Divide et Impera
Optimal Deterrence Mechanisms Against Cartels and Organized Crime*

Giancarlo Spagnolo

University of Mannheim
C.E.P.R.

Last revised: August 2003

Abstract

Leniency programs reduce sanctions for law violators that self-report. We focus on their ability to deter cartels and organized crime by increasing incentives to "cheat" on partners. Optimally designed "courageous" leniency programs *reward* the first party that reports with the fines paid by all other parties, and achieve the *first best*: complete and costless deterrence. "Moderate" leniency programs that only reduce or cancel sanctions may deter organized crime (a) by protecting an agent that defects from fines and from other agents' punishment; and (b) by increasing the *riskiness* of crime/collusion, in the sense of Harsanyi and Selten (1988).

JEL Classification: L13, L40, K21, K40

Keywords: Antitrust; Cartels; Collusion; Competition policy; Crime deterrence; Law enforcement; Leniency; Organized crime; Risk dominance; Self-reporting.

---

*This is a heavily revised and substantially extended version of my "Optimal Leniency Programs." I am grateful to Paolo Buccirossi, Tore Ellingsen, Philipp Festerling, Paul Heidhues, Massimo Motta, Paul Milgrom, Michele Polo, Andrea Prat and Patrick Rey for comments or stimulating discussions. Thanks also to participants at presentations in Stockholm, Mannheim, Lausanne (EARIE 2000; ESEM 2001), and Berlin (WZB-CEPR 2000 Conference on Antitrust Issues in International Markets). The financial support of the Swedish Competition Authority and the European Commission (Marie Curie Fellowship) is gratefully acknowledged. Correspondence to: *Department of Economics, Universität Mannheim, L7 3-5, D-68131 Mannheim, Germany.* E-mail: *gianca@uni-mannheim.de.*

The son said to him, "Father, I have sinned against heaven and against you. I am no longer worthy to be called your son". But the father said to his servants, "Quick! Bring the best robe and put it on him. Put a ring on his finger and sandals on his feet. Bring the fattened calf and kill it. Let's have a feast and celebrate." (Luke 15, 21-23)

This paper characterizes optimal law enforcement policies against cartels and organized crime, and evaluates their potential deterrence and welfare effects. The focus is on *leniency programs*, reduced sanctions for wrongdoers that report information on their partners' misbehavior to the law enforcing agency. These schemes attracted much attention in recent years thanks to the new *Corporate Leniency Policy* for Antitrust violations introduced by the US's Department of Justice (DoJ) in 1993.[1] This policy is widely regarded as a tremendous success. Since its introduction, an unprecedented number of cartels has been detected and successfully prosecuted, enormous fines (up to US$ 500 millions) have been levied against participants, and several top executives from different countries have served jail sentences in the US.[2] This celebrated success led Australia, Canada, the European Union, Germany, New Zealand, the UK and other countries to introduce analogous programs, and many more to discuss their possible introduction (OECD, 2001).

Although breaking down adversary coalitions by playing members against each other is a consolidated practice since Julius Cesar – who named it *Divide et Impera* – we cannot be sure that current leniency policies are the success they are claimed to be. The optimistic view that the increase in convicted cartels reflects an increase in cartel deterrence is plausible, but the actual change in active cartels caused by the Corporate Leniency Policy cannot be observed. In principle the observed increase in detected cartels could even be due to an increase in cartel activity. And even if current leniency programs do increase deterrence, we do not know whether differently designed ones would have done better. This calls for theoretical analysis.

The issue is not only relevant to Antitrust policy. As an illegal activity involving many agents, cartels can be considered a form of *organized crime*, certainly not the most harmful. Many more dangerous and far reaching forms of crime are organized and share crucial features with cartels: all of Mafias' and gangs' activities, terrorism, all forms of corruption (where at least two parties are involved, a briber and a bribee), all kinds of illegal trade (drugs, arms and people tafficking, where at least a buyer and a seller are involved), large frauds (including financial ones), and any other form of crime exercised at too large a scale for an isolated individual. In fact, any criminal activity above a certain scale must be organized. And needless to say, the costs of organized crime to society are enormous. Understanding the optimal design of law enforcement policies against organized crime, therefore, is of primary importance.

All forms of organized crime share with cartels three fundamental and intertwined features that distinguish them from ordinary crime, and that are crucial to the design of effective

[1] Toghether with the companion Individual Leniency Policy. The DoJ introduced a leniency policy for cartels already in 1978, but the old policy was much less generous than the new one. As a result, very few firms applied for leniency before 1993.

