.

[111] Since there exists a continuum of possible equilibria, firms need some coordination to select the fully collusive outcome giving them the per-period profit Π_M.

illustrates the game at period $t = 2$.

For any $t > 2$, if up to the previous period the AA has not started an investigation, with probability α it opens an inquiry in t, firms decide whether to reveal, and so on.

INSERT Figure 4.3 Game tree, at $t = 2$.

We restrict attention to the case where $\delta \geq (\Pi_D - \Pi_M)/(\Pi_D - \Pi_N)$: in the absence of anti-trust policy, collusion would arise at equilibrium.

Solution To look for the sub-game perfect equilibria of the game, consider first the "revelation game" which is played once an investigation is opened by the AA. If a firm reveals, it gets a payoff of $\frac{\Pi_N}{1-\delta} - R$ independently of the action chosen by the other firms. If a firm does not reveal any information but at least one other firm does, then the former firm receives a payoff of $\frac{\Pi_N}{1-\delta} - F$. Finally, if no firm reveals any information, each firm receives an expected payoff

$$p(\frac{\Pi_N}{1-\delta} - F) + (1-p)\frac{\Pi_M}{1-\delta}. \tag{29}$$

INSERT Table 4.1 Payoff matrix in the "revelation game"

It is easy to check that the n-tuple (*reveal,.., reveal*), that we denote as (R,..,R), in which all firms choose to cooperate with the AA obtaining a reduction in fines, is always a Nash equilibrium. Instead, the n-tuple (*no reveal,.., no reveal*), or (NR,.., NR), is an equilibrium: (1) if $pF < R$, always; (2) if $pF \geq R$ if the following condition holds:

$$p \leq \frac{\Pi_M - \Pi_N + R(1-\delta)}{\Pi_M - \Pi_N + F(1-\delta)} = \tilde{p}(\delta, R, F). \tag{30}$$

Note that when the (NR,.., NR) equilibrium exists, it would also be the one selected by standard criteria of equilibrium selection such as Pareto-dominance or risk dominance.[112] Therefore, firms reveal information only

[112] It is also reasonable to believe that if firms are able to coordinate on the collusive outcome, they will also be able to coordinate not to reveal when this is an equilibrium which gives them a Pareto superior payoff.

if $p > \tilde{p}$. Note that (a) when no leniency programme exists $R = F$ and $\tilde{p} = 1$: firms will never collaborate with the AA even after an investigation has started; (b) to induce revelation the best the AA can do is to set $R = 0$.

As for the decisions taken by the firms at $t = 1$, we have to find the discounted sum of profits if a firm decides to collude and compare it with the discounted sum of profits if it decides to deviate from the collusive strategy. This comparison must be done for both cases, when firms will decide to reveal once investigated, and when they will prefer not to cooperate with the AA.

(1) *Collude and reveal*: $p > \tilde{p}$. In this case firms reveal if an investigation is opened by the AA. Define Π_R as the expected profit immediately before an investigation is opened. It is easy to see that:

$$\Pi_R = \alpha(\frac{\Pi_N}{1-\delta} - R) + (1-\alpha)(\Pi_M + \delta\Pi_R) \tag{31}$$

which can be rewritten:

$$\Pi_R = \frac{(1-\alpha)\Pi_M + \alpha(\frac{\Pi_N}{1-\delta} - R)}{1-\delta(1-\alpha)}. \tag{32}$$

If a firm decides to set the collusive price, then its expected discounted payoff will be:

$$V_{CR} = \Pi_M + \delta\Pi_R = \frac{\Pi_M + \delta\alpha(\frac{\Pi_N}{1-\delta} - R)}{1-\delta(1-\alpha)}. \tag{33}$$

If instead a firm decides to deviate from the collusive strategy, then its payoff is given by:

$$V_D = \Pi_D + \frac{\delta\Pi_N}{1-\delta}. \tag{34}$$

Collusion can arise if $V_{CR} \geq V_D$, that is if the following condition is satisfied:

$$\alpha \leq \frac{\Pi_M - \Pi_D + \delta(\Pi_D - \Pi_N)}{\delta(\Pi_D - \Pi_N + R)} = \alpha_{CR}(\delta, R). \tag{35}$$

Figure 4.4 illustrates α_{CR} in the plane (p, α), for given values of δ and R: this locus does not depend on p since in the region considered here firms cooperate with the AA once an investigation is opened.

Below the line, firms prefer to collude even though they anticipate that, if an investigation is opened, collusion would collapse because firms would reveal information to the AA. Above the line, firms, anticipating revelation, prefer to deviate, and the collusive outcome never occurs.

Note also that the less generous the leniency programme (the higher the reduced fine R) the lower α_{CR}: if firms expect that in case an investigation is opened they can reveal and get away with a small fine, they will have an extra incentive to choose the collusive strategy. In other words, a generous leniency policy might stimulate ex-ante collusion. (I shall come back to this issue below.)

INSERT Figure 4.4. Equilibrium solutions for given policy parameters.

(2) *Collude and not reveal*: $p \leq \tilde{p}$. In this case, firms anticipate that even if an investigation is started, no firm will reveal any information. Only by beginning an investigation and then proving the firms guilty, can the AA break the cartel.

Write the expected profit immediately before knowing if an investigation is opened as:

$$\Pi_{NR} = \alpha[p(\frac{\Pi_N}{1-\delta} - F) + (1-p)(\frac{\Pi_M}{1-\delta})] + (1-\alpha)(\Pi_M + \delta\Pi_{NR}), \quad (36)$$

whence:

$$\Pi_{NR} = \frac{\alpha[p(\frac{\Pi_N}{1-\delta} - F) + (1-p)(\frac{\Pi_M}{1-\delta})] + (1-\alpha)\Pi_M}{1-\delta(1-\alpha)}. \quad (37)$$

If a firm follows the collusive strategy its expected discounted payoff is given by:

$$V_{CNR} = \Pi_M + \delta\Pi_{NR} = \frac{\Pi_M(1 + \frac{\delta\alpha(1-p)}{1-\delta}) + \delta\alpha p(\frac{\Pi_N}{1-\delta} - F)}{1-\delta(1-\alpha)}. \quad (38)$$

As before, a firm which deviates obtains a payoff $V_D = \Pi_D + \frac{\delta\Pi_N}{1-\delta}$. Some simple but boring algebra shows that the inequality $V_{CNR} \geq V_D$ is satisfied for:

$$\alpha \leq \frac{(1-\delta)[\Pi_M - \Pi_D + \delta(\Pi_D - \Pi_N)]}{\delta[pF(1-\delta) + p(\Pi_M - \Pi_N) + \Pi_D(1-\delta) - \Pi_M + \delta\Pi_N]} = \alpha_{CNR}(\delta, p, F), \quad (39)$$

when $p\left[F(1-\delta)+\Pi_M-\Pi_N\right] > \Pi_M-\Pi_D+\delta(\Pi_D-\Pi_N)$, and always satisfied otherwise. Figure 4.4 illustrates the curve, which decreases with p: other things being equal, an increase in the probability of being found guilty will make collusion less likely (and at low enough p, collusion is always preferred). Of course, when F rises, given p, collusion will be less profitable.

Figure 4.4 illustrates the equilibrium solutions of the game. Note that if no LP is introduced ($R = F$) firms have no reason to reveal information to the authority once an investigation is opened, and the equilibrium outcomes would be defined uniquely by the line α_{CNR}. Below the line, firms would collude (CNR); above it, they would not (NC), because any proposed agreement would break down immediately.

Reduced fines ($R < F$) modify the situation: when $p < \tilde{p}$ firms do not reveal if monitored, and the analysis above still applies: the line α_{CNR} distinguishes the equilibrium where firms collude and not reveal from that where no collusion occurs. When $p \geq \tilde{p}$ firms anticipate that they reveal information if monitored: above α_{CR} they prefer not to collude and below α_{CR} they initially collude and then reveal if monitored.

To understand the role of leniency on the sustainability of collusion, consider what happens when, starting with a situation in which no LP is used, we introduce reduced fines. This has two effects which are shown in Figure **??**. On the one hand, the LP might have an adverse, pro-collusive effect. By reducing the expected value of the fine to be paid if an investigation is opened, the LP might give an incentive to collusion. This occurs in the area (a) included between the thin section of the curve α_{CNR} and the line α_{CR}. In this region, no collusion can be sustained in the industry if full fines are given, but under a LP firms would engage in collusion and, if monitored, they would reveal and pay the reduced fine $R < F$.

On the other hand, there exists an area (b) where collusion will break down (because the firms reveal information) if the AA starts monitoring the industry, whereas in the absence of a LP collusion could stop only after a successful complete investigation. This is the area below the dotted part of the curve α_{CNR}.[113]

---

[113] If the Leniency Programme was unanticipated, firms would decide whether to collude or not on the basis of an expected fine $R = F$ and therefore would not cooperate unless $\alpha < \alpha_{CNR}$. When unexpectedly the leniency program is introduced, collusion would break down in all the area (2), without any adverse effect arising.

Implementing the optimal policy A leniency programme is therefore not unambiguously optimal, as it might introduce a trade-off between lower ex-ante deterrence and ex-post desistence of cartels. Motta and Polo (1999, 2001) formally analyse the optimal anti-trust policy, and show that leniency should be used when the AA has limited resources. A full analysis requires maximising the objective function of the AA, that is welfare, subject to its budget constraint. However, even without a full analysis, note that the AA will rank the regions as follows: NC>CR>CNR. Cartels entail an allocative efficiency loss, and so the AA wishes to deter or break them. In the first case (NC), cartels are deterred; in the second (CR), cartels are broken as soon as an investigation is started, because firms reveal information to the AA; in the third, they are broken only if an investigation is started and prosecution is successful.

Intuitively, if the AA had a high budget, it could set high values of (p, α) and would be able to achieve full deterrence by using the full fine F, while introducing a leniency programme might result in lower deterrence (i.e., ending up in region (a)). Instead, if the AA had lower budget, full deterrence can never be achieved and it becomes better to implement the outcome where firms collude and reveal (CR) by granting maximum discounts ($R = 0$) rather than having full collusion (CNR).

