Leniency Programs and Cartel Prosecution

Massimo Motta
European University Institute, Florence
Universitat Pompeu Fabra, Barcelona

Michele Polo
Bocconi University and IGIER, Milan *

February 13, 2001

Abstract

We study the enforcement of competition policy against collusion under Leniency Programs, which give reduced fines to firms which reveal information to the Antitrust Authority. Leniency Programs make enforcement more effective but they may also induce collusion, since they decrease the expected cost of misbehaviour. We show that in the optimal policy the former effect dominates, calling for Leniency Programs when the Antitrust Authority has limited resources. We also show that these Programs should apply to firms that reveal information even after an investigation is started.

*We benefited from discussions with Luis Cabral, Federico Ghezzi, Patrick Rey, Giancarlo Spagnolo, Thomas Von Ungern-Sternberg and seminar participants at the University of Salerno, Bocconi University, the European University Institute and the 1999 EARIE Meeting.

1 Introduction

The enforcement of competition policy against collusion and price fixing agreements is one of the main fields of antitrust intervention. In the design of the policy we find today richer and more complex mechanisms than those based simply on an increase in fines. Since 1978 the US Antitrust Division of the Department of Justice has allowed for the possibility of avoiding criminal sanctions if specific conditions occurred. In 1993 this policy has been redesigned in the *Corporate Leniency Policy*, which establishes that criminal sanctions can be avoided in two cases: either if a colluding firm reveals information before an investigation is opened, as it was in the previous regime, or if the Division has not yet been able to prove collusion when a firm decides to cooperate.[1] The new Leniency Policy has shown in the first years of application a significant success in terms of the number of cases that the Division has been able to open and successfully conclude.

The European Union has introduced in 1996 a new regulation in which more generous fine reductions can be given to firms which cooperate with the antitrust authority *before* an inquiry is opened, by providing evidence of a collusive agreement in which they have been involved, while limited reductions can be granted if cooperation occurs after the opening of a case.[2]

Although it is too early to evaluate the effects of this new policy, this paper argues that - as also the US experience suggests - this leniency program might lead to more effective enforcement against cartels. However, our analysis also indicates that the design of a leniency program might be improved by extending it to cover situations where firms reveal information even after an investigation has started.

More generally, the objective of this paper is to investigate the deterrence

[1] Some additional restrictions on the firms entitled to benefit from this regime are introduced, as the fact that only the first can be given a fine reduction, and that it must be a junior partner in the cartel.

[2] See European Union (1996). To be more precise, a 75-100% reduction in fines can be given if firms reveal information before an inquiry is opened; a lower reduction (50-75%) can be granted if cooperation occurs after an investigation has started, *but that investigation has failed to provide sufficient grounds for initiating a procedure leading to a decision*; a 10-50% reduction in fines can be given for partial cooperation, such as providing additional evidence or not contesting the facts on which the Commission bases its allegations. Notice that while in the US the regime applies to criminal sanctions (which include both fines and incarceration), in the EU reductions are referred only to monetary fines. Criminal sanctions do not exist under EU competition law.

and desistence properties of a Leniency Program[3] in antitrust cases.[4] The Antitrust Authority is motivated by the maximization of social welfare and aims at minimizing the occurrence of collusion among firms by committing on a set of policy parameters which includes full and (possibly) reduced fines and an allocation of internal resources that determines the probability that a cartel is reviewed and the probability that it is found guilty. After observing the Antitrust Authority's decisions, firms play an infinitely repeated game where they decide in the first period whether they want to deviate or collude, and, in the following periods, whether they want to reveal information about the cartel to the Authority or not. In particular, we consider two possible collusive strategies. One prescribes firms *never to reveal* (if one firm revealed, it would trigger Nash reversal forever) and to go back to collusion after a possible condemnation, the other requires firms *to reveal* to the Authority as soon as an investigation review is opened, and to go back to collusion after the procedure is closed.

In this setting we show that a Leniency Program might have two possible effects, depending on the policy parameters chosen. The positive effect is that it might lead firms to desist from colluding. If it convinces firms to reveal information whenever an investigation is opened, society will benefit not only because of a (temporary) cessation of collusive pricing, but also because the Authority's resources are saved (information received by firms brings about the punishment with certainty and allows to avoid the costly prosecution stage of the investigation). But the Leniency Program might also give rise to a perverse effect. Since it allows colluding firms to pay reduced fines, it may have ex-ante a pro-collusive effect, given that it decreases the expected cost of anticompetitive behaviour. A priori, therefore, it would be difficult to conclude that a Leniency Program unambiguously increases welfare, without considering which policies are implementable and desirable.

Our analysis of the optimal enforcement policies shows that, if the Antitrust Authority has limited resources, and is therefore unable to prevent collusion ex-ante, the use of Leniency Programs improves welfare, by sharply increasing the probability of interrupting collusive practices and by shortening the investigations. Hence, in a second best perspective, fine reductions may be desirable because they allow to better implement ex-post desistence

[3] In this paper we use the term Leniency Programs as referring to a reduction in monetary fines.

[4] After our work, other papers on the use of leniency programs in antitrust have been written. See Rey (2000) for an excellent survey.

from collusion and put saved resources to better use.[5]

The key mechanism of Leniency Programs is the rule that allows firms to receive fine reductions even after an investigation is opened. In this situation, the probability of paying the fine increases compared to the case when firms are not yet under scrutiny, and the exchange of reduced fines with cooperation becomes attractive. Conversely, we prove that limiting eligibility to the case where the inquiry has not yet been opened eliminates the incentive to reveal and the effectiveness of the program.

The enforcement problem we study has several ingredients: We analyze the design of self-reporting incentives, having a group of (and not a single) defendants and considering ongoing (and not single episode) infringements and benefits. This paper is therefore related to several strands of literature that have often considered some of these features separately. The closest to our work are perhaps the studies on optimal enforcement under self-reporting schemes. Malik (1993) and Kaplow and Shavell (1994) are probably the first to have identified the potential benefits of schemes which elicit self-reporting by violators.[6] Self-reporting may reduce enforcement costs[7] and improve risk-sharing, as risk-averse self-reporting individuals face a certain penalty rather than the stochastic penalty faced by non-reporting violators (who pay only if caught).[8] There are two main differences between these papers and ours. First, they consider individual violators, rather than a group of violators like our colluding firms, which requires us to analyse in a game theoretic setting the conditions for self-reporting. Second, they deal with an illegal action which is taken and gives benefits only once, whereas we analyse ongoing infringements and benefits. These two elements explain

[5] If resources are very low, the Authority is not able to credibly prove firms guilty with a probability sufficiently high to induce revelation, and Leniency Programs become ineffective.

[6] A recent paper in this field is Innes (1999), who considers an extension of the environmental self-reporting schemes.

[7] In Malik (1993), who applies self-reporting to environmental violations, self-reporting decreases auditing costs but increases penalty costs. It is the relative importance of the auditing and punishment technologies which determine the desirability of the scheme. Kaplow and Shavell (1994) also note that if the imposition of penalties occurs more frequently under self-reporting, administration costs may increase.

[8] See also Arlen and Kraakman (1997) who analyse the effects of different corporate liability regimes, and the incentive schemes for corporations to monitor and report wrondoings of their employees. At the other extreme, Tokar (2000) analyses whistleblowing of employees, who under the US False Claims Act are given monetary incentives to file cases against employers which defraud the US Government.

why - unlike the earlier papers on self-reporting - an optimal programme might be one which gives generous penalty discounts in case of collaboration with the Antitrust Authority. A similar result can be found in Livernois and McKenna (1999), where a repeated game is played between the regulator and a polluting firm, and where by self-reporting a firm will return to compliance, which decreases future profits.

Another strand of literature related to our paper is that on plea-bargaining, where an individual is given the option to plead guilty in exchange of a less harsh penalty rather than waiting for a court decision. Landes (1971) has showed that this allows to save the prosecution costs (a motive which appears in our paper as well), while Grossman and Katz (1983) have identified the possible beneficial insurance and screening effects of settlements.[9] In the plea bargaining literature[10] the enforcer balances the goal of condemning the guilty agents and not condemning the innocent ones with the minimization of resources devoted to enforcement. However, the issue of deterrence is generally not addressed: agents have (possibly) already committed a crime, and in most papers, whether the agent is innocent or guilty and how strong is the evidence against him (agent's type) is exogenous. The effects of the legal procedures on preventing the crime (collusion) or making it to cease are instead at the center of our analysis. In order to focus on deterrence, we make the simplifying assumption that there are no judicial type-I errors (innocents will never be found guilty) and that firms are symmetric (either they are all guilty or they all innocent). Consequently, we cannot address the insurance value or the possible self-selection effects of Leniency Programs[11].

Our work also shares some features with studies on multi-defendant settlements, where a single plaintiff faces many defendants, a literature initiated by Easterbrook, Landes and Posner (1980) and Polinski and Shavell (1981). In particular, Kornhauser and Revesz (1994) analyse the case where there exists joint and several liability and the plaintiff's probabilities of success are highly correlated across defendants. Their setting presents game theoretic aspects similar to the case we analyse, as the decision of one of many defendants between settling or not is very similar to the decision of one of many colluding firms between revealing or not information to the Antitrust

[9] But if innocent defendants are more risk-averse than guilty defendants, the former might plead guilty even if they are not.

[10] See also Reinganum (1988) for a plea bargaining model with asymmetric information.

[11] For the same reason, we do not have multiple equilibria with different crime rates, as identified by Schrag and Scotchmer (1997).

Authority.[12]

Finally, the issues studied here have some relationships with the literature on tax amnesties, even if the models used in that literature are very different from our own.[13] In particular, despite the different settings, we believe that our results might represent a contribution to the understanding of the effects of fully anticipated (or permanent) tax amnesties. While they might have the effect of reducing compliance, they could still be beneficial in a second-best perspective, when the tax authorities have not enough resources to avoid tax evasion.

The paper continues as follows. In section 2 we set up the basic model, in which every firm which decides to cooperate with the Antitrust Authority is given a fine reduction. In section 3 the firms' decisions given the policy parameters are studied, while in section 4 we analyze the optimal policies. Section 5 deals with the case where leniency applies only if information is disclosed before an inquiry is open; concluding remarks follow in section 6.

2 The Model

We consider a group of perfectly symmetric firms (an industry) which consider colluding taking into account the enforcement activity of the Antitrust Authority (AA from now on). The AA is able to commit to a certain enforcement policy, which might entail the use of Leniency Programs (LP hereafter). LP grant reduced fines to those firms which cooperate in the investigation by revealing information which proves the existence of a collusive agreement.[14] The content of the collusive agreement, therefore, has to prescribe both the market conduct and the behaviour towards the AA. A cartel, for example, may prescribe to its members to replicate the monopoly configuration and

[12] See also Kobayashi (1992), where by offering a plea discount to one defendant the prosecutor obtains information which raises the probability of conviction of other defendants. Kobayashi shows that if the culpability of an individual is positively correlated with the amount of incriminating evidence about the other defendants, the prosecutor will offer the highest plea discount to the most culpable defendant, to maximise deterrence. While our symmetric setting does not allow us to analyse this case, this result would carry over to a properly modified extention of our model.

[13] See Andreoni (1991), Malik and Schwab (1991) and Das-Gupta and Mookherjee (1996).

[14] Antitrust law in most countries prohibits collusive agreements among firms independently of their success in restricting competition. Accordingly, revealing information to the AA consists of reporting evidence of any coordination device established by the firms to promote and/or renew collusion. This would be considered as a proof of collusion.

to refuse any cooperation with the AA during the inquiries, or, conversely, it may allow the members to reveal information if the AA opens a review of the industry.

We now describe the policy choices of the AA, moving then to the firms' strategies and to the timing of the game.

Enforcement choices The AA goal is the maximization of a utilitarian welfare function. Four parameters summarize the enforcement policy.