[2] See Spratling (1998, 1999) and Hammond (2000, 2001) for an overview.

law enforcement policies. The first feature is that cooperation among several agents is required to perform the criminal activity, so that free riding, "hold-up", and "moral hazard" issues become relevant. The second one is that organized crime takes the form of ongoing criminal relations: instead of the isolated criminal act with given benefit b and harm h familiar from the law enforcement literature (the next section presents a literature review), organized crime delivers *flows* of present and expected future benefits and damages. The third is that cooperating wrongdoers, by acting together, inevitably end up having information on each others' criminal behavior that could be reported to third parties. This third feature is clearly a consequence of the first, but also the second is. Criminal organizations suffer of an intrinsic "governance problem" since to curb moral hazard and ensure internal cooperation they cannot rely on explicit contracts enforced by the legal system. This is why, as cartels, all forms of organized crime *must* take the form of criminal *relations*. Then implicit/relational contracts can be sustained by the "carrot" of expected future gains from the criminal activity, together with the "stick" of the (often very harsh) punishments against cheaters.

Because organized crime requires cooperation between multiple wrongdoers and a certain degree of internal trust, one way prosecutors have traditionally fought it is by shaping private incentives to play one party against the other, by ensuring that they find themselves in a situation as close as possible to a Prisoner's Dilemma. This is the idea behind leniency programs.[3] These policies have been used more or less explicitly to fight most forms of organized crime. Most notably, they have been extensively and successfully used in the US and Italy to fight Sicilian Mafia, and are routinely used (and misused) in the US to fight drug-dealing and related crimes.[4]

From a theoretical viewpoint, the Prisoner's Dilemma game is perhaps the first and best known model of a leniency policy: the sanctions for a prisoner are reduced to induce him to confess and prove guilty his former partner. The Prisoner's Dilemma refers to a situation in which the joint law violators have already been detected, and the leniency program seeks to elicit additional information to facilitate prosecution. Leniency programs are also advocated and implemented as a way to directly *deter* organized crime, by inducing *undetected* wrongdoers to spontaneously self-report and "turn in" their partners. The idea is undermining trust between wrongdoers with the increased risk that one of them unilaterally reports to enjoy the benefits of the leniency program (which are typically restricted to the first reporting party). Indeed, a crucial new feature of the Corporate Leniency Policy is its "Amnesty Program" – Section A – that "automatically" awards full immunity from all sanctions to the first, and only the first cartel member that spontaneously reports information before an investigation is opened.[5] According to DoJ officials, it is this new feature that led

[3] To be precise, agents involved in organized crime are in a Prisoner's Dilemma-like situation already without the leniency program, since each of them can cheat on the others "running away with the money". But typically the situation is repeated, and criminal/collusive agreements can be sustained by reputational forces. What leniency programs do is changing the payoffs in this dynamic game, so that the choice between colluding and defecting-and-reporting again looks similar to a static Prisoner's Dilemma.

[4] The misuse occurs when prosecutors and courts rely exclusively (or mainly) upon a testimony obtained in exchange for leniency. A useful introduction to this incredible practice is at *http://www.pbs.org/wgbh/pages/frontline/shows/snitch/.* Throughout the paper we will assume that the party applying for leniency must report "hard information" against his partners to obtain it, and that his testimony is not required nor admitted.

[5] The amnesty typically includes all the firm's employees. In February 2002 the European Commission

many companies to come, and often rush forward with information on their cartel.[6] This paper focuses on this aspect of leniency programs, their ability to deter organized crime by undermining trust, increasing wrongdoers' incentives to spontaneously report information when they are not under any sort of investigation.

We study a stylized model of law enforcement, in the spirit of Gary Becker (1968), that includes the possibility to offer reduced or negative fines (rewards financed by fines imposed on convicted wrongdoers) to agents that report "hard" information on their criminal organization, and to impose higher fines against repeated/recidivous offenders. The other crucial departure from Becker's contribution is that our model is dynamic: we focus on cartels and organized crime, so that the isolated criminal act is replaced by a criminal relation, an equilibrium in the dynamic game between multiple wrongdoers.

First we determine the optimal law enforcement policy in the absence of leniency policies. Besides establishing a benchmark, this exercise delivers some interesting insights. It shows that absent leniency programs, law enforcing agencies should commit not to target agents that unilaterally defect from collusive/criminal strategies, and should make this policy public. The reason is close to the logic of leniency: if agents know that they will not be fined for their past crimes if they defect from the collusive/criminal agreement, they are more prone to do it, which makes such agreements harder to sustain.

We then analyze how leniency programs affect the collusive/criminal game. We find that when these programs are sufficiently generous, they can be exploited by colluding agents who may agree to report each period, enjoy leniency and avoid part or all fines. This increases the value of a collusive/criminal agreement and reduces the overall deterrence effect of the law enforcement policy. We also find that when these programs are sufficiently generous, they directly increase agents' incentive to unilaterally defect and report information. An agent that defects and reports will be sure not to be fined for past collusive behavior, and can cash the reward (if one is present). This increases the value of defecting and the overall deterrence effect of the law enforcement policy. Taking these two effects into account, we then characterize the optimal law enforcement policy with leniency.