Fine reductions only before the inquiry is opened An alternative leniency policy where firms can apply for fine discounts only *before* an inquiry is opened would correspond to the same game as above, with the difference that at the beginning of time $t = 2$, firms simultaneously choose whether to reveal the cartel to the AA, or not; if no firm reveals, the AA opens an investigation with probability α, proving them guilty with probability p, and after the investigation is resolved, payoffs are realized.

Under this alternative regime the Leniency Programme is completely ineffective. An equilibrium in which firms choose to collude and reveal does not exist. By colluding when expecting the cartel to be broken by information given to the AA, a firm would get $V_c = \Pi_M + \delta(\Pi_N/(1-\delta) - R)$. By deviating, it would get $V_d = \Pi_D + \delta\Pi_N/(1-\delta)$. Since $\Pi_D > \Pi_M$ and $R \geq 0$, it follows that $V_c < V_d$.

In the case where firms were entitled to fine discounts *after* the opening of an investigation, the expected profit from collusion decreases when the event 'opening of an investigation' realizes, leading firms to reveal. In the case we are considering here, instead, nothing new happens between the moment firms decide on collusion and the moment they are asked to cooperate with the authorities to break down the cartel. If Leniency Programmes are to be effective in breaking down cartels, they should be extended to benefit firms which reveal after the industry is put under monitoring.

5 Joint-Ventures and other horizontal agreements

I analyse in this section horizontal agreements other than collusive ones. Section 5.1 deals with joint ventures in general; section 5.2 looks at research joint ventures; section 5.3 at other forms of cooperation involving technology and innovations.

5.1 Joint ventures

There are a number of horizontal agreements that go under the name of *joint ventures* (JVs). These are agreements between competitors that create a new entity that carries out some activities in stead of the partners. An example of JVs are *research joint ventures*: in that case, two or more partners cease doing independent research and development (perhaps in a particular field only) to do it together, within the new entity jointly set up. There are many other possible activities that could make the object of a joint venture. *Production* JVs and *marketing and sales* JVs are other examples. In the former, two (or more) partners delegate production to a common entity; in the latter, they continue to manufacture independently, but market jointly.

Given their nature, joint ventures are a hybrid entity that can lie somehow in between cartels and mergers. Consider for instance a sales JV between two competitors, whose only purpose is to set prices or quantities in the final market without any additional activity. This is nothing else than a cartel between the two competitors (and should accordingly fall under article 81 of the Treaty in the EU and under section 1 of the Sherman Act in the US). At the other extreme, consider two competitors that give all their research, production and sales assets in a certain sector to a newly created firm whose ownership they share. Since the partners cease any independent business in the sector, the JV is akin to a merger, and should be treated accordingly by the law.[114]

Apart from the extreme cases where the JV is nothing else than a cartel, since partners simply agree on some decisions rather than carrying out some activities together,[115] the economic analysis of the effects of a joint venture is not very different than that of mergers (see chapter 5). In both cases, there might be a trade-off between market power and efficiency. For instance, an operation whereby two competitors decide to delegate their marketing and sales activities to a jointly owned entity might have anti-competitive effects because the joint venture would have larger market power than if the partners operated independently. However, it might also entail pro-competitive effects if the coordination of activities through the joint venture allows them to rationalise their distribution and marketing efforts. The net impact of the operation can be assessed only after having analysed the industry (to see whether there is indeed scope for exercising market power) and the likely efficiency gains of the joint venture. Like for mergers, a good screening device is to look at the initial market power of the partners: if they have only small market shares, there would be little point in a detailed analysis of the operation, and the joint venture should be cleared.

Additional care should however be devoted to ancillary restraints. Unlike pure mergers, where the partners disappear and their assets are entirely transmitted to the merged entity, in a joint venture partners might continue activity in the sector. Therefore, their decisions - and welfare as a consequence - might be affected by restraints to their behaviour. anti-trust authorities should avoid restraints unnecessary to the life of the JV, such as output and price restrictions in markets unaffected by the JV.[116]

5.2 Research joint ventures

Collaborative agreements between firms in research and development (R&D) activities deserve a special treatment due to the particular nature of R&D

[114] It is difficult a priori to say whether a joint venture should be considered as a horizontal agreement or as a merger by the law. In the EU, the distinction was not very clear, and I have probably not been the only one to struggle to understand the difference between cooperative and concentrative JVs (the former used to fall under article 81, the latter under the Merger Regulation). After the amendement (Regulation 1310/97) of the Merger Regulation, *full-function* JVs are treated as mergers, whereas JVs that perform only few specific functions are evaluated under article 81. See Ritter et al. (2000: ch. 6, E).

[115] See Correia (1998: 738-9) and Werden (1998: 714).

[116] See also Werden (1998: 723-725).

(more generally, of knowledge), which is often characterised by *spillovers* and *non-rivalry*.

Spillovers are a distinctive feature of R&D because technology and know-how often flow from a firm to another, for instance through imitation (and reverse-engineering), and workers' mobility. This reduces the extent to which firms can *appropriate* the results of their R&D efforts, which in turn reduces their incentive to invest in R&D. True, there exist various forms of protection of intellectual property rights (patents, copyrights, trade secret laws) but often they are imperfect.

R&D is also non-rival in the sense that it can be used by other parties without its value being diminished. Knowledge might be costly to create the first time, but once it is there, its diffusion does not modify its nature. In this sense, ex-post one would like R&D to spread as much as possible in society.

Further, one would like to avoid firms to repeat large investments to get knowledge which exists already: diffusion would avoid duplications of costs.

As a result of these two features, the market left alone is unlikely to give rise to socially optimal levels of research. Research joint ventures might help to cope with these problems and create appropriate levels of R&D. If firms collaborate in research, they will share the cost of R&D, thereby increasing their incentives to invest; they will also have immediate access to R&D output, thereby increasing diffusion; finally, they can coordinate their effort, thereby avoiding duplication of investments.

This simple reasoning provides the rationale for *research joint ventures*, even when they involve competing firms. However, allowing competitors to carry out R&D together might also have costs. For instance, the main motive behind doing R&D might be the hope to get a lead over rivals: if all the main firms in the industry take part in a cooperative venture, the incentive to carry out R&D might dramatically fall; if the joint venture does not only concern the research stage but also the production and marketing stages, cooperation might extend to the product market and result in collusive behaviour. Therefore, while there are good reasons to allow for research joint ventures among competitors, such horizontal agreements should be limited in scope.

Technical section 5.4 provides a very simple analysis of the welfare effects of research joint ventures. This is an issue which has attracted a lot of attention in the last fifteen years. Overall, this line of research shows that allowing competitors to carry out joint research projects might be a sensible policy, but with two caveats. First, such a policy is beneficial only insofar as there exist high enough spillovers. Indeed, if spillovers are high, firms are unable to appropriate their R&D effort, and co-operation increases R&D. Unfortunately, the need for clear rules and for avoiding arbitrariness and influence, on the one hand, and the difficulty of measuring spillovers on the other, make it impossible to limit the policy to particular sectors or technologies where more spillovers are likely to arise. However, a R&D project is composed of many stages, from basic research to applied research, to development, production and marketing. Externalities being higher at the earlier stages of this path, it makes sense to exclude marketing operations from R&D cooperative agreements among competitors. In other words, one should make sure that cooperation does not extend too close to the product market. One way to do it is to require that the R&D results are used independently by collaborating firms.

This leads us to the second point, which is that the authorities should make sure that the R&D cooperation does not "spill over" to the product market competition stage. It is conceivable that firms that start to work very closely on R&D projects might start to extend the coordination of their behaviour onto other spheres of the life of the firms.[117] Anti-competitive effects of a research joint venture are less likely if the collaborating firms have limited market power. Therefore, a market share criterion might be used as a proxy to screen potentially harmful ventures from others, the former having to be cleared only after a thorough investigation.

Furthermore, competition authorities should scrutinise *ancillary restraints* to the agreement for possible anti-competitive clauses. This might be the case, for instance, if the agreement setting up the RJV divides (geographical or product) markets among the participating firms, thus relaxing competition artificially; or if it calls for the payment of per-unit royalties to the joint venture: this would increase the unit costs of output, and would have the effect of increasing prices (see also below, section 5.3.1).[118] If payments are due to old patents owned by the firms, or by the joint venture, fixed rather than per-unit of output payments would avoid distortion of competition.

[117] See Martin (1985).

[118] A similar effect would have a rule which establishes that the cost of financing the joint venture falls on partners according to their ex-post output (that is, the output they sell and which incorporates the results of the joint venture). Instead, financing the joint venture on the basis of the ex-ante market shares would not distort the incentives to compete ex-post.

5.2.1 Treatment of research joint ventures in US and EU law

In the US and in the EU, research joint ventures have been seen with favour since the early Eighties. Due also to political and industrial policy reasons (it was felt at the time that the success of domestic firms in international markets crucially depended on their advances in technology, and that more incentives to innovate should be given to domestic firms), a more lenient anti-trust approach was followed for firms collaborating in R&D.

In the US, the National Cooperative Research Act of 1984 (15 USC 4301) established that joint ventures in R&D should be subject to a *rule of reason* rather than being *per se* illegal, and that - if properly notified to the anti-trust agencies - they should be subject to single rather than treble damages in case a court would find the partners guilty of anti-trust infringement.

In the EU, a *block exemption* issued in 1985 exempts from article 81 (then article 85) of the Treaty certain forms of cooperative agreements in R&D,[119] subject to certain conditions. In particular, partners should be free to use the results of R&D independently in production and distribution (to avoid restriction of competition), and their combined market share should not exceed 20% (if they are competitors).[120] Note, however, that in case the joint venture does not qualify for the block exemption, it is still possible for the parties to apply for an individual exemption to article 81.

5.3 Other forms of cooperation regarding technology

There are a number of other agreements that competing firms might have to share R&D results, or to fix technological standards. Like research joint ventures, they can have beneficial welfare effects, although some caution is needed to prevent ancillary clauses that might end up with distorting competition. In what follows, I briefly discuss them.