- The full fines $F \in [0, \overline{F}]$ for firms that are proved guilty and that have not cooperated with the AA, where $\overline{F}$ is exogenously given by the law.
- The reduced fines $R \in [0, F]$ specified by a LP together with the eligibility conditions.[15] We shall consider initially the benchmark case in which all [16] the firms that cooperate even after an investigation is opened can be granted reduced fines R.[17]
- The probability $\alpha \in [0, 1]$ that the firms are reviewed by the AA. (This review - or monitoring - stage is the first stage of an investigation.)
- The probability $p \in [0, 1]$ that the AA successfully concludes the investigation when firms do not cooperate. (This prosecution stage is the second and last stage of the investigation.)

Once opened the investigation, the AA has to conclude it with a decision. We assume that the AA does not commit (type I) judicial errors: if an industry where firms are not colluding is investigated, the firms will

[15] Spagnolo (2000), who builds on the previous version of this paper (Motta and Polo (1999)), considers the case of negative fines (i.e., rewards) for firms which provide information to the AA. However, offering rewards for firms that have colluded is generally politically unfeasible. It also raises a number of other issues, such as the possible creation of further incentives for collusion.

[16] Throughout the paper, we 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 interpreted as the case where each firm has access to the minutes of the meetings which take place among all the colluding firms, or has copies of letters, faxes or e-mail messages which all the firms have used to coordinate on the collusive outcome. Since an important component in the working of cartels is the coordination of moves among participants, the access of each partner to some information regarding the others seems realistic.

[17] In Motta and Polo (1999) we also consider the case where only the first comer is eligible for leniency. We find there that restricting the LP to the first firm is sub-optimal, but otherwise results are qualitatively similar.

be acquitted with probability one and the investigation does not enter the prosecution stage. An investigation on colluding firms can be ended in two ways: either some cartel member reveals information to the AA, in which case the participants are found guilty with probability 1 (and there is no need to enter the prosecution stage), or nobody reveals information. In this case the AA has to go on with the investigation, trying to prove the firms guilty, which occurs with probability p (type II errors might occur).

We assume that the AA has an exogenous budget that can be used to promote enforcement: hence, we have fixed rather than variable enforcement costs. The important decision regarding the policy parameters will be the allocation of internal resources, which determines the trade-off between the monitoring and prosecution rates α and p . In section 4 we discuss in details the enforcement technology and costs.

Moreover, when the AA proves firms guilty, it is able to impose compliance for a certain (one) period, for instance by requiring reports on prices and market strategies, obtaining competitive behaviour.

We initially treat the policy parameters as exogenous, focusing on the game played by the firms once the policy is set. When moving to the analysis of the optimal policies we will describe the constraints of the AA and explain how the policy parameters are determined.

Firms' collusive strategies We analyse two different collusive strategy profiles of firms.

- In the first one, CR (Collude and Reveal), firms start colluding in period 1, with per period profits Π_M and will collude in the market if no deviation occurs and no inquiry is opened; if the AA opens a review, firms will reveal information to the AA, pay the (reduced) fine R and compete non cooperatively for one period, with profits $\Pi_N < \Pi_M$. In the period after revealing and being fined, since no deviation from the equilibrium strategy occurred and if the AA is monitoring another industry, they go back to the collusive action. If a deviation either in the marketplace or in the revelation policy occurs, they use Nash punishment forever with profits Π_N in every period.

- In the second collusive strategy, CNR (Collude and Not Reveal), firms start colluding in period 1 and will collude in the market if no deviation occurs and no inquiry is concluded; if a review is opened, they do not

reveal any information to the AA (and collude during the investigation); if they are proved guilty, they will pay the fine F, compete non cooperatively for one period and then revert to the collusive behaviour; if they are not proved guilty, they go on colluding. If, however, a firm deviates in the marketplace or reveals information to the AA, Nash punishment starts and goes on forever.

Hence the firms combine the usual grim strategies of the supergame literature[18] with a revelation policy which is agreed upon, and they interpret any deviation from either the market strategy or the revelation policy as a break-down of the cartel. Moreover, firms collude until they are not proved guilty, and restart collusion after an inquiry is concluded, as long as no deviation from the prescribed strategy occurred. These strategies are consistent with the idea that if the conditions for collusive behaviour are satisfied, firms tend to coordinate their actions as long as they are not forced to take non-cooperative actions by a sentence of the AA, and they restart collusion once the AA moves its attention to other industries.[19]

Timing of the game To summarise, the game starts with the AA setting the policy parameters; then firms decide whether to collude or to deviate from the proposed agreement.[20] Afterwards, the AA opens an investigation with probability α, which can take one period if the firms cooperate, or two periods if no firm reveals. If firms are condemned, they are forced to play non cooperatively for one period and to pay the (reduced or full) fine. After an investigation is concluded, the game restarts[21] with the collusive strategy if no deviation from the equilibrium strategy occurred, while it goes on with the punishment phase if some firm deviated either from the market or from the revelation strategy.

We can now proceed to analyse the equilibrium of the game. We first consider (section 3) the subgame perfect equilibria in the repeated game

[18] See Friedman (1971) and, for a textbook presentation, Tirole (1988, ch. 6) .

[19] In Motta and Polo (1999) we consider an alternative scenario, in which a firm proved guilty does not collude anymore, while if the AA is unable to condemn the firms, the cartel is no more reviewed in the future. That case implies that if found guilty, a firm is costantly monitored by the AA, while a second inquiry is never opened if the firms were found innocent (not proved guilty). The results do not change in this alternative scenario.

[20] The game once set the policy parameters is stationary and therefore we can restrict the attention to deviations at $t = 1$.

[21] Notice that if, after being condemned, firms want to go back to collusion and they are reviewed again, they can reveal the evidence on the ongoing cartel to the AA.

among firms once the policy parameters F, R, α and p have been set by the AA. We shall identify in the (α, p) space the regions corresponding to the different equilibria, which identify the incentive compatibility constraints when the AA designs the optimal policy. In section 4 we shall consider the optimal policy choices of the AA, thus finding the solutions of the whole game.

3 The firms' decisions

¿From the description of the strategies, three possible outcomes are relevant. In a NC (No Collusion) equilibrium collusion does not arise, because each participant prefers to deviate rather than to join a collusive agreement. In this case full ex-ante deterrence is reached. Alternatively, firms collude and reveal if monitored (CR) or they collude but refuse to reveal any information if an investigation is opened (CNR): in both cases a cartel starts, and the AA obtains ex-post desistence (for one period) when it is able to condemn the firms. When examining the conditions for the existence of a collusive equilibrium, we have to consider two possible deviations: a deviation from the market strategy, and a deviation from the revelation policy agreed upon by the firms. In what follows, we study the incentive compatibility constraints of the firms and determine the equilibrium outcomes.

3.1 CR: Collude and Reveal

We consider first the conditions for a CR equilibrium to exist, in which firms collude in the market and reveal information to the AA if a review is opened. From the timing of the game, we can easily obtain the value of the CR strategy, V_{CR}:

$$V_{CR} = \Pi_M + \delta \tilde{V}_{CR} \tag{1}$$

where Π_M are the profits from collusion and $\delta \in (0, 1)$ is the discount factor. $\tilde{V}_{CR}$ is:

$$\tilde{V}_{CR} = \alpha(\Pi_N - R) + (1 - \alpha)(\Pi_M) + \delta \tilde{V}_{CR} = \frac{\alpha(\Pi_N - R) + (1 - \alpha)\Pi_M}{1 - \delta} \tag{2}$$

where $\Pi_N < \Pi_M$ are the non cooperative profits. The first term gives the profits when the firm reveals, is condemned to pay R and plays the

non cooperative Nash equilibrium for the current period; the second term corresponds to the profits if no inquiry is opened in the period. From the next period, the game restarts. Substituting in the previous expression we obtain

$$V_{CR} = \frac{\Pi_M}{1-\delta} - \frac{\alpha\delta(\Pi_M - \Pi_N + R)}{1-\delta} \tag{3}$$

Notice that the first term corresponds to the value of collusion in the standard case where no antitrust intervention is considered. The value of collusion becomes smaller when antitrust enforcement takes place for two reasons: the firms pay the fine R if found guilty, and they have a profit loss $\Pi_M - \Pi_N$ when the AA forces them to interrupt the collusive behaviour for a certain (one) period.

If an investigation is opened, there is no incentive to deviate from the revelation policy agreed upon in a CR equilibrium: by not revealing when the other firms are expected to cooperate with the AA, the (deviating) firm would get the full fine F instead of the reduced fine R, and would break the cartel, with further future losses. Hence, to establish the conditions for a CR equilibrium the relevant constraint is that the firm cannot be better off by deviating (in the market) from the beginning. In this case the value of the deviating strategy is

$$V_D = \Pi_D + \delta\frac{\Pi_N}{1-\delta} \tag{4}$$

where $\Pi_D > \Pi_M$ are the profits in the deviation phase, which is followed by Nash punishment forever. The following Lemma states the conditions for a CR subgame perfect equilibrium to exist.

Lemma 1 *For given policy parameters (F, R, α, p), a subgame perfect equilibrium exists in which firms collude and reveal when monitored if*

$$\alpha < \alpha_{CR}(R) = \frac{\Pi_M - (1-\delta)\Pi_D - \delta\Pi_N}{\delta(\Pi_M - \Pi_N + R)} \tag{5}$$

Proof: The condition immediately follows from the inequality $V_{CR} > V_D$. ∎

Notice that $\alpha_{CR} \geq 0$ for $\delta \geq (\Pi_D - \Pi_M)/(\Pi_D - \Pi_N)$ which is the usual critical discount factor when firms collude with no threat of prosecution. For the remaining of the paper we focus on the interesting case when this

minimal condition holds. It is easy to show that $\alpha_{CR}(R)$ is decreasing in R and $\alpha_{CR}(0) < 1$. Hence, granting more generous discounts increases the threshold value α_{CR} relaxing the constraint for a CR equilibrium and making collusion more attractive. In this sense, LP have a pro-collusive effect since they decrease the expected cost of anticompetitive behaviour. Notice that the probability of independent prosecution p plays no role in a CR equilibrium, since the evidence to prove the existence of the cartel is provided by the colluding firms themselves. Since firms stop collusion for one period and pay a reduced fine every time they are reviewed, we need a sufficiently low α in order to make firms better off colluding rather than deviating.

3.2 CNR: Collude and Not Reveal

We proceed as before, deriving first the value of the game in a CNR equilibrium. From the timing of the game we obtain:

$$V_{CNR} = \Pi_M + \delta \tilde{V}_{CNR} \tag{6}$$

and

$$\begin{aligned} \tilde{V}_{CNR} &= \alpha\{\Pi_M + \delta[p(\Pi_N - F) + (1-p)\Pi_M]\} + (1-\alpha)(1+\delta)\Pi_M + \\ &\quad + \delta^2 \tilde{V}_{CNR} \\ &= \frac{\Pi_M}{1-\delta} - \frac{\alpha\delta p(\Pi_M - \Pi_N + F)}{1-\delta^2} \end{aligned} \tag{7}$$

The first term gives the profits if a review is opened, the firms continue to collude and do not cooperate with the AA and are condemned after one period with probability p. The second term corresponds to the stream of profits if the firms are not investigated in the period: since some other industry is being reviewed and then prosecuted, the firms have safe collusive profits for two periods. After two periods, the game restarts. Substituting in the previous expression we get:

$$V_{CNR} = \frac{\Pi_M}{1-\delta} - \frac{\alpha\delta^2 p(\Pi_M - \Pi_N + F)}{1-\delta^2} \tag{8}$$

In order to establish if CNR is a subgame perfect equilibrium we have to consider two constraints: *i*) the firms prefer to collude and not reveal rather than deviate (in the market) from the beginning; *ii*) when reviewed by the

AA, they prefer not to reveal rather than cooperate in the investigation. This latter condition is trivially satisfied if $R = F$, i.e. if no LP is introduced, because by deviating (revealing) firms pay the full fine for sure and lose the future collusive profits. When, instead, reduced fines are granted, this second constraint must be accurately considered.