We find that an optimal leniency policy is restricted to the first party that reports. Allowing more agents to obtain leniency makes the program more easily exploitable, since then fewer agents must pay the full fine each period; and it reduces the maximal reward that can be offered to the first agent that reports. The optimal policy also maximizes fines. High fines are valuable not only because they reduce the expected value of collusive-criminal relations, as in Becker (1968), but also because they allow to finance higher rewards for agents that self-report. Most importantly, we find that unless fines are exogenously constrained to be very small, *the optimal policy offers to the first reporting agent a reward*

revised its eight-year-old leniency policy exactly to introduce this feature: complete, "automatic" exemption from fines for the first firm that spontaneously self-report before an investigation is opened (European Commission, 2002). The previous version, introduced in 1996, left substantial discretion to the Commission, which may have scared potential applicants (no firm spontaneously self-reported under the old scheme).

[6] According to Scott Hammond, Director of Criminal Enforcement of the DoJ Antitrust Division, about 50% of the leniency applications are now spontaneous reports falling within Section A of the Corporate Leniency Policy (personal communication). Elsewhere, he claimed that "over the last five years, the Amnesty Program has been responsible for detecting and prosecuting more antitrust violation than all of our [other investigating tools]" (2001). Similar statements can be found in Spratling (1998, 1999).

equal to the sum of all fines paid by his former partners. Maximal rewards maximize the deterrence effect of the law enforcement policy.

Large fines can finance a large reward for the first reporting agent, and a sufficiently large reward would lead an agent to defect from any collusive/criminal agreement, report and cash the reward. Since agents know this, when fines and rewards are sufficiently high no agreement is sustainable and *the optimal policy implements the first best, complete and costless deterrence*. Investigations by the law enforcing agency are then redundant, and being costly, should be avoided. This is perhaps the most striking result of the paper.[7] The crucial ingredient behind it is the third among the earlier-mentioned distinctive features of organized crime: the fact that others posses information on an agent's crime. Leniency elicits the information agents have on their partners, ensuring that a wrongdoer may have to pay a fine even when the probability of being directly detected by the law enforcement agency is zero. We name "courageous" these optimal, "high powered" leniency policies offering rewards to the first reporting agent, as rewarding former wrongdoers is sometimes regarded as immoral, even though the Bible suggests the very opposite.

Since political and institutional constraints may prevent offering rewards, we go on to analyze constrained-optimal "moderate" leniency programs, where reduced fines are bounded to be non-negative (rewards are excluded). Moderate leniency programs are "low powered" incentive schemes, and as such they cannot achieve the first best. However, they are not irrelevant with respect to deterrence, as it has sometimes been argued (even by who writes). We identify two effects ensuring that even moderate leniency programs restricted to the first, spontaneously reporting party may make collusive/criminal agreements harder to sustain. The first is a "protection from fines" effect, and is present when the reduced fines of the moderate leniency program are below the expected fine of a defecting agent that does not report. By increasing the expected payoff of an agent that defects and reports above that of an agent that only defects, the moderate leniency program make collusive/criminal agreements harder to sustain. The second is a "protection from punishment" effect, present when collusive/criminal agreements are sustained by two-phase punishment strategies and repeated offenders are punished harder than first time ones. A report then raises fines and reduces expected profits from further collusion, limiting the costs agents are willing to incur to punish the agent that defected in the first place.

There is at least a third important reason why even moderate leniency programs may have deterrence effects. As often stressed by DoJ officials, leniency may generate "breakdowns in trust" among wrongdoers, it may increase the perceived "riskiness" of collusion. To capture this effect, we introduce *risk dominance* considerations in the sense of John Harsanyi and Reinhadt Selten (1988). Within a simplified version of the model, we show that even moderate leniency programs always strictly increase the riskiness of collusive/criminal agreements. Moreover, we find that riskiness increases strictly more when eligibility to the program is restricted to the first reporting party, offering theoretical support to DoJ officials' claim that the first comer rule is crucial in generating breakdowns of trust in cartels, and the consequent rushes to report. Optimal policies with respect to maximizing the riskiness of collusive/criminal agreements are otherwise identical to those making agreements harder

[7] To our knowledge, it is the first time that the first best is achieved in a model of law enforcement á la Becker.

to sustain in equilibrium: they prescribe maximal fines, and a maximal reward for the first reporting agent.

The remainder of the paper is organized as follows. We start with a short review of the literature. Section 2 describes the model. Section 3 derives the optimal law enforcement policy in the absence of leniency programs. Section 4 analyzes how leniency programs affect sustainability of collusive/criminal agreements in equilibrium. Section 5 characterizes the optimal law enforcement policy when rewards are feasible. Section 6 considers constrained-optimal policies where leniency programs must be “moderate”. Section 7 characterizes the effects of leniency when agents care about strategic risk. Section 8 discusses robustness issues and concludes. All proofs are in the Appendix.