5.3.1 Cross-licensing and patent pooling

Cross-licensing occurs when two firms allow each other to use own technology protected by patents. In principle, it is possible that such an agreement might be used to restrict competition in the marketplace. This is the case,

[119] EC Regulation 418/85, extended by Regulation 151/93 in 1993.

[120] The R&D block exemption applies also to joint distribution if the parties' production does not exceed 10% combined market share.

for instance, when the technologies at stake are substitutable, and the cross-licensing agreements contain per-unit royalties, that reduce the incentive to market aggressively (for a technical argument see exercise 2).[121]

However, it often happens that firms have essential (or "blocking") patents that are needed for further technological progress or for final production. Suppose that two firms could introduce a new process (or manufacture a new good) that requires use of complementary technologies protected by two patents, each owned by one of the two firms. In this situation, cross-licensing might allow technological advance (or production).

Furthermore, if the patents are complementary, it would be better to have both firms having both patents. Recall that when two firms produce complementary goods each of them fails to internalise the externality that it imposes on the other, resulting in too high prices at equilibrium relative to a firm which has both (we have seen an application of this principle in the case of vertically related products, where vertical integration would get rid of double marginalisation). For complementary technologies, the same principle would hold, and cross-licensing would reduce prices (see also exercise 19).

Note also that the best situation for competition would arise when cross-licenses are royalty-free, or when they specify fixed payments rather than unit royalties, as the latter would amount to higher variable costs and reduce output.

Similar reasoning applies to a *patent pool*, which might be a firm or other organisation that holds the patent rights of two or more firms, and license them to third parties as a package. Again, if the patents are essential inputs in the technological process, their availability as a package is highly desirable, and to the extent they are complementary, having a them pooled would keep royalties lower.[122]

Patent pools might have the additional positive feature of decreasing transaction costs, as one firm would be spared time-consuming bilateral negotiations and could instead deal with another party only.[123]

---

[121] See Eswaran (1994) for a repeated game in which cross-licensing facilitates collusion.

[122] Of course, not all patent pools necessarily package only complementary patents. In the *Summit/VISX* case, the two leaders in laser eye surgery formed a patent pool and the FTC challenged the agreement because it involved competing, rather than complementary, patents (see Shapiro, 2000).

[123] See Merges (2001).

5.3.2 Cooperative standard-setting

There are many examples (from compact discs, to television, from encripting of digital musical to digital TV, from Web protocols to videocassettes) where firms that compete in the development of a new technology decide to set common standards. The question is whether such cooperation in fixing standards is welfare improving or rather anti-competitive. It turns out to be a difficult question, that has no a priori answer. Following Shapiro (2001), I briefly summarise the main arguments pro and against cooperative standard-setting. Understanding of network externalities and of their effects is crucial here (see also chapters 2 and 7).

The main benefit from common standards is that consumers will belong to the same single network. Therefore, they will be able to communicate directly with each other in physical networks (for instance, each consumer can exchange files with all computer users who have the same operative system), and they will enjoy a larger variety in indirect networks (for instance, software programmers will face a larger market and will develop a larger number of applications). It is not a small advantage, think for instance of the inconveniences if you could call only half of the population because the other half use a different telephone standard!

An additional benefit is that consumers do not risk being stranded with a product that uses a standard that turns out to be abandoned later. A standards war can create a lot of uncertainty as to which standard will win, resulting in consumers delaying their purchase for fear of being stranded with the wrong product,[124] and the market not taking off when it could.

A common standard also implies fiercer competition, since consumers will face greater choice within the same standard (although smaller choice across standards, see below). Instead, if standards were different, consumers would tend to be locked with a given product standard, and competition would be reduced.

However, setting a common standard means that there is no competition for the dominant standard. This might have several implications. First of all, it is not clear that the industry will pick up the best standard, or the one that would prevail if there was market competition for the dominant standard. A standard might be chosen on criteria that follow more the different bargaining powers of the firms involved than other, more worthy, consid-

[124] See Dranove and Gandal (2000) for an interesting account of the standard war between DVD and DVXX.

erations (government organisations might also be involved in the standard setting process, which might add bureaucracy but help find a good compromise). In other words, it is possible that competition for the standard would help select the standard which is more appreciated by the market, whereas under a cooperative standard this selection process would not occur. Second, precisely because network externalities are important, absent a cooperative standard the market will usually end up with one standard, with incompatible products disappearing altogether. After an initial phase where different standards co-exist, one will usually prevail. Anticipating the fact that eventually the market will *tip* towards one standard or the other, firms engage in intensive competition in these initial phases. Consumers will benefit from this competition, even though they will pay higher prices after the standards war has settled and one firm will prevail. This mechanism is very similar to the one which occurs under switching costs: ex-ante competition is very intense to acquire a large customer base, but ex-post competition is relaxed once consumers are locked-in with a particular product (see also chapter 2).

Therefore, cooperative standard-setting implies a clear trade-off between ex-ante and ex-post competition: it eliminates the former but it intensifies the latter. It is very difficult to conclude a priori whether the net effect is positive or not. Shapiro (2001) emphasises that the case for cooperative standard setting is very strong if a standards war is likely to delay (or fail altogether) the market because consumers fear being stranded with a product that has the losing standard.[125]

It also helps when the product to be developed needs complementary technologies owned by different firms. In this case, cooperative standard-setting would allow to pool mutually blocking patents together, with the same benefits described in the previous section.

Overall, it makes sense to allow for cooperative standard-setting, but with some caution. In particular, ancillary agreements should be carefully scrutinised. For instance, side clauses which call for cross per-unit of output royalty payments should be avoided because they tend to relax competition in the product market.

[125] However, it is also true that some negotiations for setting the standard cooperatively might be very slow. A case in point is the standard for High-Definition Television, which took ten years to be set, considerably delaying the take-off of the market for digital TV.

5.4 Cooperative R&D*

Consider a simple model due to d'Aspremont and Jacquemin (1988), to analyse the effects of cooperative agreements in R&D.

The demand function for the homogeneous good produced by the two firms is $p = a - Q$, with $Q = q_1 + q_2$, the sum of individual outputs. Firm i ($i = 1, 2$) is characterised by marginal costs $c_i = C - x_i - lx_j$, where x_i is the R&D investment made by firm i, and $l \in [0, 1]$ is a parameter that indicates the spillover from the R&D investment x_j made by the rival firm. The cost of R&D is given by the function $gx_i^2/2$, where $g > 4/3$ is a parameter expressing the efficiency of R&D production.[126] (See chapter 2 for very similar models of innovation.)

Consider two cases. In the first, firms compete both in R&D and in quantities. In the second, they cooperate in their R&D investment decisions, but then compete in the marketplace, that is they choose quantities non-cooperatively.

Competition in both stages In the first case, firms simultaneously invest in R&D at the first stage; they then simultaneously choose quantities in the second stage.

Last stage: quantity competition At the last stage of the game, each firm i chooses q_i to maximise its profit function $\pi_i = (a - q_i + q_j - c_i(x_i, x_j))q_i - gx_i^2/2$, given (x_i, x_j). It is easily checked that the (Cournot) equilibrium output is given by:

$$q_i^c = \frac{a - 2c_i(x_i, x_j) + c_j(x_i, x_j)}{3} = \frac{a - C + x_i(2 - l) + x_j(2l - 1)}{3}. \qquad (40)$$

After substitution, one can check that profits are given by:

$$\pi_i(x_i, x_j) = \left(\frac{a - C + x_i(2 - l) + x_j(2l - 1)}{3}\right)^2 - \frac{g}{2}x_i^2. \qquad (41)$$

[126] The restriction $g > 4/3$ derives from the stability conditions of the game, which are stricter than the second-order conditions (that also require g to be large enough). See next footnote.

First stage: R&D competition At the first stage of the game, each firm will choose x_i to maximise $\pi_i(x_i, x_j)$. By taking the first derivatives $d\pi_i/dx_i = 0$ and solving the system of FOCs,[127] and focusing on the symmetric equilibrium $x_i = x_j = x^c$ we obtain the equilibrium R&D level for each firm:

$$x^c = \frac{2(a-C)(2-l)}{9g-4-2l+2l^2}. \tag{42}$$

From this solution, one can then obtain by substitution the equilibrium levels of quantity and profit:

$$q^c = \frac{3(a-C)g}{9g-4-2l+2l^2}; \quad \pi^c = \frac{(a-C)^2\left(9g-8+8l-2l^2\right)}{\left(9g-4-2l+2l^2\right)^2}. \tag{43}$$

Consumer surplus is given by

$$CS^c = \frac{(a-p^c)Q^c}{2} = \frac{18(a-C)^2g^2}{\left(9g-4-2l+2l^2\right)^2}. \tag{44}$$

Finally, welfare is:

$$W^c = 2\pi^c + CS^c = \frac{4(a-C)^2\left(9g-4-4l+l^2\right)}{\left(9g-4-2l+2l^2\right)^2}. \tag{45}$$

Cooperation in R&D I will now consider the case where firms compete in quantities in the product market, but cooperate in their R&D investment decisions.[128] Since firms take non-cooperative decisions in the product market place, the last stage is the same as before.

[127] The second-order condition requires $g > 2(2-l)^2/9$. The stability conditions for the R&D stage require $(d^2\pi_i/dx_idx_{-i})/(d^2\pi_i/dx_i^2)= \left|[-2(2-l)(2l-1)]/[2(2-l)^2-9g]\right| < 1$. These are satisfied for $g > 2(2-l)(1-l)/3$. A sufficient condition for stability to be met is that $g > 4/3$, which also satisfies the second-order condition.

[128] I assume here that firms continue to have the same degree of spillovers even when they cooperate. Perhaps it would be natural to suppose that when they cooperate they also share completely their R&D results, leading to $l = 1$; see Motta (1996) for such a formalisation. In this case, the private and social benefits from R&D cooperation would be stronger.