The conditions for the first constraint to hold are stated in the following Lemma. As just claimed, if $R = F$, this condition establishes the existence of a CNR equilibrium.

Lemma 2 *For given policy parameters (F, R, α, p), when $R = F$ a CNR equilibrium exists if*

$$\alpha < \alpha_{NC}(p) = \frac{(1+\delta)(\Pi_M - (1-\delta)\Pi_D - \delta\Pi_N)}{\delta^2 p(\Pi_M - \Pi_N + F)} \tag{9}$$

Proof: Setting $V_{CNR} > V_D$ and rearranging we obtain the condition that guarantees that firms prefer to collude and not reveal rather than to deviate. Since it is never optimal to reveal when $R = F$, a CNR exists. ∎

The curve $\alpha_{NC}(p)$, shown in figure 1.a, is decreasing in p: a higher probability p of being condemned must be balanced by a lower probability α of being reviewed, in order to maintain the firms indifferent between CNR and deviate. The curve $\alpha_{NC}(p)$ is defined for $p \geq p_{NC}$, where $p_{NC} > 0$ is defined by $\alpha_{NC}(p_{NC}) = 1$ and is:

$$p_{NC} = \frac{(1+\delta)(\Pi_M - (1-\delta)\Pi_D - \delta\Pi_N)}{\delta^2(\Pi_M - \Pi_N + F)} \tag{10}$$

Consider now the second constraint for a CNR subgame perfect equilibrium, which becomes relevant when LP are introduced: for a CNR equilibrium to exist we want that in the subgame starting when the AA opens an investigation, the firms prefer not to reveal. The value of the game if the firm reveals once monitored, deviating from the prescriptions of a CNR equilibrium is:

$$V_R \mid \alpha = \frac{\Pi_N}{1-\delta} - R \tag{11}$$

If instead the firm does not reveal, according to the equilibrium strategy, the value of the game from that point on is:

$$V_{NR} \mid \alpha = \Pi_M + \delta[p(\Pi_N - F) + (1-p)\Pi_M] + \delta^2 \tilde{V}_{CNR} \tag{12}$$

Substituting the expression for $\tilde{V}_{CNR}$ and rearranging we obtain:

$$V_{NR} \mid \alpha = \frac{\Pi_M}{1-\delta} - \frac{\delta p(1-\delta^2(1-\alpha))(\Pi_M - \Pi_N + F)}{1-\delta^2} \tag{13}$$

Not revealing is optimal once monitored if the conditions stated in the following Lemma are satisfied. Coupled with the previous condition, we can establish the existence of a CNR equilibrium when $R < F$.

Lemma 3 *For given policy parameters (F, R, α, p), when $R < F$ a CNR equilibrium exists if $\alpha < \min\{\alpha_{NC}(p), \alpha_R(p)\}$, where*

$$\alpha_R(p) = \frac{(1+\delta)\{\Pi_M - \Pi_N + R(1-\delta) - \delta p(1-\delta)(\Pi_M - \Pi_N + F)\}}{\delta^3 p(\Pi_M - \Pi_N + F)} \tag{14}$$

Proof: See Appendix B. ∎

The locus $\alpha_R(p)$, shown in figure1.a, corresponds to the policy combinations that make firms indifferent between revealing and not revealing once a review is opened, and $\alpha < \alpha_R(p)$ guarantees that firms do not cooperate with the AA in the subgames starting with the opening of a review. This locus is decreasing in p: a higher probability of independent prosecution makes not revealing a less appealing choice once the firms are monitored; to keep firms indifferent between revealing and not revealing, we need a lower probability α of being reviewed, which increases the continuation value of the equilibrium strategy CNR.

3.3 CNR vs. CR

Comparing the conditions for the existence of a CR ($\alpha \leq \alpha_{CR}$) and a CNR ($\alpha \leq \min\{\alpha_{NC}(p), \alpha_R(p)\}$) subgame perfect equilibrium when $R < F$, we can notice that there are regions of parameters that admit both types of equilibria. We assume that in this case the firms are able to coordinate on the equilibrium that gives higher payoffs. The conditions for selecting the more profitable equilibrium are stated in the following Lemma.

Lemma 4 *For given policy parameters (F, R, α, p), the CNR subgame perfect equilibrium dominates the CR equilibrium if $\alpha < \alpha_{NC}(p)$ and*

$$p < p_{CNR}(R) = \frac{(1+\delta)(\Pi_M - \Pi_N + R)}{\delta(\Pi_M - \Pi_N + F)} \tag{15}$$

Proof: See Appendix B. ∎

A CNR equilibrium is preferred to a CR one if the probability of being condemned is sufficiently low: since a higher p reduces the value of the CNR equilibrium while it does not affect the value of the CR equilibrium, there exists a threshold level p_{CNR} over which firms prefer to reveal information to the AA and obtain reduced fines for sure. Notice that $p = p_{CNR}$ can be rewritten as:

$$p\frac{\delta(\Pi_M - \Pi_N + F)}{1+\delta} = \Pi_M - \Pi_N + R \tag{16}$$

Then, for $p = p_{CNR}$, the expected average losses suffered with probability p in a CNR equilibrium (the left hand side term) equal the average losses suffered with certainty in the CR case (the right hand side expression).

We have therefore concluded the analysis of the subgame perfect equilibria in the repeated game played by the firms. Figure 1.a shows in the space (α, p) for a given value of $R < F$ the curves that identify the constraints in the different equilibria. The following proposition summarizes the results.

Proposition 1 *In the repeated game played by the firms from $t = 1$ on, once the policy parameters (F, R, α, p) are set, there exist three subgame perfect equilibria:*

- *For $p \in [0, p_{CNR})$ and $\alpha \in [0, \min\{1, \alpha_{NC}(p))$ there exists a (Pareto dominant) subgame perfect equilibrium in which firms collude and do not reveal (CNR).*
- *For $p \in [p_{NC}, 1]$ and $\alpha \in [\max\{\alpha_{NC}(p), \alpha_{CR}\}, 1]$ there exists a unique subgame perfect equilibrium in which firms do not collude (NC).*
- *For $p \in [p_{CNR}, 1]$ and $\alpha \in [0, \alpha_{CR})$ there exists a (Pareto dominant) subgame perfect equilibrium in which firms collude and reveal if monitored (CR).*

Figure 1.b describes the three regions of parameters [22] in the space (α, p) for a given value of $R < F$. When both α and p are high (the north-east

[22] It should be remarked that the case of perfectly symmetric cartels that we are considering implies that the curves $\alpha_{NC}(p), p_{CNR}(R)$ and $\alpha_{CR}(R)$ which identify the equilibrium outcomes are the same for any industry and group of firms; hence, every industry will react in the same way, choosing the same equilibrium outcome NC, CR or CNR. For an analysis of the heterogeneous cartels case in a slightly different setting see Motta and Polo (1999).

region) deterrence is very effective and the value of colluding is decreased, making deviation more attractive: as a result, no cartel arises (NC). When α is low but p is high (the south-east region) refusing to cooperate with the AA is not rewarding since firms will likely receive the full fine. Since deterrence is not very effective because α is low, firms prefer to collude and reveal (CR). Finally, for low p deterrence is not effective and firms prefer to collude; but now they prefer not to reveal information (CNR), since the ability of the AA to condemn them to the full fine is very low.

Figures 1.a and 1.b about here

The lower R, the larger the region CR, which disappears when $R = F$: in this latter case we have only the NC and CNR regions, identified by the locus $\alpha_{NC}(p)$. We can compare the outcomes of the game played by the firms with and without Leniency Programs. There are two effects at work: under LP, there is a region of parameters (labelled 1 in the figure) which induces CR under Leniency Programs that would rather prevent collusion when no reduced fines are granted: since the expected cost of misbehaviour is lower, LP have a pro-collusive effect in this case. However, when collusion arises, the use of LP allows to obtain ex-post desistence more easily, by inducing revelation and by shortening the investigations, for certain values of the policy parameters labelled region 2. In order to establish which effect prevails, we have now to move to the implementation of the optimal policies.

Before studying the optimal policies, however, we would like to stress that fine discounts in our setting must be more generous than in Kaplow and Shavell (1994). They show that to induce violators to self-report it is enough to set the reduced fine $R \leq pF$. In our setting, firm would reveal if $p > p_{CNR}$, which can be rewritten as:

$$R \leq \frac{\delta}{1+\delta} pF - \frac{(1+\delta(1-p))}{1+\delta} (\Pi_M - \Pi_N). \tag{17}$$

There are two reasons why this condition differs from the findings of Kaplow and Shavell (1994). The first just lies in the formal specification of the model, and the fact that when firms do not reveal they can be fined only with a one period lag and every two periods (hence the term $\delta/(1+\delta)$ which multiplies pF). The second one is more important, and is due to the fact that by revealing the firm also foregoes with certainty some (appropriately weighted and discounted) future profits that, by not revealing, it would get with probability $(1-p)$ if not found guilty (hence the second term in the

inequality above). Therefore we should expect that whenever self-reporting diminishes the future profits of violators (which might occur not only in the case of interruption of a collusive practice, but also in cases where self-reporting increases the compliance of polluters, or reveals black assets to the tax authorities), fine discounts have to be more generous.

4 Optimal enforcement

In the previous section we have studied the firms' decisions given policy parameters. We now move to the endogenous choice of such parameters. The analysis of the optimal enforcement choices of the AA is built in two steps. We first consider the policy combinations α, p that the AA can implement given its enforcement technology, i.e. its budget constraint; then we derive the Iso-welfare curves in the (α, p) space where the equilibrium outcomes of the game among firms were identified; finally we discuss under which conditions leniency programs should be used.

4.1 Budget Constraint

In this section we specify the enforcement technology and derive the locus of implementable policies, that we define as the AA budget constraint. The AA is (exogenously) endowed with a sunk per-period budget; we assume that setting the fines at any level is not costly, while increasing the probability of enforcement requires resources. The budget allows to hire L officers, a richer budget implying a larger organization. The L officers are organized in teams of l units, which run both the review and prosecution stages, where l can be interpreted as the overall time spent by the members of a team on a single case during the period (equivalently, the crucial decision of the AA could be described in terms of the number of cases that are run in a period). The choice of l determines the probabilities of reviewing a cartel (α) and of proving firms guilty (p).

A team opens an investigation by selecting randomly an industry within a pool of N symmetric industries [23] which are potentially collusive. Given

[23] Since an industry can be investigated repeatedly, no matter whether it was previously found guilty or not, the number N of industries subject to review remains constant over time. Moreover, to obtain a stationary game structure, we assume that if a not colluding industry is reviewed, the AA closes the case in one period with no fine. This guarantees that if cartels start in any period after $t = 1$, the policy parameters are the same as in the case of cartels starting to operate at the beginning of the game.

the total labour force L, choosing the dimension of the teams determines the number[24] of cases n that can be treated in a period, $n = L/l$. In the case of symmetric cartels we are considering, all the industries choose the same equilibrium behaviour NC, CR or CNR: in this latter case the investigations last two periods. Hence, the AA opens n new cases each period (NC and CR) or every two periods (CNR). Let $l_N = L/N$ be the dimension of (time spent by) a team if all the N industries were reviewed at the same time: since the interesting case is when the AA has scarce resources, in our discussion l_N will be considered very small. The probability that an industry is reviewed is therefore

$$\alpha = \frac{n}{N} = \frac{l_N}{l} \tag{18}$$

The dimension of the team influences also the probability p of proving firms guilty when they do not cooperate. We assume that a minimum scale of the team (time spent) l_0 is needed in order to obtain some evidence, and that prosecution has decreasing returns, according to the function:

$$p = g(l - l_0) \tag{19}$$

with $g(0) = 0$, $\lim_{l\to\infty} g(\cdot) = 1$, $g'(\cdot) > 0$, and $g''(\cdot) < 0$ for $l \geq l_0$, and $\lim_{l\to l_0} g'(\cdot) = \infty$. A higher l_0 implies that, for given $l > l_0$ the probability p is lower, and corresponds to cases where the investigations are more complex (for instance because courts require higher standard of proofs) or the productivity of the teams is lower.