1 Literature review

Despite the prominence of the Prisoner's Dilemma game in economics and the importance of organized crime in society, until very recently there was no economic investigation of the effects of leniency programs on cartels and organized crime.

The first paper addressing the effects of leniency policy on cartels, to our knowledge, is one by Massimo Motta and Michele Polo (2003). Their approach is complementary to our. We look for the optimal design of law enforcement policies with leniency programs in a model of crime deterrence á la Becker (1968), where detection and conviction are identified with a single probability and the cost of enforcement is a choice variable for the policy maker. In this sense, we are mainly focusing on the optimal design and deterrence effects of Section A of the US Corporate Leniency Policy. Motta and Polo's model is instead in the spirit of the plea bargaining literature, with an exogenous cost of law enforcement and with detection leading to conviction only with some probability. Their model is designed to answer a specific question: whether firms that report information when being already under investigation should also be eligible to leniency. Their main focus, therefore, is on the value of Section B of the Corporate Leniency Programs, and their central result is that it may indeed increase deterrence by making prosecution more effective, a result on which we fully agree.

Besides in focus, our paper and Motta and Polo's differ crucially in both assumptions and results. These authors do not consider rewards nor strategic risk; they assume that firms sustain collusive agreements with ”grim trigger” strategies (we allow for generic punishment strategies), and that a defecting firm cannot be convicted for having taken part to a cartel, nor can report information on former partners (we consider both these possibilities). Under their simplifying assumptions – required to solve their more complex model – the “protection from fines” and the “protection from punishment” effects do not emerge, and since risk dominance issues are not considered a moderate leniency program appears unable to induce agents to spontaneously self-report. This leads to their conclusions that to have any effect a leniency program must be open to firms under investigation, that the same lenient treatment should be offered to all firms independent of who reports first (under their assumptions removing the “first comer rule” – the benefit of being the first firm to report – has no cost), and that leniency programs are second-best (i.e. if the Antitrust Authority has sufficient resources to deter cartels through fines and inspections it should not introduce any leniency program).

These conclusions conflict directly with our central result that an optimally designed leniency program can deliver the *first best*, and with our other (Sections 6 and 7) results that it is always optimal to have a leniency program, even if moderate, and that an optimal leniency program restricts eligibility to the best treatment (rewards or full amnesty) to the first firm that reports, as done in reality.[8] This contrast is due to the restrictive assumptions Motta and Polo worked with. As we show, if agents can be convicted for past crimes and can self-report when they defect, if they can use (optimal) two-phase punishment strategies, or if they are susceptible to strategic risk considerations, even moderate leniency programs have a direct deterrence effects which would be just lost by offering the same amount of leniency to all reporting parties.

Our work is also closely related to a paper by Cecile Aubert, William Kovacic and Patrick Rey (in progress). This paper was developed independently and simultaneously to our, and also considers rewards in Antitrust enforcement, although in a model that – as far as we know (the paper was not yet circulated) – shares most of Motta and Polo's assumptions. Aubert et al. focus on the interaction between the Individual and the Corporate Leniency Policies, i.e. on the costs and benefits of creating an internal agency problem, a conflict of interests between the firm and individual employees, by allowing the latter to cash a reward when reporting their own firm's misbehavior. They also consider the possibility of secret reports. Rey (2001) offers an excellent survey that discusses some of this paper's results.

Giancarlo Spagnolo (2000) highlighted a potential drawback of current (moderate) leniency programs. It shows that when information is durable, with moderate leniency programs the threat of reporting the cartel to the Antitrust Authority in case of a defection could be used to enforce collusion in occasional competitive situations like auctions. The optimal schemes discussed in the present paper, though, would also solve that kind of problems.

Cristopher Ellis and Wesley Wilson (2002) developed a model of the current leniency programs that offers a new perspective. They show that a moderate leniency program may induce firms to report information in order to damage competitors and obtain a strategic advantage. Their result, together with our, helps explain the rush of cartel breakdowns with spontaneous reports that has taken place in the US these last years. However, these authors also find that leniency programs may end up having the perverse effect of stabilizing those cartels that it could not deter, reinforcing the results in Spagnolo (2000).

Our paper builds on Becker (1968), therefore it is related to the economic literature on optimal law enforcement stemming from that seminal contribution.[9] Within this literature, our paper is closest to the recent work on self-reporting. Focusing on individual wrongdoers committing isolated crimes, Louis Kaplow and Steven Shavell (1994) have shown how reducing sanctions against wrongdoers that spontaneously self-report lowers law enforcement costs by reducing the number of wrongdoers to be detected. These authors also show that

[8] Motta and Polo's conclusions also conflict with the statements of DoJ officials arguing that treating better the first reporting agent is crucial to the success ofthe leniency program; and are somewhat inconsistent with the recent US experience, where 50% of leniency applications were spontaneous reports falling under Section A of the Corporate Leniency Policy (we believe, though, that only part of these applications were really spontaneous reports).