At the first stage of the game, however, firms will choose x_1, x_2 to maximise their joint profit $\pi^J(x_i, x_j) = \sum_{i \neq j} \pi_i(x_i, x_j)$:

$$\pi^{JV}(x_i, x_j) = \sum_{i \neq j}^{2} \left[\left(\frac{a - C + x_i(2 - l) + x_j(2l - 1)}{3} \right)^2 - \frac{g}{2} x_i^2 \right]. \tag{46}$$

Focusing on the symmetric equilibrium, we find:

$$x^{JV} = \frac{2(a - C)(1 + l)}{9g - 2(1 + l)^2}. \tag{47}$$

By substitution the equilibrium levels of a firm's quantity and profits are:

$$q^{JV} = \frac{3(a - C)g}{9g - 2(1 + l)^2}, \quad \pi^{JV} = \frac{(a - C)^2 g}{9g - 2(1 + l)^2}. \tag{48}$$

We can now compute consumer surplus as:

$$CS^{JV} = \frac{(a - p^{JV})Q^{JV}}{2} = \frac{18(a - C)^2 g^2}{(9g - 2 - 4l - 2l^2)^2}, \tag{49}$$

and welfare as:

$$W^{JV} = 2\pi^{JV} + CS^{JV} = \frac{4(a - C)^2 (9g - 1 - 2l - l^2)}{(9g - 2 - 4l - 2l^2)^2}. \tag{50}$$

The effect of a joint-venture in R&D It is now possible to compare the results under the two regimes, and check that $l \geq 1/2$ guarantees that R&D investments, outputs and welfare are higher under the regime of cooperative R&D than under the regime of competition in both stages. Indeed, $x^c > x^{JV}$ amounts to the inequality:

$$\frac{18g(a - C)(1 - 2l)}{(9g - 2 - 4l - 2l^2)(9g - 4 - 2l + 2l^2)} > 0, \tag{51}$$

which holds for $l < 1/2$; $q^c > q^{JV}$ can be rewritten as:

$$\frac{6g(a-C)(1-l-2l^2)}{(9g-2-4l-2l^2)(9g-4-2l+2l^2)} > 0, \tag{52}$$

which also holds for $l < 1/2$. Finally, $W^c > W^{JV}$ can be rewritten as:

$$\frac{36g^2(a-C)^2(1-2l)(9g-4+l(36g-12-12l-4l^2))}{(9g-2-4l-2l^2)^2(9g-4-2l+2l^2)^2} > 0; \tag{53}$$

Since the last term at the numerator of the LHS is always positive (given that $l \leq 1$ and that $g > 4/3$), again the inequality holds for $l < 1/2$. In other words, a research joint-venture raises welfare if spillovers are large enough ($l \geq 1/2$).

When spillovers are large enough, non-cooperating firms anticipate that they would not be able to appropriate the result of their R&D investments, and reduce accordingly their investments. Under R&D cooperation, they internalise the effect of the spillover and this results in higher R&D expenditures.

One can check that profits are always higher under the cooperative R&D regime, which means that there would be no need for subsidies (or other financial incentives) for firms to coordinate their R&D levels: they have private incentives to do so.

Extensions and discussion This simple model has given rise to a vast literature dealing with the effects of R&D cooperation. A number of extensions have been considered, such as considering price competition, generalising cost, demand functions, and the number of firms, and so on.[129] The main qualitative results of the analysis do not change across all these specifications, though.

Leahy and Neary (1997) analyses R&D cooperation in a general model, and compares it with alternative R&D policies to increase R&D investments, such as R&D and output subsidies. They conclude that the former policy is better than the latter, but they also warn that the welfare improvement to be obtained through R&D cooperative agreements is unlikely to be of a high

[129] I confess I have contributed to raise the number of papers on this topic well above the social optimum: Motta (1992) looks at the effects of R&D joint ventures upon entry and Motta (1996) considers them in an international economics context.

order of magnitude. They also stress that R&D cooperation is privately profitable, so there is no need for further government incentives beyond allowing competing firms to engage in cooperative projects.

6 Exercises

Exercise 1 ** (From Tirole (1988)). Consider* n *firms producing homogenous goods and choosing prices in each period for an infinite number of periods. Suppose that the market meets every two periods, rather than every period. Write the incentive constraint for collusion under trigger strategies and show that collusion is less likely to hold than when the market meets every period.*

Exercise 2 ** Consider* 2 *firms producing homogenous goods and choosing prices in each period for an infinite number of periods. Each of the two firms owns a share* σ *of its rival. This share is small enough for each firm to keep full control on its own activities and decisions: the rival is a minority shareholder, who is not represented in the board and just receives a share* σ *of the firm's profits. Is the likelihood of collusion affected by this cross-ownership?*

Exercise 3 *There are only three sellers in a given industry. One day, one of the firms sends the following fax to the two competitors: "In the interest of fair competition, and for the sake of market transparency, we hereby inform you that the Board of our company has decided that from the next quarter our sales prices will be increased by 10%." Do you think that the competition authority should allow or forbid sending such faxes? Why?*

Exercise 4 *In a given sector there are* n *firms that sell directly to consumers. They sell a highly homogenous good which is subject to very frequent shocks, giving rise to high price instability. Every week, these firms communicate to the central office of the trade association the price at which they will sell the product in the incoming week. The trade association then publishes the (following week) prices of all the firms in national newspapers. The national anti-trust authority argues that this practice allows firms to exchange information thereby increasing the likelihood of collusion.*

Exercise 5 *Consider a homogenous good industry where n firms produce at zero cost and play the Bertrand game an infinite number of periods. When firms choose the same price, they earn a per-period profit $\Pi(p) = p\alpha D(p)/n$. When a firm i charges a price p_i lower than the price of all the other firms, it earns a profit $\Pi(p_i) = p_i \alpha D(p_i)$, and all other firms obtain zero profits. The parameter α represents the state of demand. Imagine that in the current period demand is characterised by $\alpha = 1$, but starting from the following period demand will be characterised by $\alpha = \theta$ in each of the following periods. All the players know exactly the evolution of the demand state at the beginning of the game. Firms have the same common discount factor, δ. (a) Assume $\theta > 1$ and consider the following trigger strategies. Each firm plays the monopoly price p_m in the first period of the game and continues to charge such a price until a profit equal to zero is observed. When this occurs, each firm charges price equal to zero forever. Under which conditions does this n-tuple of strategies represent an equilibrium? In particular, show how θ and n affect such a condition, and give an economic intuition for this result. (b) Can other prices be sustained at equilibrium under strategies similar to the ones above? Under which condition? (c) Assume now $\theta < 1$, and find the conditions under which the n-tuple of strategies delineated above represent an equilibrium.*

Exercise 6 *In a country there are three large automotive firms, which have together 100% of the market (the existence of a prohibitive tariff excludes imports). During the last weeks, these firms have been engaged in a promotional campaign in the major newspapers of the country. In these announcements, separately made by each of the firms, a firm announces a price for each of the models produced and says it will sell at such a price to any buyer who shows the newspaper cut containing the announcement. The anti-trust agency has opened an investigation, suspecting a collusive agreement. The announcements are very similar, since they use similar wording and even prices appear to be very close to each other for comparable models. Give your opinion on the basis of what has been reported here.*

Exercise 7 *Consider a sector Y where two firms sell a homogenous good at a marginal cost c. They have the same discount factor δ. The firms play the Bertrand game for an infinite number of times. That is, in each period they have to choose the price at which they sell. They face a market demand $D(p)$. When they charge the same price, one firm sells a market share*

λ (with: $1/2 < \lambda < 1$) and the other a market share $1 - \lambda$. (It does not matter why one firms has a higher market share than the other.) The two firms have the following trigger strategy. In the first period they set a price $\bar{p}$, where $\bar{p} > c$. In the following periods, they choose the same price $\bar{p}$ if both firms have chosen $\bar{p}$ in all previous periods. Otherwise, they choose the one-shot Bertrand equilibrium price forever. (a) Find the condition under which these trigger strategies represent an equilibrium. (b) Practitioners often argue that symmetry among firms facilitates collusion. Does this simple model supports their claim? (c) Under which condition the trigger strategies above can sustain a fully collusive equilibrium where the firms charge the joint-profit maximising price p_m in each period?

Exercise 8 *The time-series of price data of a given industry reveals that prices are not stable over time. During the last ten years, one can observe periods in which prices are high and periods where prices are low (about 30% less than the high price level). (a) Can one infer from the existence of periods with low prices that there is no collusion in this industry? Why? (b) Can one infer from the existence of periods with low prices that there is collusion in this industry? Why?*

Exercise 9 *Can you explain why and under what circumstances resale price maintenance can facilitate collusion?*

Exercise 10 *In a small town, Fiesole, there is a street market twice a week, all year round and every year (so we can approximate the game being played with an infinite horizon game). In a market day, n vendors sell their perfectly homogenous apples (they produced them at a marginal cost c and without fixed costs). When they arrive in the morning to the market, they each have to simultaneously hand in to the market authorities a sealed envelope where it is written at which price they will sell apples during the day. Once opening the envelopes, the prices are public, but they cannot be changed for the whole day. (Assume also that they are not capacity constrained: each of them brings to the market enough apples to satisfy all the demand at any price above marginal cost.) In another small town far away, Schriesheim, a very similar situation and game occur, with just two small differences: the street market takes place once a week only, and there are only 2 vendors that sell their (homogenous) apples (they also have marginal cost c and no fixed cost). Would you expect that a collusive equilibrium (with trigger strategies) would more likely arise in Fiesole or in Schriesheim? Show formally your point.*

Exercise 11 ** Consider 2 firms producing homogenous goods and choosing* quantities *in each period for an infinite number of periods. Demand in the industry is given by $p = 1 - Q$, Q being the sum of individual outputs. Firms are identical: they have zero constant marginal costs, and the same discount factor δ. Consider the following trigger strategies. Each firm sets the output $q \in [1/4, 1/3]$ at the beginning of the game, and continues to do so unless a deviation occurs. After a deviation, each firm sets the quantity q^{cn} which is the Nash equilibrium of the one-shot game. (a) Find the condition for collusion to arise in this industry. (b) Show that the lower q (i.e., the more collusive the output choice) the more likely the collusive trigger strategies above are sustained at equilibrium.*