There is a trade-off in the enforcement policy between opening more reviews, which requires smaller teams, and being able to successfully conclude them, which is more likely if the teams are larger. To obtain the budget constraint, we can notice that $g(\cdot)$ is increasing and we can invert it, defining $f(p) = g^{-1}(p)$. Then $l = l_0 + f(p)$. Since $\alpha = l_N/l$, for $l > l_0$ the budget constraint of the AA, is

$$\alpha_{BC}(p) = \frac{l_N}{l_0 + f(p)} \tag{20}$$

for $p \leq g(L - l_0)$, which is the highest feasible probability, obtained if all the officers L are allocated to a single case. Figure 2 shows how the budget constraint can be constructed as a function of the dimension of the teams l. Once chosen $l \leq L$ (SW quadrant) both p (SE quadrant) and α (NW

[24] To ease the exposition and analysis we treat n as defined over the real line.

quadrant) are obtained, and the features of the α and $g(\cdot)$ curves determine the shape of the budget constraint $\alpha_{BC}(p)$ (NE quadrant). We can notice that it is downward sloping, initially concave and then convex. In Appendix A these properties are formally established. Moreover, since $l_N = L/N$, an increase in the budget L shifts up the vertical intercept of $\alpha_{BC}(p)$. A similar effect occurs if the minimum team size l_0 decreases.

The slope of the $\alpha_{BC}(p)$ curve is:

$$\frac{d\alpha_{BC}}{dp} = -\frac{l_N f'(p)}{(l_0 + f(p))^2} = -\frac{l_N}{l^2 g'(l - l_0)} = -\frac{\alpha}{p\epsilon_l} < 0 \tag{21}$$

where $\epsilon_l = g'(l - l_0)l/g(l - l_0)$ is the elasticity of p with respect to the dimension of the team l and depends also on the minimum team size l_0. For given l_0, let us define l_1 as the dimension of the team such that the elasticity is one, and let the corresponding probability be $p_1 = g(l_1 - l_0)$.

Figure 2 about here

In the Appendix it is proved that ϵ_l is greater (lower) than 1 if l is lower (higher) than l_1 and that both l_1 and p_1 are increasing in the minimum team size l_0. In figure 2 (SE quadrant), l_1 and p_1 are shown, and correspond to the point where the line from the origin to the g curve is tangent to the curve itself, i.e. where the marginal (g') and the average (g/l) probabilities are equal. The probability p_1 will play a crucial role in the analysis of the optimal enforcement.

4.2 Welfare gains

We can now derive a welfare measure of the antitrust activity. When cartels can arise in the market, antitrust intervention improves social welfare by preventing or interrupting collusion. We evaluate these effects through a utilitarian welfare function with equal weights on consumer and producer surplus. Moreover, the fines are assumed to be pure transfers that do not affect the aggregate welfare. The traditional deadweight loss (DWL) measures therefore the welfare gains associated with a successful intervention. We evaluate the welfare effects of antitrust enforcement by comparing the equilibrium outcomes NC, CR and CNR with the situation where collusion arises because no policy intervention is promoted.

As mentioned when describing the enforcement technology, the AA faces N industries which can potentially promote collusion. In a NC equilibrium,

no cartel arises and therefore the AA realizes the following welfare gains:

$$W_{NC} = N\frac{DWL}{1-\delta} = NK \tag{22}$$

where K is the present value of avoiding the cartelization of an industry. When the firms coordinate on a Collude and Reveal equilibrium, the AA opens in each period n reviews which induce revelation and end up with the n industries behaving non cooperatively for one period (with DWL gains), until the AA moves its attention to another industry. In this case the AA operates ex-post, obtaining desistence for a certain (one) period. Hence, starting from the second period, the AA obtains $nDWL$ gains per period, with a total welfare gain

$$W_{CR} = \frac{\delta nDWL}{1-\delta} = \delta nK = \delta\alpha W_{NC} \tag{23}$$

Comparing the welfare gains under NC and CR, the latter is lower for two reasons: it interrupts cartels only with probability α in each period (and $\alpha \leq \alpha_{CR} < 1$ in the CR region), and it intervenes ex-post and not ex-ante (δ). Finally, in a Collude and Not Reveal equilibrium the AA interrupts n cartels for one period ($nDWL$) only with probability p, and taking two periods to conclude the procedure. The welfare gains are therefore

$$W_{CNR} = \frac{\delta^2 npDWL}{1-\delta^2} = \frac{\delta}{1+\delta}pW_{CR} = \frac{\delta^2\alpha pW_{NC}}{1+\delta} \tag{24}$$

Compare the welfare gains of antitrust intervention with and without LP for given α in a right (CR) and left (CNR) neighborhood of p_{CNR}: CNR gives a lower welfare gain because ex-post desistence occurs with a lower probability (p) once a case is opened, and because it takes more time ($\delta/(1+\delta)$) to reach a decision.

Consider now the iso-welfare curves in the (α, p) space associated with the three outcomes. We initially focus on the case in which no LP are used, i.e. $R = F$: no collusion (NC) and collude and not reveal (CNR) are the two possible outcomes. In all the region NC the welfare gains are the same, as they do not depend on α and p. In the CNR region, the welfare gains depend on both α and p and the indifference curves are downward sloping hyperboles. If the welfare gain is W the iso-welfare curve in the CNR region is

$$\overline{\alpha}_{CNR}(p) = \frac{(1+\delta)W}{\delta^2 NKp} \tag{25}$$

Hence, $\overline{\alpha}_{CNR}(p)$ is an equilateral hyperbole as $\alpha_{NC}(p)$ is. In fact, it is easy to verify that the two curves overlap for a welfare gain in the CNR region equal to

$$W = NK\frac{\Pi_M - (1-\delta)\Pi_D - \delta\Pi_N}{\Pi_M - \Pi_N + F} \tag{26}$$

This welfare gain is the highest that can be realized in a CNR equilibrium, and corresponds to the highest iso-welfare curve in the CNR region.

Consider now the iso-welfare curves when LP are introduced. Now three outcomes can occur: NC, CNR and CR. In the CR region, W_{CR} depends only on α, and therefore the iso-welfare curves are horizontal. For a certain welfare gain W, the iso-welfare curve in the CR region is flat at

$$\overline{\alpha}_{CR} = \frac{W}{\delta NK} \tag{27}$$

To find the (same) iso-welfare curve that passes through the CNR and CR regions when LP are used, the following argument applies: let us fix a welfare gain W; in the CNR region the iso-welfare curve corresponds to the $\overline{\alpha}_{CNR}(p)$ curve already discussed. Once entering the CR region, the iso-welfare curve jumps down at $\overline{\alpha}_{CR} = W/\delta NK$ and becomes horizontal. Notice that $\overline{\alpha}_{CR} = W/\delta NK < \overline{\alpha}_{CNR}(1) = (1+\delta)W/(\delta^2 NK)$, i.e. the flat portion $\overline{\alpha}_{CR}$ of the iso-welfare curve is below the value of the downward sloping portion when prolonged to $p = 1$, $\overline{\alpha}_{CNR}(1)$. In figure 3 the thick line represents the iso-welfare curve passing through the CNR and CR regions.

Figure 3 about here

4.3 Optimal policies

We can now identify the optimal policies given the iso-welfare curves, the budget constraint and the incentive compatibility constraints that identify the equilibrium outcomes in the game played by the firms for given policy parameters. We first characterize the optimal policy when the AA wants to implement one of the three outcomes NC, CNR and CR. Then we compare the implementable outcomes and select the best one.

As a general point in all the equilibrium outcomes, it is always optimal to set $F = \overline{F}$ since increasing the fines is not costly and allows to obtain more favourable (lower) boundaries $\alpha_{NC}(p)$ and p_{CNR}.

Proposition 2 *The optimal policies that implement the NC, CNR and CR outcomes are:*

- *If* $\alpha_{BC}(p) \geq \alpha_{NC}(p)$ *for some* $p \in [p_{NC}, 1]$ *the NC outcome can be implemented. The optimal policy picks up the tangency point along the* $\alpha_{NC}(p)$ *curve where* $p = p_1$*, i.e. where the elasticity of the probability function* $p = g(\cdot)$ *is 1. If no tangency point exists, the optimal policy entails the corner solution* $p = p_{NC}$ *and* $\alpha = 1$*.*

- *The optimal policy to implement CNR sets* $R = F$ *and picks up the tangency point between the iso-welfare curve* $\overline{\alpha}_{CNR}(p)$ *and the budget constraint* $\alpha_{BC}(p)$ *at the point* $p = p_1$*, i.e. where the elasticity of the probability function* $p = g(\cdot)$ *is 1. If no tangency point exists, the optimal policy entails the corner solution* $p = g(l_N - l_0)$ *and* $\alpha = 1$ *or* $p = g(L - l_0)$ *and* $\alpha = l_N/(l_0 + f(p))$*.*

- *If* $g(L - l_0) \geq p_{CNR}(0)$*, a CR outcome can be implemented. The optimal policy sets* $R = 0$*,* $p = p_{CNR}(0)$ *and* $\alpha = l_N/(l_0 + f(p))$*.*

Proof: See Appendix B. ∎

Proposition 2 describes the optimal combination of policy parameters (α, p) which implement each of the three possible sub-game perfect equilibrium outcomes. This amounts to finding, within each equilibrium region, the highest welfare curve subject to a given budget constraint. For regions NC and CNR the optimal point will be given by the tangency point between the budget curve and the iso-welfare curves (or by a corner solution, as stated in Proposition 2).

Figure 4 shows how the equilibria NC, CR and CNR can be optimally implemented for different budget constraints. Given a budget α^1_{BC}, the best policy combination that implements NC is described by point E^1, i.e. at the tangency point of the budget constraint and the $\alpha_{NC}(p)$ curve corresponding to $p = p_1$. Given a budget α^2_{BC}, the highest iso-welfare curve attainable in the region CR is reached at the corner solution E^2 with $p = p_{CNR}(0)$: Above this point and to the right the budget constraint is not satisfied; below this point welfare is lower; to the left, the equilibrium CR does not exist. Finally, the CNR outcome is optimally implemented for budget α^3_{BC} at the tangency point with the iso-welfare curve $\overline{\alpha}_{CNR}(p)$, again for $p = p_1$.

Figure 4 about here

Note that - although the welfare gains in a CR equilibrium depend only on α - the optimal policy has to satisfy the constraint $p = p_{CNR}$ to be credible. If the AA, relying on firms' cooperation, chose to open a very large number of reviews (to reach a very high α) by organizing very small teams, then it would be unable to successfully conclude any of the reviews with a probability p sufficiently high to induce revelation.

An important result obtained by Proposition 2 is also that the optimal leniency scheme calls for $R = 0$. To understand why this is the case, suppose that $R > 0$. By setting $p = p_{CNR}$ the AA makes firms indifferent between revealing or not. By giving a more generous discount $R = 0$ firms strictly prefer to reveal, and the AA can still induce revelation reducing p. In turn, this frees up some resources of the AA, which can organize smaller teams and open more reviews, still obtaining firms' revelation.

We have found the optimal policies needed to implement the three different outcomes NC, CNR and CR. Our last step is to identify the conditions for selecting the outcome associated to the highest welfare gain. The following Proposition states the result.