[9] Michael Polisky and Steven Shavell (2000) and Nuno Garoupa (1997) offer encompassing overviews of this literature.

when agents are risk averse, offering leniency to wrongdoers that self-report increases welfare by reducing the overall risk agents bear. Both these insights apply to leniency policies in general. Relatedly, Arun Malik (1993) discusses the role of self-reporting in reducing auditing costs in environmental regulation; while Innes (1999a,b) highlights the value of the early remediation of damages that fine reductions for self-reporting wrongdoers allow for. These papers highlight important benefits that a lenient treatment of self-reporting agents may bring about, but none of them considers its peculiar ability to elicit information and undermine trust in cartels and organized crime.

We are not the first to discuss organized crime from an economic point of view. Polo (1995) and Diego Gambetta and Peter Reuter (1995) already noted that criminal organizations suffer of internal enforcement problems; Kai Konrad and Stergios Skaperdas (1997, 1998) emphasized the role of reputational forces and credible threats in making contracts in the criminal world self-enforcing; and Garoupa (2000) emphasized the vertical-hierarchical aspect of criminal organizations and its effect on optimal law enforcement. None of these papers, though, discuss why and how leniency policies may be used to deter organized crime.

2 The model

Let there be an economy with many oligopolistic industries – or a society with many potential criminal organizations – each of which can be represented by a discounted infinitely repeated (oligopolistic or criminal) game between a number of risk neutral agents. Let there also be a benevolent Legislator who – having forbidden welfare–reducing collusive/criminal behavior – sets the parameters of the law enforcement policy. We assume that the Legislator sets and commits to law enforcement policy parameters first. Then, having observed those policy choices, agents interact in the oligopolistic (criminal) supergame.

TIMING

- Step 1: The Legislator commits to law enforcement policy parameters
- Step 2: Agents observe the policy parameters and start interacting

2.1 Cartels/organized crime

We will focus on industries where the exercise of collusive market power generates deadweight welfare losses that dominate any potential dynamic gain, and on criminal organizations that produce net social losses.[10] A representative industry (potential criminal organizations) i consists of $N \geq 2$ symmetric firms (agents) interacting repeatedly in the infinite, discrete time denoted by $t = 1, 2...$, and discounting future through the common factor δ_i, with $0 < \delta_i < 1$.

We assume there is a continuum of industries (potential criminal organizations) that differ only with respect to agents' discount factor δ_i, and that can be ordered one a line

[10]This means that if law enforcement was costless complete crime/cartel deterrence would be optimal. In classical models of optimal law enforcement, optimal deterrence may be partial even when law enforcement is costless (see e.g. Polinsky and Shavell, 2000). This is because cases are included where the benefit from a crime exceeds the harm it causes, so that from a purely utilitarian perspective the crime should not be deterred.

with uniform density with respect to the such factor $\delta_i \in [\delta_{\min}, \delta_{\max}]$. We also assume that agents in all industries (potential criminal organization) use the same strategies to support collusive agreements, and that the stage game – the static oligopolistic (criminal) interaction – has at least a pure strategy Nash equilibrium, with π^n denoting agents' lowest payoff from a static equilibrium.

Remark 1 *For expositional convenience, in the remainder of the paper we will phrase our discussion mainly in terms of a collusive agreement between oligopolistic firms. However, the reader should keep in mind that all reasoning and results directly apply to the other forms of organized crime discussed in the introduction.*[11]

In the absence of law enforcement, firms can sustain a stationary collusive agreement in subgame perfect Nash equilibrium if the value of sticking to the agreement – the discounted sum of expected payoffs from respecting the agreement V_i^c – exceeds the value V_i^d of defecting unilaterally and then being subject to the punishment phase that disciplines the agreement.[12] The correspondent algebraic condition for a representative firm in industry i, is

$$(V_i^c =)\ \frac{\pi^c}{1-\delta_i} > \pi^d + \delta_i V_i^p\ (= V_i^d),$$

or, normalized by $(1-\delta_i)$,

$$\pi^c > (1-\delta_i)\pi^d + \delta_i v^p, \tag{1}$$

where π^c denotes a firm's static payoff from sticking to the collusive agreement; π^d that from unilaterally deviating from the agreement and choosing the static best response (of course $\pi^d > \pi^c > \pi^n$); V_i^p denotes the discounted sum of payoffs expected at the beginning of the punishment phase following a firm's unilateral defection; and v^p denotes the time-average payoff a firm that defected unilaterally earns after the defection, so that $v^p = (1-\delta_i)V_i^p$ (the superscript p is for "punished").[13] Of course it must be $v^p < \pi^c < \pi^d$, since to enforce collusion cartel members must penalize defecting ones in one way or another. Finally, π^b will denote the payoff a colluding firm obtains when one of its partners defects unilaterally, where $\pi^b < \pi^n < \pi^c < \pi^d$.