Exercise 12 ** Consider n firms producing homogenous goods and choosing* quantities *in each period for an infinite number of periods. Demand in the industry is given by $p = 1 - Q$, Q being the sum of individual outputs. All firms in the industry are identical: they have the same constant marginal costs $c < 1$, and the same discount factor δ. Consider the following trigger strategies. Each firm sets the output q^m that maximises joint profits at the beginning of the game, and continues to do so unless one or more firms deviates. After a deviation, each firm sets the quantity q^c which is the Nash equilibrium of the one-shot game. (a) Find the condition for collusion to arise in this industry. (b) Indicate how the number of firms in the industry affects the possibility of reaching the tacit collusive outcome, and discuss. (c)** Do you know any other punishment strategy under which firms can sustain the collusive outcome under weaker conditions?*

Exercise 13 *Consider two perfectly identical firms that sell a homogenous product and consider collusion. The fully collusive price in the market is given by p_m, and give firms a profit π_m each. Firms also have the same discount factor δ. They play the Bertrand game an infinite number of periods. There also exists an anti-trust authority, that investigates the industry in every period. If firms collude, the authority will find them guilty with a probability p and will accordingly give them a fine F. If they are found colluding, also assume that the authority will prevent them from colluding in the future: they will forever earn market profit π_n each, where the index n stands for Nash. If firms do not collude, they cannot be fined. (a) Focus on simple trigger strategies with Nash reversal forever. Write the incentive constraints for collusion to be sustained at equilibrium, and discuss the effects that p and F have upon collusion. (b) Consider a value of the discount* *factor high enough for collusion to be sustainable. Are prices other than p_m sustainable at equilibrium of this infinite horizon game? (c) Do you know of any other strategies that could allow firms to sustain collusion under a slacker condition?*

Exercise 14 *Very often, in cartel investigations, lawyers, judges and economic experts state that collusion is more difficult in periods of declining demand than in periods of boom. (a) Can you briefly mention theoretical arguments which support this view?*

Exercise 15 *If one discovers that two or more firms change prices by the same proportion within few days one after the other(s), then this is proof that these firms have a collusive agreement.*

Exercise 16 *Consider a sector where n firms play a non cooperative game that will continue in each of the following periods with a probability α. (That is, in any period T, the game can either continue with probability α, or stop forever with a probability $1-\alpha$.) In each period where the game takes place, a firm has to choose a certain action. If all firms choose the 'collusive' action, they will all share the market and gain a profit Π_m/n. If one firm deviates from the collusive action, this firm will earn $\Pi_d > \Pi_m/n$. Firms have the following strategies. In the first period of the game, they take the collusive action. In any of the following periods where the game takes place, they choose the collusive action if all firms have obtained profits Π_m/n in any of the previous periods. Otherwise, they choose a punishment action which results in all firms making profits Π_c for all future periods of the game. (a) Show that if the discount factor δ (which is equal for all firms) is high enough, these strategies form an equilibrium of the game. (b) Show that collusion is the more likely to happen the stronger the punishment, the lower deviation profits, and the smaller the number of firms. (c) Is the collusive equilibrium the unique equilibrium of this game?*

Exercise 17 *Consider the model of section 5.4, where two firms face a demand $p = a - Q$ and have costs $A_i = (C - x_i - lx_j)q_i - (g/2)x_i^2$, where $0 \leq l \leq 1$ is a spillover parameter and $g > 4/3$ is an investment efficiency parameter. They sequentially choose first R&D levels x_i and then output levels q_i. (a) Suppose that firms could cooperate at both the R&D and the output stage of the game. Find the equilibrium R&D investments and outputs. (b)*

Suppose now that firms behave non-cooperatively in both stages. Find the equilibrium R&D and outputs. (c) Compare the equilibrium outcomes under the two regimes. [A possible variant of this exercise is to find R&D and outputs when firms behave cooperatively at the R&D stage, and then compare the results with those of full cooperation.]

Exercise 18 **(Anti-competitive Cross-licensing) Consider two firms that play the following game. In the first stage, they jointly decide whether they want to cross-license their technologies. The technologies are assumed to be perfect substitutes and are used to produce the same homogenous good. At this stage, if cross-licensing is agreed upon, they also jointly decide the same per-unit of output royalty* c_L *for the cross-license. In the following stage, they compete on quantities. Assume for simplicity that the only unit cost, if any, is given by* c_L*. Assume linear demand* $p = 1 - Q$*, where* Q *is total output. Show whether at equilibrium they will decide to cross-license and at which royalty level.*

Exercise 19 ** (Pooling of complementary patents) To produce a certain (homogenous)* n *final good manufacturers need two complementary technologies, whose patents are owned by two firms* A *and* B*, who license separately the technologies at a unit royalty fee* w_i $(i = A, B)$*. The game is as follows. In the first stage, the patent-holders independently and simultaneously decide the royalty level. In the second stage, the manufacturers compete à la Bertrand, and incur unit costs* $c + w_A + w_B$*. They face market demand* $q = 1 - p$ *(as usual, if several firms all charge the same lowest price, demand is equally shared among them; zero demand goes to firms having higher prices). (a) Find the equilibrium values of royalties and final prices. (b) Consider an alternative situation where the two patent-holders assign the right of exploitation of their patents to a* patent pool*. It is now the pool which sets the value of both royalties. Find equilibrium values of royalties and final prices under the patent pool and compare them with the previous case. (c) Show that forming the patent pool is both profitable for the patent-holders and good for consumers.*

Exercise 20 *Suppose there are two separate geographic markets,* a *and* b*, and two firms,* 1 *and* 2*, located resp. in* a *and in* b*. Unit transportation costs (productions costs are zero) from one market to the other are given by* $t < 1/2$*. Firms produce a homogenous good whose demand in country* $k = a, b$

is $p_k = 1 - 2Q_k/s$.[130] *We assume that each firm* $i = 1, 2$ *simultaneously chooses quantities* q_{ia} *and* q_{ib} *to be sold in both markets in each period of an infinite horizon game with common discount factor* δ. *Consider the following trigger strategies, which define a market allocation collusive scheme. At the beginning of the game, each firm sells only in its own market* ($q_{2a} = q_{1b} = 0$). *If a deviation occurs (that is, if a firm starts to export), then both firms revert to the Nash equilibrium forever (at the Nash equilibrium, both firms export). Find the condition for collusion and show how they vary with* t.

6.1 Solution of exercises

Solution of exercise 1. The IC becomes: $\pi(p_m)(1 + \delta^2 + \delta^4 + ...)/n \geq \pi(p_m)$. Write $\delta^2 = d$. Then the IC can be written $d \geq 1 - (1/n)$, whence: $\delta \geq \sqrt{1 - (1/n)}$. Since $\sqrt{x} > x$ when $x < 1$, the incentive constraint is less likely to be satisfied than under the standard case.

Solution of exercise 2. Yes, cross-ownership makes collusion more likely. To see why, consider that firm i's profit will be given by $\pi_i + \sigma\pi_j$, and that if firm i deviates, in the deviation period $\pi_j = 0$. The incentive constraint for collusion can then be written as: $\pi(p^m)(1 + \sigma)/\left(2(1 - \delta)\right) \geq \pi(p^m)$, which simplifies to $\delta \geq (1 - \sigma)/2$.

If there were no cross-ownership, $\sigma = 0$, and the condition would be less likely to hold.

Solution of exercise 5 Let us denote by $p^c \in (c, p_m]$ the collusive price. At time $t = 0$, $\alpha = 1$, whereas at time $t \in \{1, 2, ...\}$ $\alpha = \theta$. The incentive constraint then becomes: $\pi\left(p^c\right)\left(1 + \delta\theta + \delta^2\theta + \delta^3\theta + ...\right)/n \geq \pi\left(p^c\right)$, or, equivalently, $\delta \geq (n-1)/(n-1+\theta) \equiv \widetilde{\delta}(n, \theta)$.

(a) When firms play the monopoly price along the collusive path, $p^c = p_m$ and the condition for the n-tuple of firms strategies to be an equilibrium is given by the incentive constraint just derived, which is the same for all firms. Now, some simple algebra shows that $\partial\widetilde{\delta}(n, \theta)/\partial\theta = -(n-1)/(n-1+\theta)^2 < 0$, whereas $\partial\widetilde{\delta}(n, \theta)/\partial\theta = \theta/(n-1+\theta)^2 > 0$.

[130] This can be rationalised as follows. If there are $s/2$ consumers in market k with the same demand $q = 1 - p$, then total demand is given $q = (s/2)(1 - p)$. By inversion the demand in the text is found. Note that $s/2$ measures the size of the market.

The higher the value of θ, the higher the one-time change in demand will be. Hence, the higher is θ, the higher the present value of the stream of profits received from $t = 1$ onwards. The opportunity cost of deviation increases with θ. Hence, the higher θ is, the less likely is that firms will disrupt the collusive agreement today (at $t = 0$). On the other hand, when n increases, the tighter the incentive constraint becomes, that is, the less likely is that collusion will be sustained at equilibrium. The intuition that underlies this result is the standard one: the higher the number of firms in a collusive agreement, the more difficult it becomes to reach and sustain collusion.

(b) Notice that the incentive constraint derived above is valid for all $p^c \in (c, p_m]$. Hence, under strategies similar to the ones above, a collusive price $p^c \in (c, p_m]$ can be sustained at equilibrium if $\delta \geq (n-1)/(n-1+\theta) \equiv \widetilde{\delta}(n, \theta)$.

(c) HERE IS WHERE I HAVE DOUBTS. THIS QUESTION WOULD MAKE SENSE FOR ME IF YOU ASKED WHAT IF $\theta < 0$? THEN, $\widetilde{\delta}(n, \theta) > 1$ BECAUSE FIRMS WOULD BE VERY TEMPTED TO DEVIATE TODAY ($t = 0$: LARGE ORDER)...

Solution for exercise 7 Let us denote by s_i the market share of firm i, $i = 1, 2$. Assume, without loss of generality, that $s_1 = \lambda$ and $s_2 = 1 - \lambda$.