Proposition 3 *If $\alpha_{BC}(p) \geq \alpha_{NC}(p)$ for some $p \in [p_{NC}, 1]$ the optimal policy entails implementing the NC outcome. If $\alpha_{BC}(p) < \alpha_{NC}(p)$ for any $p \in [p_{NC}, 1]$ and if $\alpha_{BC}(p_{CNR}(0)) \geq \delta p_1 \alpha_{BC}(p_1)/(1+\delta)$, the optimal policy implements a CR outcome. If $\alpha_{BC}(p) < \alpha_{NC}(p)$ for any $p \in [p_{NC}, 1]$ and if $\alpha_{BC}(p_{CNR}(0)) < \delta p_1 \alpha_{BC}(p_1)/(1+\delta)$, or if $g(L-l_0) < p_{CNR}(0)$, the optimal policy implements a CNR outcome.*

Proof: See Appendix B. ∎

Figure 5 about here

Proposition 3 identifies the conditions under which the Antitrust Authority selects the outcome NC, CNR or CR using the optimal policies analyzed in Proposition 2. If the budget constraint is sufficiently high to implement the NC outcome, this is the optimal policy. For intermediate budgets, the CNR and CR outcomes are implementable and we have to compare the welfare gains obtained in the two cases: the condition $\alpha_{BC}(p_{CNR}(0)) \geq \delta p_1 \alpha_{BC}(p_1)/(1+\delta)$ ensures that the CR outcome can be implemented with higher welfare gains than the CNR outcome. Finally, when this condition fails to hold, or for very low budgets, the CNR outcome is the best policy.

Let us discuss briefly the case of intermediate budget levels. In figure 5 we have an example in which $\alpha_{BC}(p_{CNR}(0)) \geq \delta p_1 \alpha_{BC}(p_1)/(1+\delta)$, which makes CR the preferred outcome (recall that we have already proved the best leniency policy calls for $R = 0$). Point A, where the budget constraint and the iso-welfare curve in the CNR region are tangent at $p = p_1$, is the best CNR outcome implementable. The associated welfare gain is $W^m = NK\delta^2 p_1 \alpha_{BC}(p_1)/(1+\delta)$. The same welfare level is obtained, in the CR region, along the flat iso-welfare curve, setting $\alpha^m = W^m/NK\delta = \delta p_1 \alpha_{BC}(p_1)/(1+\delta)$. But we see that the budget constraint enters in the CR region at $\alpha^M = \alpha_{BC}(p_{CNR}(0))$ (point B) which is higher than α^m: welfare gains W^M at point B (in region CR) are therefore higher than W^m at point A (in region CNR).

It is now easier to see how the budget of the AA determines the optimal policy. Note first that different levels of the budget L shift up or down the $\alpha_{BC}(p)$ curve but leave unchanged the probability p_1 which is chosen in both the NC and CNR outcomes. Let us start from a rich budget L and a large minimum team size l_0 such that the NC outcome can be implemented. In this case the optimal policy selects the NC outcome and sets $p = p_1$ and $\alpha = \alpha_{NC}(p_1)$ as shown in figure 4. Now let us decrease the budget L so that only CR and CNR are implementable, and let the condition $\alpha_{BC}(p_{CNR}(0)) \geq \delta p_1 \alpha_{BC}(p_1)/(1+\delta)$ be satisfied: CR is now preferred to CNR (refer again to figure 4). Since p_1 does not change as the budget L continues to shrink, CR continues to be the optimal outcome of the policy until the budget is so poor that p_{CNR} cannot be obtained even working with a single team, i.e. until $g(L-l_0) < p_{CNR}(0)$. At this point we revert to a CNR outcome, which is implemented in the corner solution $p = g(L-l_0)$, i.e. processing one case in each period. Hence, with a sufficiently high minimum size of the team l_0, decreasing budgets L induce the sequence NC-CR-CNR of policy outcomes.

This sequence continues to hold as long as the minimum size l_0 is high enough. When l_0 is very low, p_1 gets close to p_{NC} and we cannot exclude that NC is implemented close to (or at) the corner solution and that, once the budget line goes below the $\alpha_{NC}(p)$ curve, the condition $\alpha_{BC}(p_{CNR}(0)) \geq \delta p_1 \alpha_{BC}(p_1)/(1+\delta)$ fails to hold, inducing a sequence NC-CNR.

We conclude that, when the budget L is in an intermediate range (that can be very wide) and when the minimum team size l_0 is non negligible, the optimal policy entails the use of LP and implements a CR outcome.

5 Fine reductions only before the inquiry is opened

As mentioned in the introduction, the initial Leniency Program introduced in the US in 1978 entitled firms with a reduction in fines only if the cooperation started before an inquiry was opened. Similarly, the regime chosen in the EU with the July 1996 Notice applies mainly to firms which reveal information before the AA has opened an official investigation. It is therefore interesting to analyze whether not granting LPs to firms which report after an investigation starts can be justified in terms of enforcement effectiveness. We will show that this is not the case.

In a "fine reduction only before the inquiry is opened" regime, the AA initially sets the policy parameters and the LP eligibility rules. Then firms decide to collude or to deviate; in the following period, before an inquiry has been opened, the firms choose whether to reveal the existence of the cartel to the AA; if no firm reveals the AA reviews the cartel with probability α and concludes the prosecution stage in the next period condemning the firms with probability p. Once a case is concluded the game restarts.

In this setting we are interested to check whether both a collude and not reveal (CNR) and a collude and reveal (CR) subgame perfect equilibrium exist. The following proposition shows that only the former one exists.

Proposition 4 *Under a "reduced fines only before an inquiry is opened" regime, a CR subgame perfect equilibrium does not exist. If $\alpha < \alpha_{NC}(p)$ a CNR equilibrium exists. Otherwise, a NC equilibrium exists.*

Proof: See Appendix B. ∎

This proposition says that under this regime LPs are not effective, since firms will never reveal after having colluded, and the condition which determines whether firms do not collude or collude (and do not reveal) is the same as when LPs do not exist. Everything is as if a fine reduction was not in place. It is not difficult to understand why. Consider the benchmark case where firms were entitled to fine discounts *after* the opening of an investigation. There the expected profit from collusion decreases when the event "opening of an investigation" realizes, since, the probability of being condemned to the full fines jumps up from αp to p. Firms may react, if the rules of the LP allow, revealing information in exchange of reduced fines. Hence, it is the increase in the probability of being fined that triggers, in the benchmark case, firms to reveal after they have started to collude.

If instead collaboration with the AA is rewarded only for reporting before an investigation starts, firms have no incentive to reveal information to the AA after the investigation starts. But they have no incentive to reveal before the investigation starts either. Before observing if a review is opened, the probability of being condemned is still αp, the same as when deciding on colluding or not. If αp is high enough, firms will abstain from collusion; but if it is low enough, they will collude and not reveal. If nothing new happens between the moment firms decide on collusion and the moment they are asked to report to the AA, they will have no incentive to defect.

This is not a new result.[25] Malik and Schwab (1991, pp.30-31) noticed that the introduction of a tax amnesty alone (without an increase in compliance effort, for instance) will not lead a taxpayer to report additional income: "He has already chosen the optimal level of evasion, and if he has not received any additional information, then this optimal level of evasion remains unchanged".[26]

This result is consistent with the US experience, where initially the LP was used only for firms which spontaneously offered evidence before the inquiry was opened. In this initial regime the program was quite ineffective while, once allowed in 1993 for reduced fines even after the inquiry was opened, the number of cases in which the firms cooperated with the judges increased significantly.[27]

6 Conclusions

In this paper we have analyzed the effects of Leniency Programs on the incentives of firms to collude and to reveal information that helps the An-

[25] We also found the same result in Motta and Polo (1999), where the setting of the game was slightly different. See also Spagnolo (2000).

[26] Of course, in the real world the perception of the risk of being caught colluding might change overtime, or a more risk-adverse management might take over. This might explain why some firms might report to the AA even if the LPs apply only to revelations which occur before the investigation starts.

[27] In the 1994 Annual Report of the Antitrust Division it is stressed that in the first year of the new regime "an average of one corporation per month came forward with information on unilateral conspiracies, compared to an average of one per year under the previous policy. The policy thus allowed the Division to extend the reach of its criminal enforcement activities with relatively little expenditure of resources" (Antitrust Division (1994), p.6-7). More recently, "[a]mnesty applications over the past year have been coming in at the rate of approximately two per month" (Spratling (1999, page 2).

titrust Authority to prove illegal behaviour. We have showed that, by reducing the expected fines, Leniency Programs may induce a pro-collusive reaction: Combinations of policy parameters which, without Leniency Programs, would prevent collusion, may induce firms to collude (and reveal if monitored) when fine reductions are given. Hence, if the resources available to the AA are sufficient to prevent collusion using full fines, Leniency Programs should not be used.

However, when the AA has limited resources, Leniency Programs may be optimal in a second best perspective. Fine reductions, inducing firms to reveal information once an investigation is opened, increase the probability of ex-post desistence and save resources of the AA, thereby raising welfare. The optimal scheme requires maximum fine reduction, that is, the firms that collaborate with the Authority should not pay any fine.

We have then showed that allowing fine reductions only to firms which report to the Antitrust Authority *before* an inquiry is opened (as initially established in the US policy in 1978, and in the spirit of the EU Notice on the non-imposition of fines), is inferior to a regime where firms are entitled to fine discounts even if they reveal *after* an inquiry is opened.

We believe that, despite the simple setting, our paper sheds some light on the desirable features of actual Leniency Programs. In particular, our analysis indicates that a Leniency Program should be *equally* applicable (and generous) to information disclosed before and after an investigation has started. The US experience, where after the 1993 policy revision a corporation is granted leniency after an investigation has begun, shows that the extension of the Leniency Program to post-investigation amnesty is a crucial ingredient for success.

References

[1] Andreoni, James (1991) "The Desirability of a Permanent Tax Amnesty", *Journal of Public Economics* 45, pp. 143-159.

[2] Antitrust Division, (1994) *Annual Report* for Fiscal Year 1994.

[3] Arlen, Jennifer and Reinier Kraakman, (1997) , "Controlling Corporate Misconduct: An Analysis of Corporate Liability Regimes", *New York University Law Review*, 72, 4, 687-779.

[4] Das-Gupta, Arindam and Mookherjee, Dilip (1996) "Tax Amnesties as Asset-Laundering Devices", *The Journal of Law, Economics & Organization*, Vol. 12, No. 2, pp. 408-431.

[5] Easterbrook, Frank H., Landes, William M. and Posner, Richard A. (1980) "Contribution and Claim Reduction Among Antitrust Defendants: A Legal and Economic Analysis", *The Journal of Law and Economics*, pp. 331-370.

[6] European Union (1996) Notice on the Non-Imposition or Reduction of Fines in Cartel Cases, *Official Journal*, v.207, p.4.

[7] Friedman, James (1971) "A Non-Cooperative Equilibrium for Supergames", *Review of Economic Studies*, Vol. 38 (113), pp. 1-12.

[8] Grossman, Gene M. and Katz, Michael L.(1983) "Plea Bargaining and Social Welfare", *American Economic Review*, Vol. 73 (4), pp. 749-57.

[9] Innes, Robert (1999) "Remediation and Self-Reporting in optimal Law Enforcement", *Journal of Public Economics*, Vol. 72 (3), pp. 379-93.

[10] Kaplow, Louis and Shavell, Steven (1994) "Optimal Law Enforcement with Self-Reporting of Behavior", *Journal of Political Economy*, Vol. 102, No. 3, pp. 583-606.

[11] Kobayashi, Bruce H. (1992) "Deterrence with Multiple Defendants: An Explanation to Unfair Plea Bargains.", *Rand Journal of Economics*, Vol. 23 (4), pp. 507-17.