We try keep as general as possible by not specifying particular punishment strategies. Specific punishment strategies will be discussed when they are important for a result. Since we want to understand the effects of leniency programs on deterrence, in these occasions

[11] To translate the results we obtain for cartels into correspondent results for other forms of organized crime, it is sufficient to reinterpret variables. For example, for corruption, the number of players N will typically be two, say a firm and a bureaucrat; collusive profits π^c can be reinterpreted as the gains from a complete collusive transaction; profits from a unilateral deviation π^d can be reinterpreted as a party's gains from cheating in the collusive transaction; and so on.

[12] We will focus on a given agreement that can be thought of as the most profitable one, such as the joint monopoly collusive agreement in case of oligopoly. An interesting extension of this paper would be to specify the underlying game and let agents choose among different levels of criminal cooperation, e.g. different degrees of collusion.

[13] The simbol V_i^x will always be used for the discounted sum of payoffs expected in industry i by an agent in phase x, and small v^x for the corresponding time-average per period payoff $v^x = (1-\delta_i)V_y^x$, which is independent from the discount factor.

we will assume that agents choose punishment strategies optimally: we will focus on collusive/criminal arrangements enforced by "optimal penal codes" which discipline defections with the strongest available punishment (Dilip Abreu, 1988). When these optimal arrangements are deterred, any other collusive/criminal agreement will also be deterred.

2.2 Information

We assume that each period a cartel is active a piece of "hard" evidence is generated, independent of whether a defection from collusive strategies occurs.[14] We can think to each stage game of the dynamic game as being composed of two substages: in the first cartel members communicate – e.g. to confirm/update collusive strategies – generating hard information; in the second they set the relevant market variables (price, output, investments...).[15]

For simplicity, we assume that all cartel members possess part of the hard evidence produced by the cartel and can costlessly transmit it to third parties if they wish; that if an agent reports the hard evidence it possesses to the law enforcing agency the cartel is convicted with probability one; and that there is "full information decay" in the sense that all hard information on a cartel active at time t vanishes at time $t+1$. It will become clear that the qualitative results of the paper do not depend on any of these simplifying assumptions.

Finally, we limit attention to the case of public information revelation by assuming that when a firm reports its hard information to the law enforcing agency, its report becomes public information at the end of the period (to obtain the conviction of a cartel prosecutors must usually disclose available information and its sources to courts and defendants).[16]

2.3 Law enforcement

The Legislator can set the following parameters of the law enforcement policy, within limits dependant on exogenous (e.g. political) factors:

1. A monetary fine F, with $F \in [\underline{F}, \overline{F}]$, that a colluding firm or a member of the criminal organization has to pay if convicted for the first time;

2. A fine F^r that a repeated (or recidivous) offender – a firm already convicted in the past – has to pay when convicted again for the same offense, with $F^r \geq F$, $F^r \in [\underline{F}^r, \overline{F}^r]$, and $\overline{F}^r \geq \overline{F}$;

3. A reduced fine RF ("reward" when $RF < 0$), with $RF \in [\underline{RF}, F]$, that a cartel member can pay/cash instead of F if – when it is not under investigation – it spontaneously reveals information to the law enforcement agency, allowing it to convict its partners;

4. The probability α by which cartel members are discovered and convicted in a period in which everybody conforms to agreed collusive strategies;

[14] A previous version of this paper assumed that evidence is produced only if no defection occurs, as in Motta and Polo (1999). The current assumption is more realistic, since in reality undercutting ones' cartel is no guarantee not be convicted for past collusive activities. In addition, this assumption allows us to model the effects of leniency programs on the "riskiness" of collusion.

[15] This timing for the stage game is suggested by Rey (2001, p. 17), and reflects well the behaviour of real world cartels, whose members meet regularly to monitor and update their agreement.

[16] The case of secret reports is considered in Aubert et al.(in progress) and Rey (2001).

5. The probability β by which a cartel member that conforms to agreed collusive strategies is convicted in a period when another member unilaterally defects;

6. The probability γ by which a unilaterally defecting cartel member is convicted; we assume this can happen only if non defecting members are also convicted, so that $\beta > \gamma$ (one could realistically assume $\beta = \gamma + \eta$, with $\eta \geq 0$).

We restrict focus on realistic parameter configurations by assuming $\alpha, \beta, \gamma < \frac{1}{2}$. In the tradition of Becker (1968), we also assume that administering fines is costless, so that these can be regarded as welfare-neutral transfers, but that increasing each of the conviction probabilities is costly. We let $c_k(k)$, with $k \in \{\alpha, \beta, \gamma\}$, denote the (social) cost functions of implementing such probabilities, and assume $c_k(0) = 0$, $c_k'(k) \geq 0$, $c_k''(.) > 0$ and, to simplify some statements, $c_k'(0) = 0$.