(a) The incentive constraint for firm i, $i = 1, 2$, is given by: $s_i(\bar{p} - c) D(\bar{p})/(1-\delta) - (\bar{p} - c) D(\bar{p}) \geq 0$, or, equivalently, $\delta \geq 1 - s_i$.

Hence, the described trigger strategies represent an equilibrium if and only if $\delta \geq \max\{\lambda, 1-\lambda\}$. But, since $1/2 < \lambda < 1$, the previous condition reduces to $\delta \geq \lambda$.

(b) Yes. Notice that in a symmetric industry structure $s_1 = s_2 = 1/2$. Therefore, the incentive constraint (which is now common for both firms) reduces to $\delta \geq 1/2$, which is the condition for the standard textbook symmetric duopoly case. Since λ is by assumption greater than $1/2$, the condition $\delta \geq \lambda$ is more stringent than $\delta \geq 1/2$. Collusion is less likely to be sustained in an asymmetric market structure.

(c) If $\bar{p} = p_m$, then the incentive constraint for firm i, $i = 1, 2$, is: $s_i(p_m - c) D(p_m)/(1-\delta) - (p_m - c) D(p_m) \geq 0$. It is straightforward to show check it amounts to $\delta \geq \max\{\lambda, 1-\lambda\}$, the same condition we have derived in (a). Thus, the same interval of values for the discount factor δ allows for a continuum of equilibrium solutions to hold.

Solution for exercise 10 Assume that one unit of time is half a week. In Fiesole, the market runs twice a week and there are n vendors. Therefore, the incentive constraint of a representative vendor in Fiesole is given by: $\pi^f(p^c)/[n(1-\delta)] \geq \pi^f(p^c)$, where $\pi^f(p^c)$ denotes the aggregate profit in the Fiesole market when all vendors set a (collusive) price $p^c > c$.

This condition in turn implies that the critical discount factor in Fiesole is given by $\delta \geq (1-1/n) \equiv \widetilde{\delta}_f$.

As far as Schriesheim is concerned, there the market runs only once a week and only two vendors exist. The incentive constraint faced by a vendor in this other small town is given by: $\pi^s(p^c)\left(1+\delta^2+\delta^4+...\right)/2 \geq \pi^s(p^c)$, where $\pi^s(p^c)$ denotes the aggregate profit in Schriesheim when both vendors fix the same collusive price $p^c > c$.

Collusion will therefore arise at equilibrium under trigger strategies if $1/(1-\delta^2) \geq 2$, which in turn implies that $\delta \geq 1/\sqrt{2} \equiv \widetilde{\delta}_s$.

Notice that in Fiesole, as compared to Schriesheim, we have more firms, which tends to hurt collusion, on the one hand, and an higher frequency of orders (market interactions), which tends to facilitate collusion, on the other. So, in order for collusion to be less likely in Fiesole than in Schriesheim, the first effect must more than compensate the second. In formal terms, $\widetilde{\delta}_f > \widetilde{\delta}_s$ if $(1-1/n) - 1/\sqrt{2} > 0$, which holds if $n \in \{4, 5, 6, ...\}$.

Solution of exercise 11 (a) First of all, it is useful to find the Cournot equilibrium of the game, which is given by $q^{cn} = 1/3$ and $\pi^{cn} = 1/9$ (see exercise **??**). By following the trigger strategies and setting output q between the joint profit maximisation output $q^m = 1/4$ and the Cournot output $q^{cn} = 1/3$, a firm has profits $\pi = (1-2q)q$. When the rival sets q, the optimal deviation can be found by choosing $q^d = \arg\max_{\widetilde{q}} \pi^d = (1-\widetilde{q}-q)\widetilde{q}$. By setting $\partial\pi^d/\partial\widetilde{q} = 0$ one can found that $q^d = (1-q)/2$, and $\pi^d = (1-q)^2/4$.

The IC can be written as $\delta \geq (\pi^d - \pi)/(\pi^d - \pi^{cn})$, and by substitution this becomes: $\delta \geq 9(1-3q)^2/(9q^2-18q+5) \equiv \overline{\delta}$, which is the condition for collusion to be sustainable.

(b) To see how the maximum degree of collusion varies with the discount factor, just compute $\partial\overline{\delta}/\partial q = -108/(5-3q)^2 < 0$. As q increases in the interval $[1/4, 1/3]$, the critical discount factor $\overline{\delta}$ decreases: the lower the degree of collusion the more easily sustainable.

Solution for exercise 12 (a) First find the quantities that maximise joint profits $\Pi = (1 - Q - c)Q$. It is easily checked that $Q = (1 - c)/2$.

Therefore, at the symmetric equilibrium individual quantities are $q^m = (1 - c)/(2n)$ and individual profits under the collusive strategy are: $\pi^m = (1 - c)^2/(4n)$.

As for the deviation profits, the optimal deviation by a firm is given by $q^d(q^m) = arg\max_q(1 - (n - 1)q^m - q - c)q$. It can be checked that $q^d(q^m) = (n+1)(1-c)/(4n)$, and that the profits obtained by deviating from the collusive output are: $\pi^d = (1 - c)^2(n + 1)^2/(16n^2)$.

Finally, it is standard to check that the Cournot quantities and profits are given by: $q^{cn} = (1 - c)/(n + 1)$ and: $\pi^{cn} = (1 - c)^2/(n + 1)^2$.

The IC for collusion is given by: $\pi^m/(1 - \delta) \geq \pi^d + \delta\pi^{cn}/(1 - \delta)$, which after substitution and rearranging becomes: $\delta \geq (1+n)^2/(1+6n+n^2) \equiv \delta^{cn}$.

Hence, under punishment strategies that involve a reversion to Cournot equilibrium forever after a deviation takes place, tacit collusion arises if and only if firms are sufficiently patient.

(b) By carrying out a simple exercise of comparative statics using the critical threshold for the discount factor, one concludes that $d\delta^{cn}/dn = 4\left(n^2 - 1\right)/\left(1 + 6n + n^2\right)^2 > 0$.

Hence, other things being equal, as the number of firms in the agreement increases, the more difficult is to reach and sustain tacit collusion (the tighter becomes firms' incentive constraint). Since firms are assumed to be symmetric, an increase in the number of firms is equivalent to a lower degree of concentration. Therefore, lower levels of concentration are associated - ceteris paribus - with less likely collusion.

(c) Yes. In order to determine the highest level of profits that a fixed number of firms can sustain as a sub-game perfect equilibrium, Abreu (1986) studied more severe punishments than reversion to Cournot forever after a deviation takes place. By being able to credibly threaten with a harsher retribution for defection, an oligopoly is able to sustain a collusive outcome for a *wider range of discount factors*. For the case of symmetric punishments, Abreu (1986) found that an optimal punishment strategy takes a very simple form. In the first period after deviation, each individual firm produces a high level of output so that all firms get negative profits. In order to induce firms to go along with the punishment first phase, the oligopoly agrees to produce the most collusive sustainable output in the ensuing periods, provided that no firm has deviated from the punishment first phase (period). Otherwise, the punishment is restarted.

Solution of exercise 13 (a) The incentive constraint each firms faces is given by: $p\,(\pi_m - F + \delta\pi_n/(1-\delta)) + (1-p)\,\pi_m/(1-\delta) \geq \pi_d + \delta\pi_n/(1-\delta)$, where π_d denotes the one-shot deviation profit. Notice that the previous condition can be rewritten as follows: $\delta\,(\pi_m - \pi_n)\,(1-p)/(1-\delta) \geq \pi_d - (\pi_m - pF)$.

Hence, collusion is self-enforcing if the long run expected loss due to the punishment are no smaller than the one shot expected net gains from deviation. Since the LHS of the previous condition increases with δ, one concludes that if the discount factor is high enough, then these strategies form an equilibrium of the game.

If p and/or F increase, then the expected fine pF increases, which in turn implies that collusion is less likely to be sustained at equilibrium, other things being equal.

(b) The described game, being a supergame, admits a continuum of solutions. If we consider exactly the same model but assume that firms, along the collusive equilibrium, set the price $\overline{p} \in (c, p_m)$, which gives firms a profit $\overline{\pi}$, then it is straightforward to show that firms' incentive constraint would be similar to the one derived in (a), the only difference being that π_m would be substituted by $\overline{\pi}$.

(c) No. Notice that the competitive outcome in this model is both the Nash equilibrium and the minmax of the stage game. Therefore, an indefinite reversion to the competitive situation constitutes the harshest credible punishment that can be imposed on deviators.

Solution for exercise 16 (a) In order for collusion to be self enforcing, then the following incentive constraint should be satisfied: $\delta\alpha(\frac{\pi_m}{n} - \pi_c)/(1-\delta) \geq \pi_d - \pi_m/n$.

Since the LHS of the previous condition increases with δ, one concludes that if the discount factor is high enough, then these strategies form an equilibrium of the game.

(b) It is straightforward to see that the incentive constraint derived in (a) is relaxed when the punishment becomes stronger (i.e., π_c decreases), the deviation profits π_d decrease and the number of firms n decrease.

(c) No. Another feasible equilibrium of this game is the one in which firms play the Nash equilibrium of the stage game at each period of the repeated game.

Solution of exercise 17. (a) The joint profit of the two firms are $\pi^M = \sum_{i=1}^{2} [(a - q_i + q_j - C + x_i + lx_j)q_i - gx_i^2/2]$ and the problem of the firm is to choose q_i, q_j which maximise it. Solving the FOCs $d\pi^M/dq_i = 0$ and focusing on the symmetric solution gives $q = ((a - C) + (1 - l)x)/4$. At the first stage of the game, the joint profit becomes $\pi^M = ((a - C) + (1 - l)x)^2/16 - gx^2$. Maximisation requires $d\pi^M/dx = 0$, which is solved by: $x^M = (a - C)(1 + l)/[4g - (1 + l)^2]$, and after replacement this gives the equilibrium output: $q^M = (a - C)g/[4g - (1 + l)^2]$.

(b) This is done in the text.