[12] Kornhauser, Lewis A. and Revesz, Richard L. (1994) "Multidefendant Settlements under Joint and Several Liability: The Problem of Insolvency", *Journal of Legal Studies*, Vol. XXIII, pp. 517-542.

[13] Landes, William M. (1971), "An Economic Analysis of the Courts", *Journal of Law and Economics*, 14, 61-108.

[14] Livernois, John and McKenna, C.J. (1999) "Truth or Consequences: Enforcing Pollution Satndards with Self-reporting", *Journal of Public Economics* 71, pp. 415-440.

[15] Malik, Arun S. (1993) "Self-Reporting and the Design of Policies for Regulating Stochastic Pollution" *Journal of Environmental Economics and Management*, Vol. 24, pp. 241-257.

[16] Malik, Arun S. and Schwab, Robert M. (1991) "The Economics of Tax Amnesties.", *Journal of Public Economics*, Vol. 46, pp. 29-49.

[17] Motta, Massimo and Polo, Michele (1999), "Leniency programs and Cartel Prosecution." Working Paper ECO No. 99/23, European University Institute.

[18] Polinski, Mitchell A. and Shavell, Steven (1981) "Contribution and Claim Reduction Among Antitrust Defendants: An Economic Analysis", *Stanford Law Review*, Vol. 33, pp. 447-471.

[19] Reinganum, Jennifer F. (1998) " Plea Bargaining and Prosecutorial Discretion", *American Economic Review*, Vol. 78 (4), pp. 713-28

[20] Rey, Patrick (2000) "Towards a Theory of Competition Policy", mimeo. (Available at http://www.univ-tlse1.fr/idei/Commun/Articles/Rey/seattle1026.pdf)

[21] Schrag, Joel and Scotchmer, Suzanne (1997) "The Self-Enforcing Nature of Crime", *International Review of Law and Economics*, Vol. 17, pp. 325-335.

[22] Spagnolo, Giancarlo (2000) "Optimal Leniency Programs", Stockholm School of Economics, Mimeo.

[23] Spratling G.R. (1999), The corporate leniency policy: answers to recurring questions, speech of the Deputy Assistant Attorney General, presented at the Bar Association of the District of Columbia, February 16, 1999.

[24] Tirole, Jean (1988) "*The Theory of Industrial Organization*", Cambridge, Mass. and London: MIT Press.

[25] Tokar, Steven (2000) "Whistleblowing and Corporate Crime", European University Institute, Mimeo.

Appendix A: The budget constraint

In this Appendix we formally prove the main properties of the budget constraint discussed in section 4.1. Since $\alpha = l_N/(l_0 + f(p))$, in the set $(\alpha, p) \in [0,1]^2$ the budget constraints intersects the boundaries at the following points: when $l_0 > l_N$ we have a vertical intercept for $p = 0$ at

$\alpha_{BC}(0) = l_N/l_0$; when $l_0 \leq l_N$ the budget constraint starts at $\alpha = 1$ and $p = g(l_N - l_0)$. The budget constraint is downward sloping and the second derivative of $\alpha_{BC}(p)$ is:

$$\frac{d^2\alpha_{BC}}{dp^2} = l_N \frac{2[f'(p)]^2 - f''(p)[l_0 + f(p)]}{[l_0 + f(p)]^3} \tag{28}$$

Since $lim_{p\to 0} f'(p) = 0$ and $f'' > 0$, for low values of p the curve is concave, while it becomes convex when p becomes large, as shown in figure 2.

Finally in the following Lemma the properties of l_1, the team size such that the elasticity $\epsilon_l = 1$, and of the associated probability $p_1 = g(l_1 - l_0)$ are presented.

Lemma 5 *$\epsilon_l \geq 1$ for $l \leq l_1(l_0)$ and $\epsilon_l < 1$ for any finite $l > l_1(l_0)$. $l_1(l_0)$ is increasing in l_0 with $l_1(0) = 0$ and $l_1(l_0) > l_0$. $p_1 = g(l_1(l_0) - l_0)$ is increasing in l_0.*

Proof: The easiest way to establish the result is by using the argument that the elasticity of the $g(\cdot)$ function corresponds to the ratio of the marginal (g') over the average (g/l) probability. Then $\epsilon_l = 1$ if $g' = g/l$. Being $g(\cdot)$ concave, for any finite $l > l_1$ the marginal probability is lower than the average probability and $\epsilon_l < 1$, the opposite occurring for $l < l_0$. For $l_0 = 0$, the marginal and average probability are equal at the origin, i.e. $l_1(0) = 0$. For any $l_0 > 0$, $\partial\epsilon_l/\partial l_0 = lg'(g' - g'')/g > 0$, i.e. starting from l_1 an increase in l_0 makes $\epsilon_l > 1$. Since $\partial\epsilon_l/\partial l = lg''/g + (1 - \epsilon_l)g'/g$, which is certainly negative for $\epsilon_l > 1$, to restore $\epsilon_l = 1$ we need an increase in l, i.e. l_1 is increasing in l_0. Since $l_1 = g/g'$ by definition, an increase in l_1 requires an increase in g/g', which can occur only if $l_1 - l_0$ increases, i.e. if l_1 increases more than l_0. We conclude that $g(l_1 - l_0) = p_1$ will increase as well when l_0 increases. ∎

Appendix B: Proofs

Proof: [Lemma 3]

When LPs are introduced, we have to check that firms prefer to collude and not reveal rather than to deviate, and that they prefer not to reveal once monitored rather than to cooperate with the AA. The first constraint corresponds to the condition $\alpha < \alpha_{NC}(p)$ in Lemma 2. The condition V_{NR} |

$\alpha > V_R \mid \alpha$ guarantees that not revealing is optimal after a review is opened, and can be rewritten as

$$\alpha\delta^3 p(\Pi_M - \Pi_N + F) \leq (1+\delta)\{\Pi_M - \Pi_N + R(1-\delta) - \delta p(1-\delta)(\Pi_M - \Pi_N + F)\} \tag{29}$$

where the term in curly brackets is non negative for

$$p \leq p_0 = \frac{\Pi_M - \Pi_N + R(1-\delta)}{\delta(1-\delta)(\Pi_M - \Pi_N + F)} \tag{30}$$

Hence, for $p > p_0$ the condition is never satisfied while for $p \leq p_0$ the condition holds for $\alpha < \alpha_R(p)$ as defined in the statement. Then a CNR equilibrium exists if both constraints hold. ∎

Proof: [Lemma 4]

The condition $\alpha < \alpha_{NC}(p)$ ensures that in a CNR equilibrium the firms prefer to collude and not reveal rather than to deviate. The condition $V_{CNR} > V_{CR}$, once, rearranged, gives $p < p_{CNR}(R)$.

Now let us establish the relations between $\alpha_{NC}(p), \alpha_{CR}$ and p_{CNR}. Substituting the latter in the first, we obtain $\alpha_{NC}(p_{CNR}) = \alpha_{CR}$, which means that the three curves intersect in the same point (α_{CR}, p_{CNR}), as shown in figure 1.a. Our last step is to show that the constraint $\alpha < \alpha_R(p)$ which is needed in a CNR equilibrium always holds if $\alpha < \alpha_{NC}(p)$ and $p < p_{CNR}$. Setting $\alpha_R(p) = 1$ and solving for p we obtain $p = p_0(1-\delta) > p_{NC}$, the latter being the value of p that solves $\alpha_{NC}(p) = 1$. That means that the curves $\alpha_{NC}(p)$ and $\alpha_R(p)$ are both decreasing and that, for $\alpha = 1$, $\alpha_{NC}(p)$ is to the left of $\alpha_R(p)$, see figure 1.a. Equating α_{CR} and $\alpha_R(p)$ and solving for p we obtain

$$p_R = \frac{(1+\delta)[\Pi_M - \Pi_N + R][\Pi_M - \Pi_N + R(1-\delta)]}{\delta[\Pi_M - \Pi_N + F][\Pi_M - \Pi_N - \delta(1-\delta)(\Pi_D - \Pi_N) + R(1-\delta)]} \tag{31}$$

Then $p_R > p_{CNR}$ since

$$\frac{\Pi_M - \Pi_N + R(1-\delta)}{\Pi_M - \Pi_N - \delta(1-\delta)(\Pi_D - \Pi_N) + R(1-\delta)} > 1 \tag{32}$$

Since $p_{NC} < p_0(1-\delta)$ and $p_{CNR} < p_R$, the curve $\alpha_R(p)$ is to the right of $\alpha_{NC}(p)$ for $\alpha \geq \alpha_{CR}$ as shown in Figure 1.a. Moreover, since $\alpha_R(p)$ is decreasing, it is to the right of the vertical locus $p = p_{CNR}$ for $\alpha < \alpha_{CR}$..

Hence, the constraint $\alpha < \alpha_R(p)$ is always satisfied in the region $\alpha < \alpha_{NC}(p)$ and $p \in [0, p_{CNR})$. ∎

Proof: [Proposition 2]

We characterize the optimal policies to implement NC, CNR and CR. Any point above the $\alpha_{NC}(p)$ curve is equivalent in terms of welfare gains, allowing to completely deter collusion. We select a point on the boundary of the NC region, i.e. along the curve $\alpha_{NC}(p)$ which allows to save resources, i.e. to implement NC with the minimum budget. Suppose that a tangency point exists between $\alpha_{NC}(p)$ and $\alpha_{BC}(p)$. The slope of the $\alpha_{NC}(p)$ curve is

$$\frac{\partial \alpha_{NC}}{\partial p} = -\frac{(1+\delta)(\Pi_M - (1-\delta)\Pi_D - \delta\Pi_N}{\delta^2 p^2 (\Pi_M - \Pi_N + F)} = -\frac{\alpha}{p} \tag{33}$$

Since the slope of $\alpha_{BC}(p)$ is $-\alpha/p\epsilon_l$, the tangency occurs when $\epsilon_l = 1$, i.e. at p_1 (or l_1). Since $\epsilon_l < 1$ for $p > p_1$ ($l > l_1$) and $\epsilon_l > 1$ for $p < p_1$ ($l < l_1$), the budget constraint curve is below the $\alpha_{NC}(p)$ curve for any other p. Hence, $p = p_1$ and $\alpha = l_N/(l_0 + f(p_1))$ are the optimal policy. For very low values of l_0 it may be that $p_1 \leq p_{NC}$, this latter being the value of p such that $\alpha_{NC} = 1$. In this case[28] the optimal policy entails the corner solution $p = p_{NC}$ and $\alpha = 1$.

Consider now the implementation of a CNR equilibrium. Since we are not interested in inducing revelation, $R = F$: only the CNR and NC regions exist in this case. The slope of the iso-welfare curves, being all of them equilateral hyperboles, is $-\alpha/p$ while that of the budget constraint curve is $-\alpha/p\epsilon_l$. The tangency point occurs therefore at $p = p_1$, and the budget constraint is always below the iso-welfare curve for any other p. For very low values of l_0 it may be that $l_N/l_1 > 1$, i.e. for any feasible value of the budget constraint the elasticity of the $g(\cdot)$ function is less than 1, i.e. in the CNR region the budget constraint is always steeper than the iso-welfare curves. In this case we have a corner solution at $\alpha = 1$ and $p = g(l_N - l_0)$.. Finally, for low values of the budget L it may be that $g(L - l_0) < l_1$ ($\epsilon_l > 1$), which implies that the budget constraint, for all $p \leq g(L - l_0)$ is always flatter than the iso-welfare curves. In this case the corner solution entails $p = g(L - l_0)$ and $\alpha = l_N/(l + f(p))$.