As for rewards (negative RF), were they financed through taxation it would of course be reasonable to assume them costly to administer. However, in this paper we will consider exclusively *self-financing* leniency programs, defined as follows.

Definition 1 *A leniency program is* self-financing *when the sum of rewards it pays to reporting agents (if any) is weakly smaller than the sum of the fines paid by other agents involved in the reported crime; i.e. when* $\sum RF \leq \sum F$ *for each cartel* i.

Therefore, as positive fines, rewards (negative fines) will also be assumed costless to administer.[17] We will see that focusing on self-financing leniency programs is not restrictive, since self-financing constraints (upper bounds on rewards) emerge endogenously in the search for globally optimal programs to avoid having agents colluding and reporting just to cash rewards. It will also become clear that the basic framework sketched in this section can be complicated in many ways without qualitatively affecting results.[18]

3 Optimal law enforcement without leniency

Consider a representative industry i when no leniency program is present, so that $RF = F$. Let V_i^{cr} denote the discounted payoff expected by a member of a convicted cartel who did not defect the period after being convicted, and v^{cr} the correspondent time-average payoff, with $V_i^{cr}(1-\delta_i) = v^{cr}$ (the superscript c is for "cooperator" and r for "recidivous"). Of course v^{cr} is a function of firms' strategies and of the legal system. By sticking to agreed collusive strategies each firm expects the profit stream

$$\alpha(\pi^c - F + \frac{\delta_i v^{cr}}{1-\delta_i}) + (1-\alpha)\pi^c + (1-\alpha)\delta_i \left[\alpha(\pi^c - F + \frac{\delta_i v^{cr}}{1-\delta_i}) + (1-\alpha)\pi^c\right] + (1-\alpha)^2\delta_i^2[...,$$

so that the value of the strategy "stick to collusive/criminal agreement" is

$$V_i^c = \frac{\pi^c + \alpha(\frac{\delta_i v^{cr}}{1-\delta_i} - F)}{1-(1-\alpha)\delta_i} = \frac{\pi^c - \alpha F + \alpha\delta_i V_i^{cr}}{1-(1-\alpha)\delta_i},$$

[17] Adding other (moral?) costs of rewarding a wrongdoer that reports would complicate exposition, reduce the set of parameter configurations where a reward is optimal, but otherwise leave all qualitative conclusions unchanged.

[18] For example, non-monetary and fit-the-crime sanctions can easily be accommodated by the model, without substantial changes in results.

which is decreasing in α, F and (weakly) in F^r.[19]

Discounted expected payoffs from defecting depend on the probability β by which the cartel is convicted a period in which a defection takes place, and on the probability γ that the defecting firm is itself convicted. The value of defecting from collusive strategies is

$$V_i^d = \pi^d - \gamma F + \frac{\delta_i(\beta v^{pr} + (1-\beta)v^p)}{1-\delta_i},$$

where v^{pr} denotes the time-average profit expected by a cartel member after he unilaterally defected and the cartel was convicted (with or without him; again the superscript r is for "repeated" or "recidivous" and $V_i^{pr}(1-\delta_i) = v^{pr}$). It can be $v^{pr} \neq v^p$ because cartel members never convicted will pay F if caught colluding again after the defection, while cartel members already convicted will pay F^r. (If $F^r = F$ it will obviously be $v^{pr} = v^p$). Similarly, if one assumes that collusive/criminal agreements are disciplined by "grim trigger" strategies (Friedman, 1971) – such that if a defection occurs all agents start playing the worse static Nash equilibrium forever – it is $v^{pr} = v^p = \pi^n$ and $\frac{\partial v^h}{\partial F^k} = 0$ for $h \in \{p, pr\}$ and $k \in \{r, 0\}$. However, if $F^r > F$ and agents use other (optimal) punishment strategies than grim trigger, expected gains from collusion will be lower for repeated offenders. For example, when the agreements are enforced by two-phase punishments à la Abreu (1986) – where all agents conform to a costly "stick" phase (e.g. a tough market war) because of the "carrot" of a subsequent return to collusion – by reducing expected gains from collusion (the carrot) a higher F^r softens the stick players can credibly threaten against a defector. In this case it will be $\frac{\partial v^{pr}}{\partial F^r} > 0$ and $v^{pr} > v^p$. A similar argument applies to the relation between F and v^p, so that when agents adopt such strategies it is $\frac{\partial v^p}{\partial F} > 0$.[20]

Remark 2 *Since it can be $\frac{\partial v^p}{\partial F} < 0$, $\frac{\partial v^{pr}}{\partial F^r} < 0$ and $v^{pr} < v^p$ only when players use suboptimal punishment strategies, in the remainder of the paper we will focus on the case $\frac{\partial v^p}{\partial F} \geq 0$, $\frac{\partial v^{pr}}{\partial F^r} \geq 0$ and $v^{pr} \geq v^p$.*