(c) Comparison of the equilibrium levels reveals that $x^M > x^C$ for $l > 7/17$, whereas $q^M > q^C$ for $3g+1-4l-5l^2 > 0$. Again, l must be high enough for monopoly to give rise to a superior outcome: if $l < (-2+\sqrt{29})/5 \simeq .677$, then $q^M < q^C$ whatever g. But for higher levels of the spillover parameter, there exist combinations of g and l for which full cooperation is better. Intuitively, there is a trade-off under full cooperation, between higher R&D due to internalisation of the spillover which restores R&D incentives, and lack of competition in the product market which reduces equilibrium outputs. The former effect is stronger when spillovers are very high.

[The variant of the exercise involves finding R&D cooperation equilibrium R&D levels and outputs, that is done in the text, and then compare them with those obtained in the full cooperation regime. It is easy to check that $x^M > x^{JV}$ for all l, whereas q^M is always strictly lower than q^{JV} ($q^M > q^{JV}$ amounts to $g > (1 + l)^2/3$ which is always true since $g > 4/3$ and $l \leq 1$).]

Solution of exercise 18 At the last stage of the game, the Cournot equilibrium is found by solving each firm' problem $\max_{q_i} \pi_i = (1 - q_i - q_j - c_L)q_i + c_L q_j$. Note that each firm has to pay the other a unit royalty on the own output, which implies that the royalties appear both as a cost and as a revenue in the profit function. From the FOCs one obtains the equilibrium output at the symmetric equilibrium as: $q^*(c_L) = (1 - c_L)/3$. The associated per-firm profits are $\pi^*(c_L) = (1 + 2c_L)(1 - c_L)/9$.

At the first stage, firms jointly decide on the cross-licensing agreement and on the level of c_L. Note that the function π^* reaches its maximum in $c_L = 1/4$. The firms will therefore have an incentive to cross-licensing and will optimally choose $c_L = 1/4$. Their outputs and profits will be equal to $q^*(1/4) = 1/4$ and $\pi^*(1/4) = 1/8$. These values correspond to the jointly profit maximisation solutions. In other words, by cross-licensing the firms manage to reach the monopoly solution, improving upon the Cournot solution obtained when $c_L = 0$: $q^*(0) = 1/3$ and $\pi^*(0) = 1/9$.

Solution of exercise 19 (a) In the last stage, given that manufacturers compete in prices, the Bertrand equilibrium applies: the market price will be $p = c + w_A + w_B$, and final demand $q = 1 - (c + w_A + w_B)$.

In the first stage, each patent-holder decides the royalty fee so as to $\max_{w_i} \pi_i = w_i(1 - c - w_i - w_j)$. From $d\pi_i/dw_i = 0$ it follows that the symmetric equilibrium is $w^* = (1-c)/3$, and the final price (by substitution) is $p^* = (2+c)/3$. Patent-holders' profits are $\pi^* = (1-c)^2/9$.

(b) Under the patent pool, there is joint profit maximisation of the patent-holders. The pool's problem is therefore $\max_{w_i,w_j} \pi_P = w_i(1 - c - w_i - w_j) + w_j(1 - c - w_i - w_j)$. Solving the FOCs gives the symmetric solution $w^P = (1-c)/4$. By substitution prices and (per-firm) profits are obtained as: $p^P = (1+c)/2$ and $\pi^P = (1-c)^2/8$.

(c) It is straightforward to see that the patent pool Pareto dominates the situation where the two patents are licensed independently. Final prices (as well as royalties) are lower (therefore, consumers are better off) and patent-holders' profits are higher. (Manufacturers in this example always get zero profits.)

Solution of exercise 20 *Collusion.* Let us first derive the market allocation scheme solution when each firm sells only in its own market. In this case, profits are given by $\pi = (1 - 2q/s)q$. From $d\pi/dq = 0$ it follows that $q_{1a} = q_{2b} = s/4$, and $q_{1b} = q_{2a} = 0$, $\pi_1^m = \pi_2^m = s/8$.

Punishment. If firms were playing the one-shot Nash game in quantities, they would sell in both markets. Their profits would be: $\pi_1 = p_a(q_{1a}, q_{2a})q_{1a} + (p_b(q_{1b}, q_{2b}) - t)q_{1b}$ and $\pi_2 = p_b(q_{1b}, q_{2b})q_{2b} + (p_a(q_{1a}, q_{2a}) - t)q_{2a}$, where $p_k = 1 - 2(q_{1k} + q_{2k})/s$, for $k = a, b$. Firm i's programme is $\max_{q_{ia},q_{ib}} \pi_i$. Therefore, finding the equilibrium solutions involves solving the system of four equations $d\pi_i/dq_{ia} = 0$, $d\pi_i/dq_{ib} = 0$, for $i = 1, 2$.

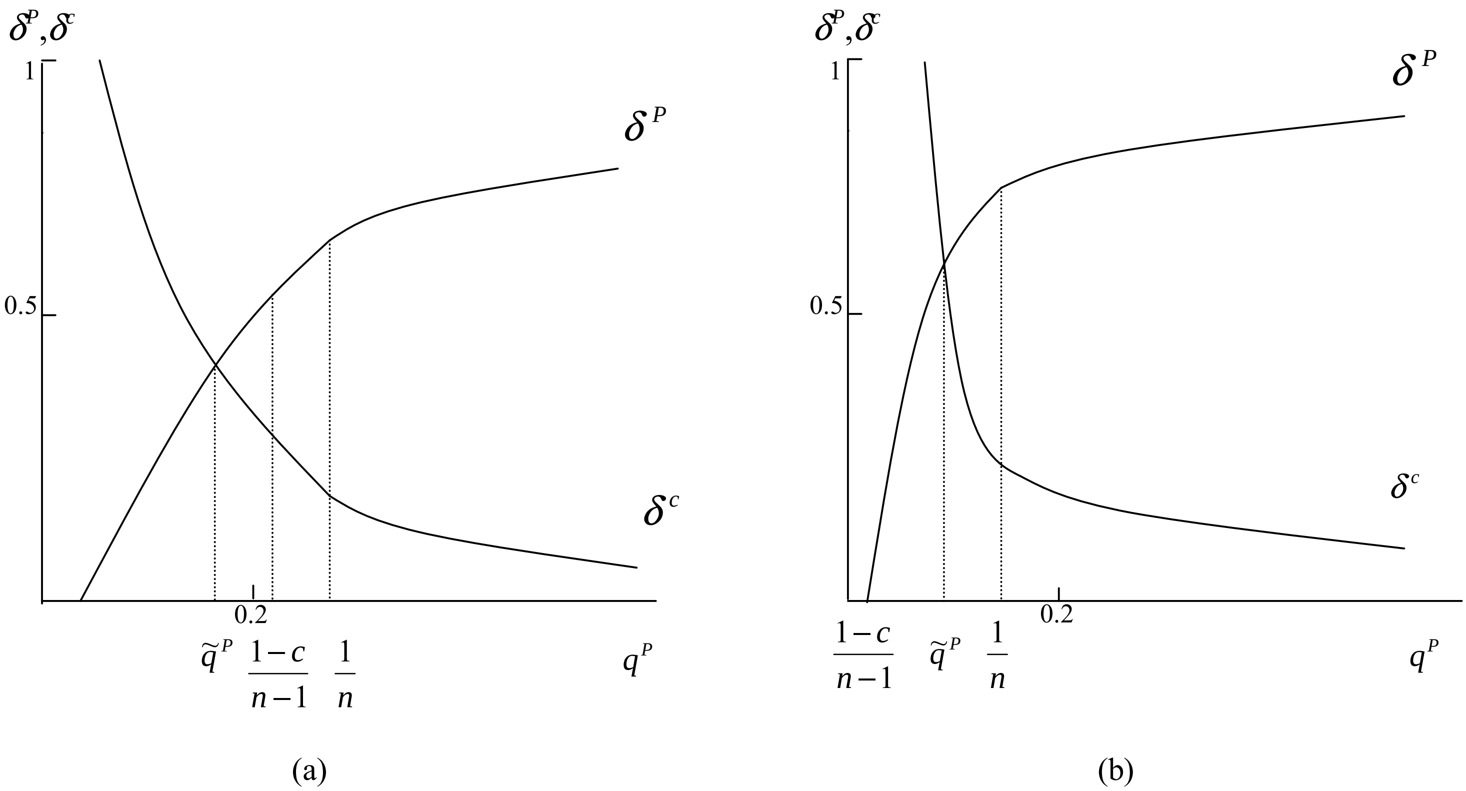
It is easy to check that the solution is given by the four prices: $q_{1a} = q_{2b} = s(1+t)/6$, $q_{2a} = q_{1b} = s(1-2t)/6$, and that firms profits are given by: $\pi_1^e = \pi_2^e = s\left[(1+t)^2 + (1-2t)^2\right]/18$.

Deviation. The optimal deviation of, say, firm 1 consists of setting its monopoly quantity in the home market a, but choosing the (positive) output q_{1b}^d that maximises its export profits π_{1b} given that the rival sets

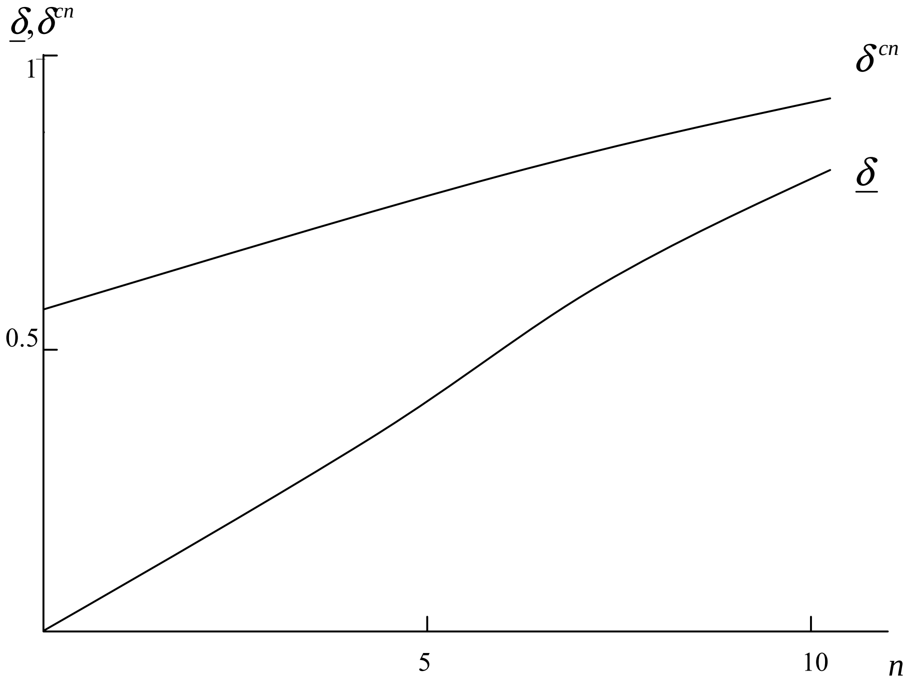
$q_{2b} = s/4$. In other words, q_{1b}^{d} is the quantity q that maximises $\pi_{1b} = (1 - t - 2\left(q + s/4\right)/s)\, q$. From $d\pi_{1b}/dq = 0$ it can be found that $q_{1b}^{d} = s(1-2t)/8$, and that $\pi_{1}^{d} = s/8 + s(1-2t)^{2}/32$ (the first component of the profit being the home monopolistic profit).