Finally, let us consider the optimal implementation of a CR outcome. The conditions stated in the Proposition imply that the budget constraint

[28] The case $p_1 = 1$, which would induce a corner solution at $p = 1$, is not relevant as it would require a minimum team of infinite dimension.

passes through the CR region. By choosing R we determine how wide this area is, and by setting α and p along the budget constraint we select a point in the CR region. Since the iso-welfare curves are horizontal, we want to select a policy combination on the highest iso-welfare curve. Since the budget constraint is downward sloping, we want to choose the lowest p (and therefore the highest α) consistent with a CR outcome, i.e. we choose $p = p_{CNR}$. Since this boundary level decreases (and α therefore increases along the budget constraint) when we reduce R, we adopt the most generous LP, i.e. $R = 0$, obtaining the lowest probability $p_{CNR}(0)$. The CR outcome can be implemented as long as $p_{CNR}(0)$ is feasible along the budget constraint, i.e. if $g(L - l_0)$ is larger. ∎

Proof: [Proposition 3]

If the budget constraint is tangent or intersects the lower boundary of the NC region, the NC outcome can be implemented. Since it is associated to the highest welfare gains, this is the optimal policy. If the $\alpha_{NC}(p)$ curve is always above the budget constraint, we have to choose between two possible outcomes, CR and CNR. If a CNR outcome is implemented, we have to choose $p = p_1$ and $\alpha = \alpha_{BC}(p_1)$. The welfare gains are $W^m = NK\delta^2 p_1 \alpha_{BC}(p_1)/(1+\delta)$. The same level of welfare W^m can be obtained in a CR equilibrium setting $\alpha = W^m/NK\delta$. Solving for α we obtain

$$\alpha^m = \frac{\delta p_1 \alpha_{BC}(p_1)}{1+\delta} \tag{34}$$

Hence, the CR outcome induced by $(\alpha^m, p_{CNR}(0))$ is welfare equivalent to the CNR outcome $(p_1, \alpha_{BC}(p_1))$. To select one of the two outcomes, we need to check whether the budget constraint passing through $(p_1, \alpha_{BC}(p_1))$ allows to implement a CR outcome preferable to $(\alpha^m, p_{CNR}(0))$. If $\alpha_{BC}(p_{CNR}(0)) \geq \alpha^m$, by moving to the CR region along the budget constraint we can implement (at least) an equivalent CR outcome. This case is shown in Figure 5, where the welfare level $W^M > W^m$ can be attained in the CR region. If on the contrary the budget constraint passes below α^m at $p = p_{CNR}(0)$, the CNR outcome is preferred. Finally, the CNR outcome is selected if CR cannot be obtained, i.e. if the highest probability $g(L - l_0)$ is lower than $p_{CNR}(0)$. ∎

Proof: [Proposition 4]

To find the conditions under which not revealing is an equilibrium, we have to check two incentive constraints. The first requires that a firm prefers to collude and not reveal rather than deviate, and is the same as in the benchmark case analized in the first part of the paper. It amounts to the condition $\alpha \leq \alpha_{NC}(p)$. The second incentive constraint imposes that the firm prefers not to reveal rather than reveal at the beginning of period 2, before an inquiry is opened. Notice that if firms do not reveal and an inquiry is opened during the same period, the prosecution stage will occur in the next period with firms proved guilty with probability p; after two periods the game restarts. Then, the value of the game if firms do not reveal is

$$V_{NR} \mid \alpha = \Pi_M + \delta\alpha[p(\Pi_N - F) + (1-p)\Pi_M] + \delta(1-\alpha)\Pi_M + \delta^2 V_{NR} \mid \alpha = \tilde{V}_{CNR} \tag{35}$$

i.e. the value of the game corresponds to the usual value of the stationary component of the game in a CNR equilibrium. The value of the game if the firm reveals, breaking the agreement, is

$$V_R \mid \alpha = \frac{\Pi_N}{1-\delta} - R \tag{36}$$

If the first incentive holds, $V_{CNR} = \Pi_M + \delta\tilde{V}_{CNR} \geq V_D$, which can also be written as $\tilde{V}_{CNR} \geq (\Pi_D - \Pi_M/\delta + \Pi_N/(1-\delta)$. We can then write:

$$V_{NR} \mid \alpha = \tilde{V}_{CNR} \geq \frac{(\Pi_D - \Pi_M)}{\delta} + \frac{\Pi_N}{1-\delta} > \frac{\Pi_N}{1-\delta} - R = V_R \mid \alpha \tag{37}$$

Hence, when the first incentive constraint holds, also the second is satisfied and $\alpha \leq \alpha_{NC}(p)$ guarantees the existence of a subgame perfect equilibrium CNR. If this condition fails, a NC equilibrium exists.

We have now to consider if an equilibrium in which firms collude and reveal before being reviewed exists. It is easy to see that this equilibrium does not exist. Indeed, by colluding and revealing a firm would get $V_{cr} = \Pi_M + \delta(\Pi_N/(1-\delta) - R)$. In fact, once firms are condemned to the reduced fines after revealing, they try to reorganize the cartel but they immediately prefer to reveal again instead of risking to be reviewed (and condemned to the full fines), loosing the possibility of reduced fines. By deviating it would get V_D which is clearly higher than V_{cr} being $\Pi_D > \Pi_M$ and $R \geq 0$. Hence, no CR equilibrium would ever arise. ∎