The incentive constraint for collusion is given by the standard condition $\pi_i^m/(1-\delta) \geq \pi_i^d + \delta\pi_i^e/(1-\delta)$, from which the critical discount factor is derived as $\delta^e = \left(\pi_i^d - \pi_i^m\right)/\left(\pi_i^d - \pi_i^e\right)$. After substitution, it can be checked that $\delta^e = 9\left(1-2t\right)/\left(13+22t\right)$, so that collusion can be achieved only if $\delta \geq \delta^e$.

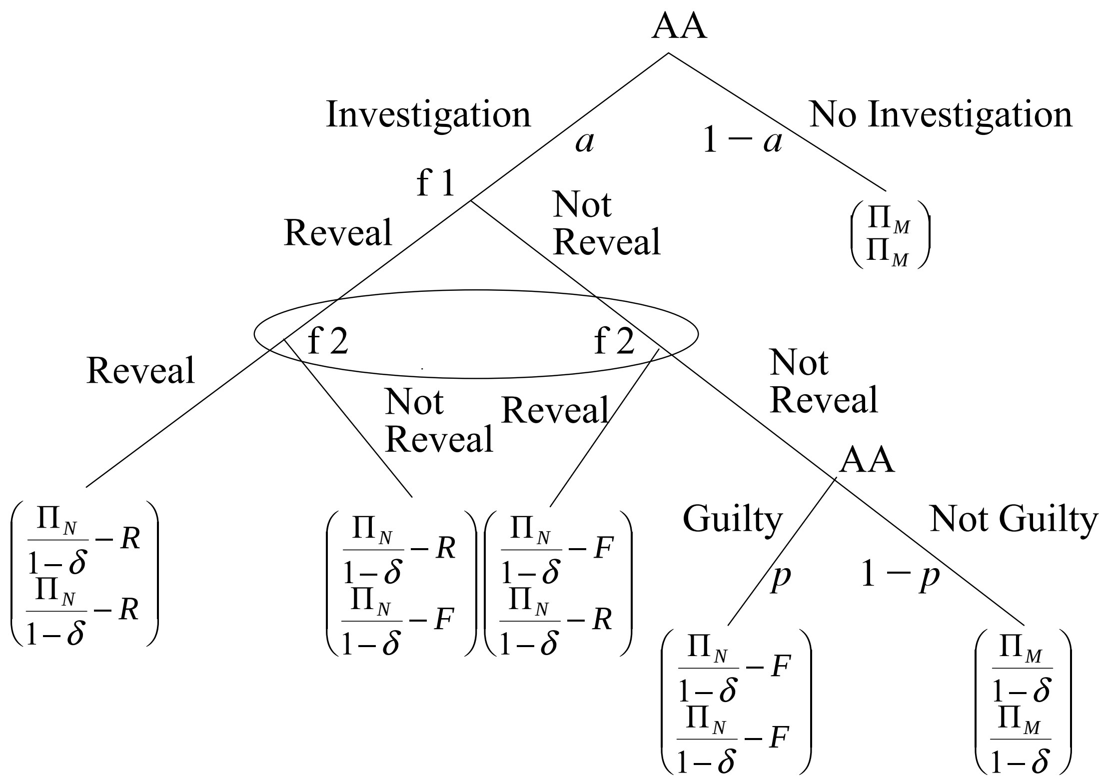
Note that $d\delta^e/dt < 0$, so that as transportation costs increase, collusion becomes easier to sustain. This is because the higher the transportation cost the smaller the market share it can be obtained through a deviation. In the limit case where $t = 1/2$, the transport cost becomes prohibitively high, so that no exports would occur at the one-shot equilibrium: a firm would not have any temptation to deviate from the (monopolistic) autarky equilibrium.

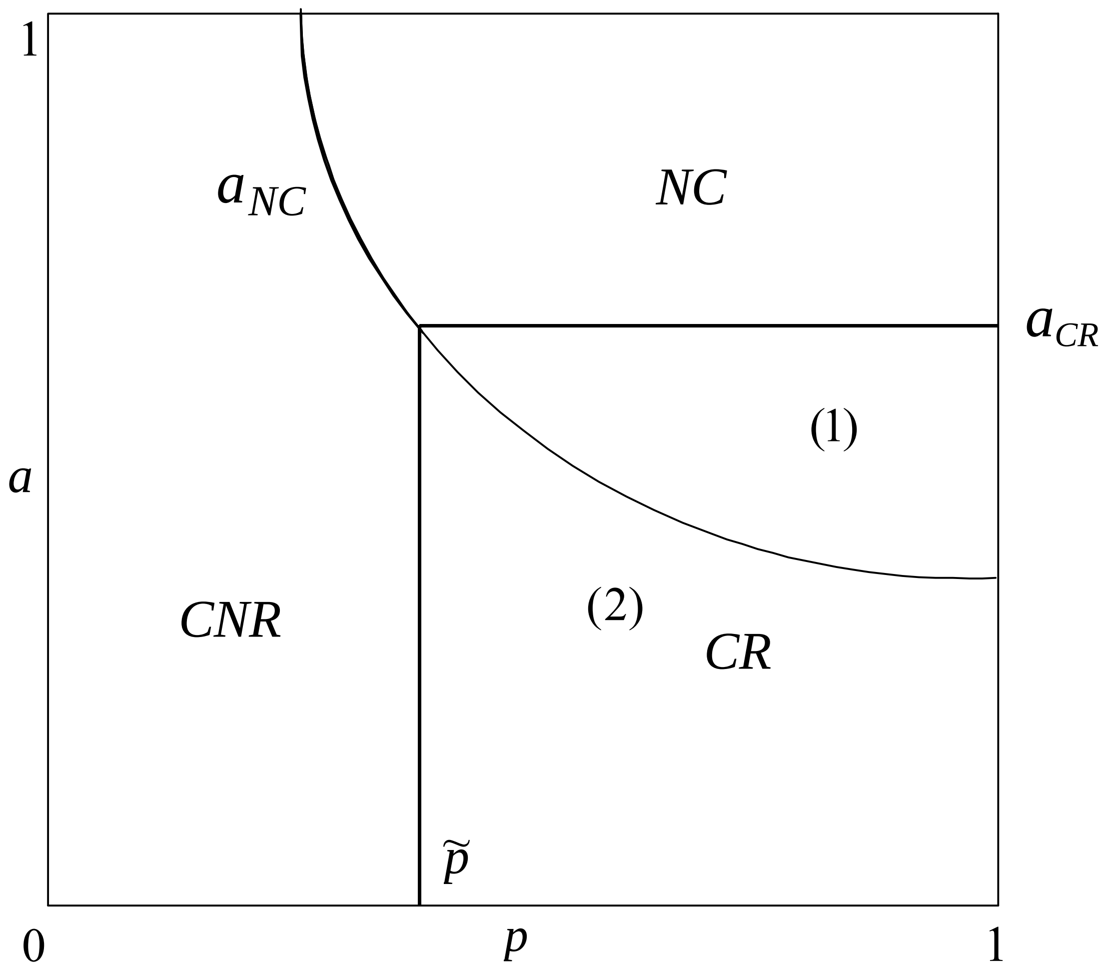


δ^P, δ^c
1
0.5
0.2
δ^P
δ^c
$\tilde{q}^P$
$\frac{1-c}{n-1}$
$\frac{1}{n}$
q^P
(a)
δ^P, δ^c
1
0.5
0.2
δ^P
δ^c
$\frac{1-c}{n-1}$
$\tilde{q}^P$
$\frac{1}{n}$
q^P
(b)



$\underline{\delta}, \delta^{cn}$
1
0.5
δ^{cn}
$\underline{\delta}$
5
10
n


AA
Investigation
a
1 − a
No Investigation
f 1
Reveal
Not Reveal
f 2
f 2
Reveal
Not Reveal
Reveal
Not Reveal
AA
Guilty
Not Guilty
p
1 − p
$\begin{pmatrix} \Pi_M \\ \Pi_M \end{pmatrix}$
$\begin{pmatrix} \frac{\Pi_N}{1-\delta} - R \\ \frac{\Pi_N}{1-\delta} - R \end{pmatrix}$
$\begin{pmatrix} \frac{\Pi_N}{1-\delta} - R \\ \frac{\Pi_N}{1-\delta} - F \end{pmatrix}$
$\begin{pmatrix} \frac{\Pi_N}{1-\delta} - F \\ \frac{\Pi_N}{1-\delta} - R \end{pmatrix}$
$\begin{pmatrix} \frac{\Pi_N}{1-\delta} - F \\ \frac{\Pi_N}{1-\delta} - F \end{pmatrix}$
$\begin{pmatrix} \frac{\Pi_M}{1-\delta} \\ \frac{\Pi_M}{1-\delta} \end{pmatrix}$


1
a_{NC}
NC
a_{CR}
(1)
a
CNR
(2)
CR
$\tilde{p}$
0
p
1