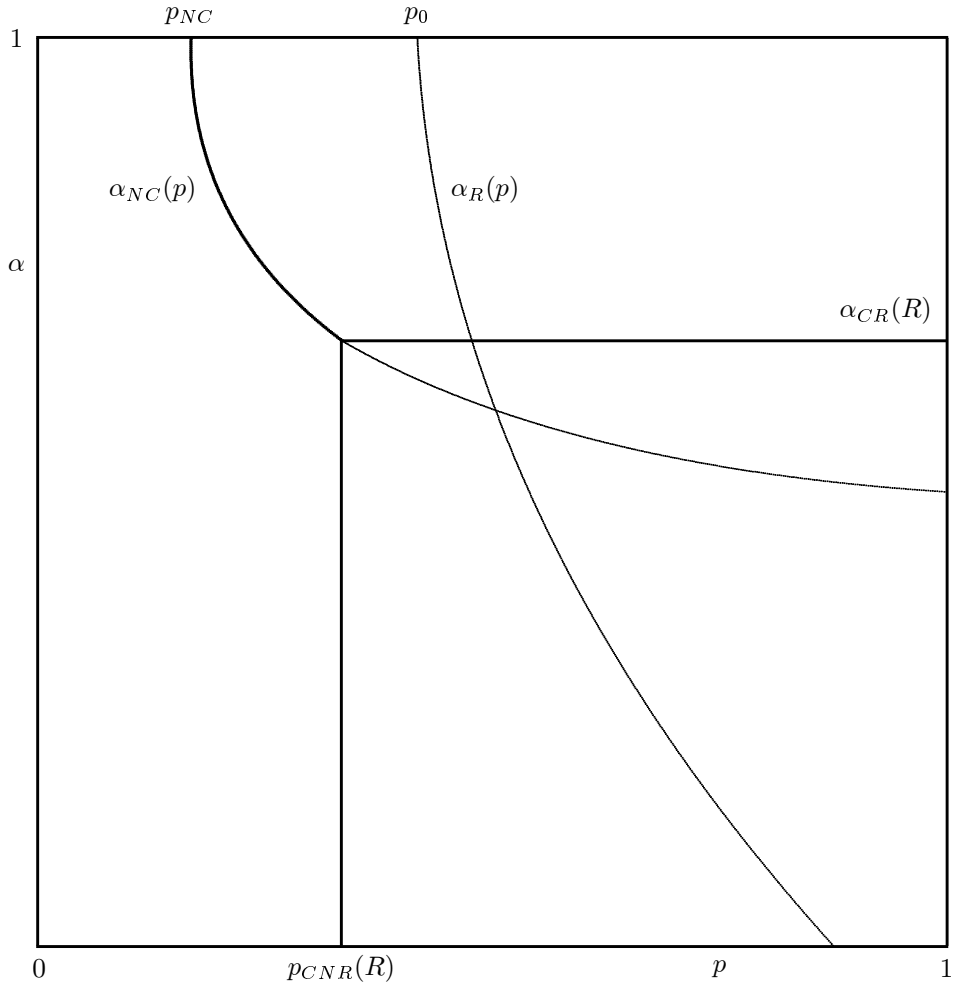

Figure 1.a: Incentive compatibility constraints

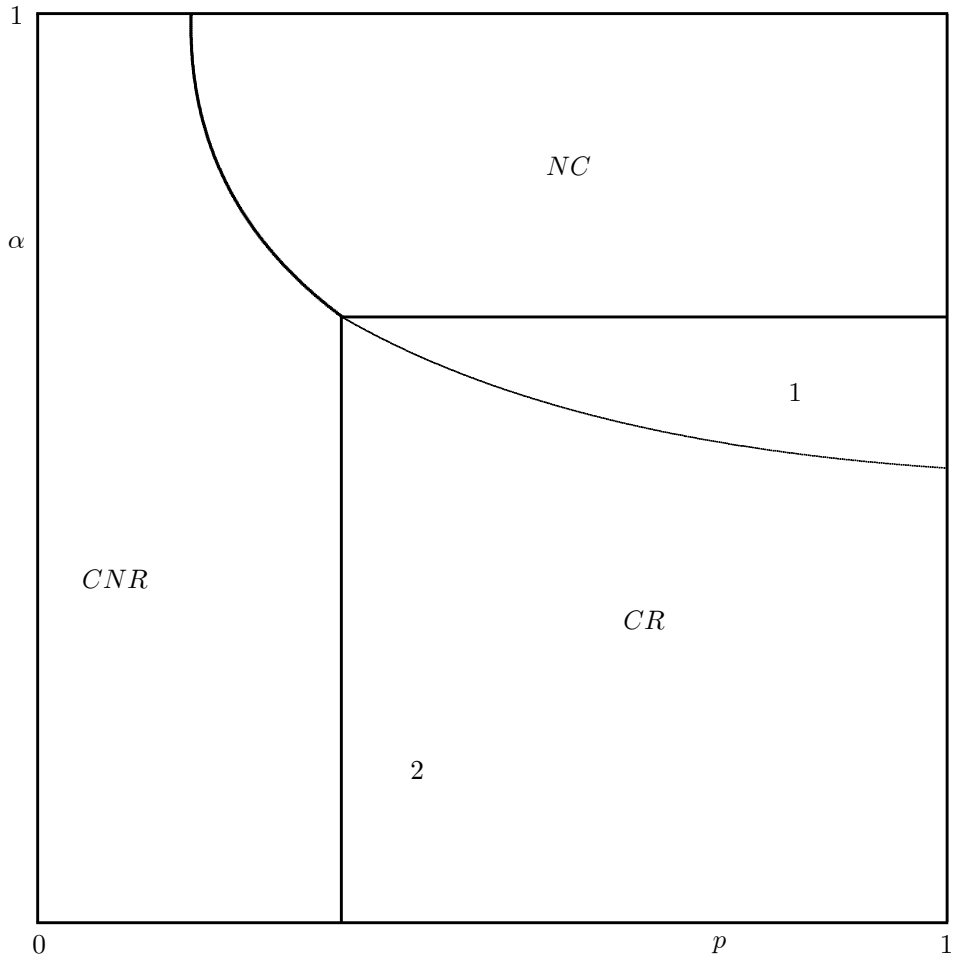

Figure 1.b: Subgame perfect equilibria

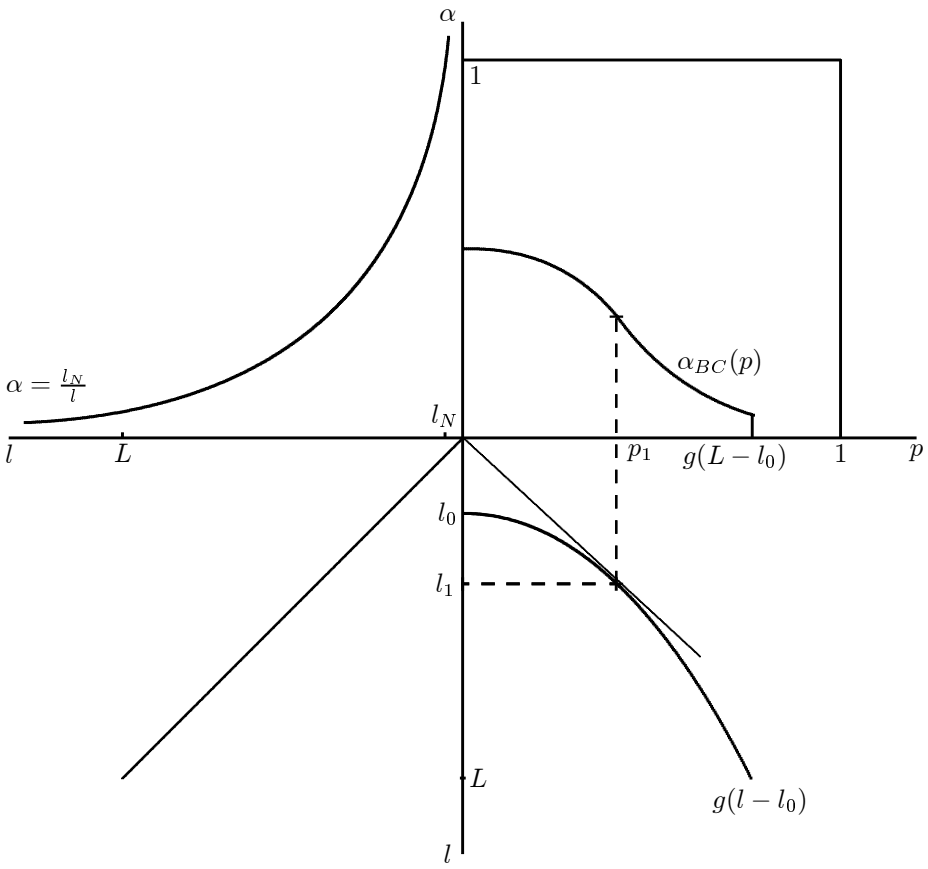

Figure 2: Deriving the Budget Constraint locus (α_{BC})

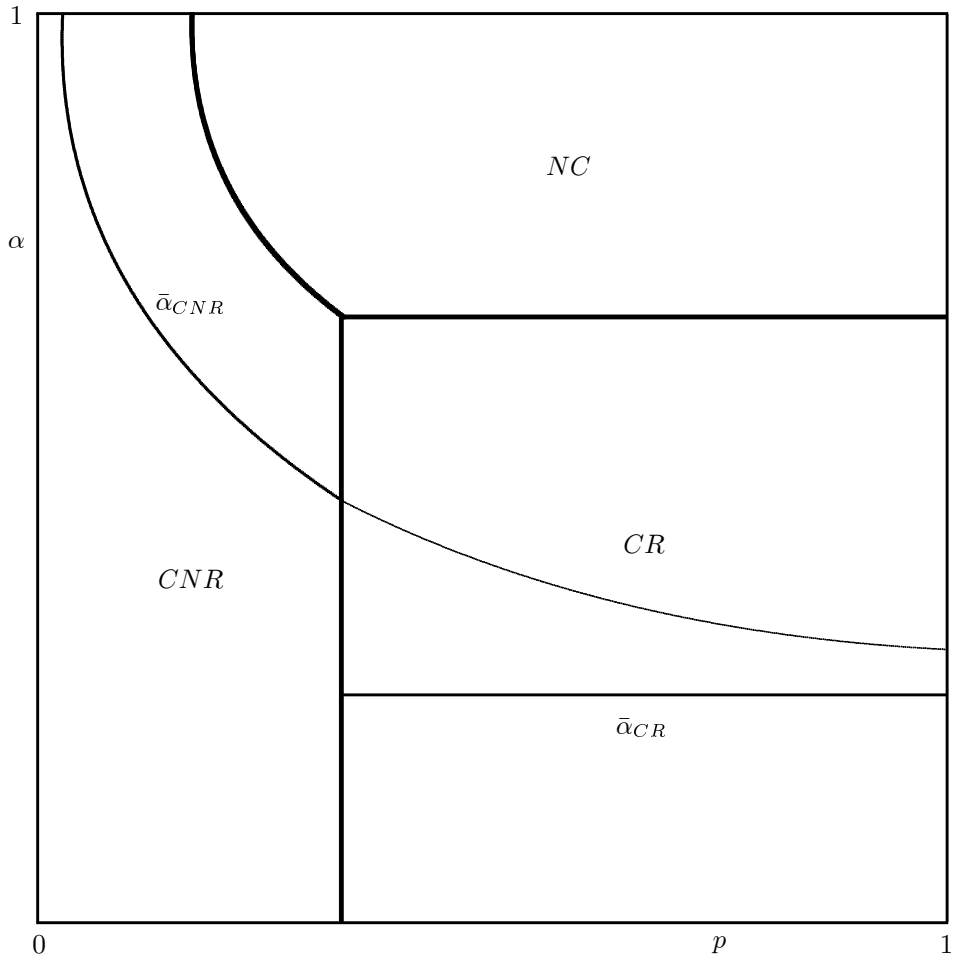

Figure 3: Isowelfare curves giving the same welfare level in the CNR and CR regions

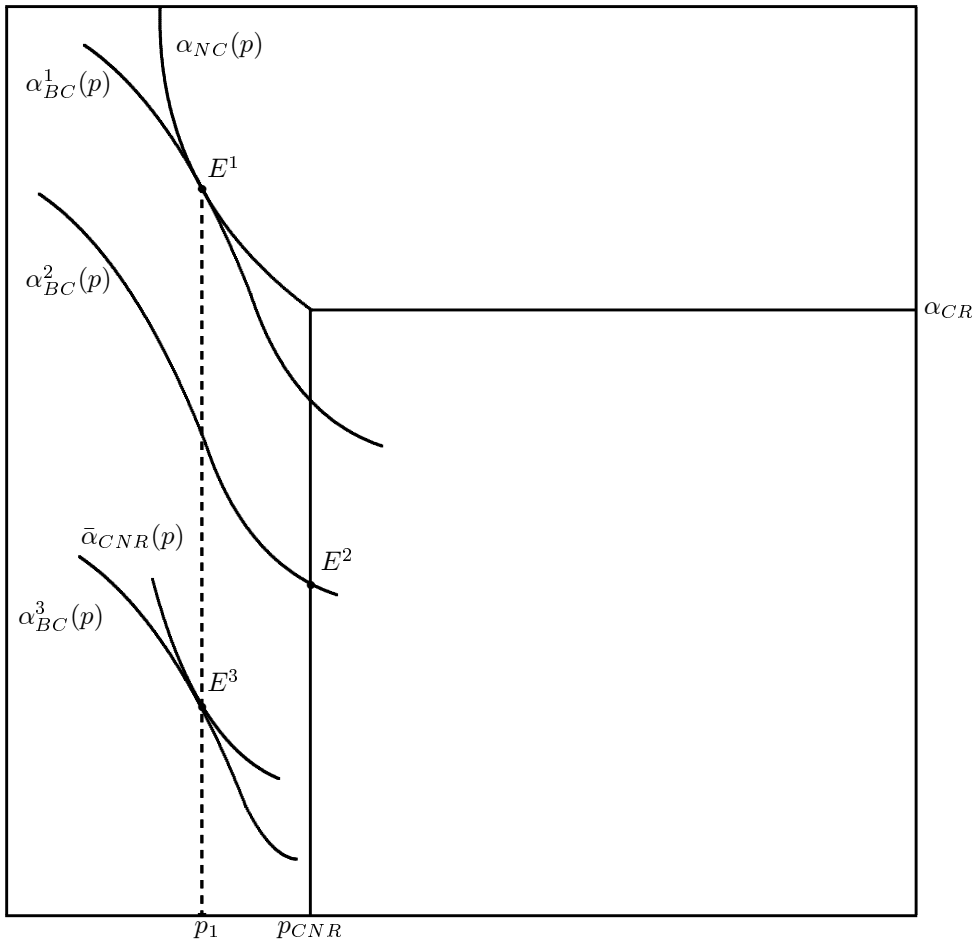

Figure 4: The optimal policies to implement NC (point E^1), CR (point E^2), and CNR (point E^3)

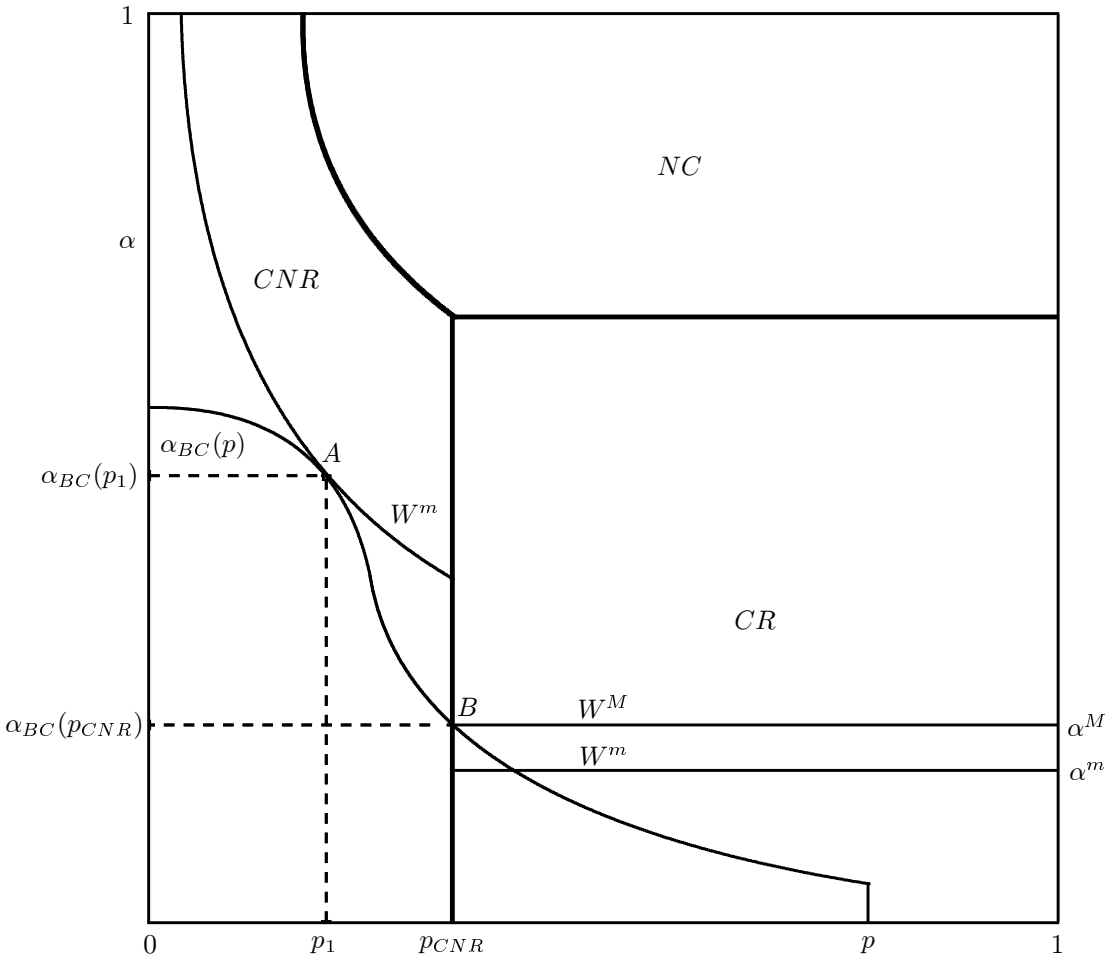

Figure 5: Welfare comparison of *CNR* (point *A*) and *CR* (point *B*) optimal policies when *CR* is preferred