Wrapping up, firms can sustain a collusive/criminal agreement in subgame perfect Nash equilibrium if, for each player i

$$V_i^c = \frac{\pi_i^c + \alpha(\frac{\delta_i v^{cr}}{1-\delta_i} - F)}{1-\delta_i(1-\alpha)} > \pi_i^d - \gamma F + \frac{\delta_i(\beta v^{pr} + (1-\beta)v^p)}{1-\delta_i} = V_i^d. \quad (2)$$

Studying this condition leads to the first, benchmark results.

Lemma 1 *Suppose there is no leniency program ($RF = F$). Then:*

[19] The value of colluding is decreasing in F^r because stronger fines for recidive wrongdoers reduce payoffs v^{cr} expected after being caught ($\frac{\partial v^{cr}}{\partial F^r} \leq 0$). To see that it is decreasing in α one can differentiate to obtain

$$\frac{\partial V_i^c}{\partial \alpha} = \frac{\delta(v^{cr} - \pi^c) - (1-\delta)F}{[1-\delta(1-\alpha)]^2},$$

negative as long as $\delta(v^{cr} - \pi^c) < (1-\delta)F$, which is satisfied since $v^{cr} < \pi^c$.

[20] Analogous arguments apply when agents enforce collusive/criminal agreements by asymmetric (weakly) renegotiation-proof punishments of the kind discussed by Joe Farrell and Eric Maskin (1989).

1. *When F and F^r can be set independently, the ex ante optimal fines are:*
   i) $F^* = \overline{F}$ *if* $\gamma < \frac{\alpha}{1-\delta_i(1-\alpha)} + \frac{\delta_i(1-\beta)}{1-\delta_i}\frac{\partial v^p}{\partial F}$*, and* $F^* = 0$ *otherwise;*
   ii) $F^{r*} = \overline{F^r}$.
2. *When F and F^r are restricted to be equal, the ex ante optimal fine is:*
   $F^* = \overline{F}$ *if* $\gamma < \frac{\alpha}{1-\delta_i} + \frac{\delta_i}{1-\delta_i}\frac{\partial v^p}{\partial F^r}$*, and* $F^* = 0$ *otherwise.*

The Lemma appears complex because we have kept everything as general as possible, but its interpretation is quite straightforward. Statements 1) and 2) are the organized crime version of Becker's (1968) celebrated result that for individual, isolated crimes the Legislator should set fines at their maximum to save on investigation costs. As one would have expected, with dynamic multiagent criminal relations things are slightly more complex than in Becker's single agent, occasional crime framework, hence the Lemma offers additional insights.

Statement 1i) gives the condition under which raising fines for first-time offenders, *caeteris paribus*, deters (or facilitates) collusion between never-convicted offenders. It says that increasing fines F for first time offenders, holding fixed other parameters of the law enforcement policy (F^r included), increases deterrence as long as the probability α of being convicted while colluding is not substantially smaller than the probability γ of being convicted when defecting unilaterally from collusive strategies. In this case the fine should be set at its maximum.[21] However, if a unilateral defection increases the probability to be caught by the law enforcing agency ($\gamma > \alpha$), say because the defection signals the existence of a cartel, a Becker-like result does not obtain: then an increase in fines deters *defections* from a cartel, rather than the cartel itself. By deterring defections, an increase in fines ends up facilitating collusion. Statement 1ii) says that to maximize ex ante deterrence of first-time collusion it is optimal to set fines against repeated offenders maximal. Statement 2 shows that also when F and F^r are restricted to be equal again it is optimal to maximize fines only if γ is sufficiently low. Lemma 1 directly leads to the first, benchmark proposition.

Proposition 1 *Absent leniency programs (i.e. with $RF = F$) an optimal law enforcement policy sets $\gamma = 0$, $F = F^*$, and $F^r = F^{r*}$.*

Absent Leniency Programs, the Legislator maximizes welfare by setting $\gamma = 0$. This result is not directly related to Becker's work, it is an insight specific to our dynamic multiagent framework. Welfare increases when (2) is more stringent (more cartels/crime are deterred), and as long as there is law enforcement ($F > 0$) any increase in γ makes (2) less stringent by discouraging unilateral defections from collusive behavior. Since increasing γ is costly, the optimal policy is to set $\gamma = 0$. In practice, *the law enforcement agency should carefully avoid prosecuting cartel members who unilaterally defected from collusive strategies, and should make this policy of public domain.* A reputation for forgiving cartel members that unilaterally defect destabilizes cartels by encouraging further defections, while at the

[21] We believe this case, and in particular the case $\alpha > \gamma$, to be the most relevant in reality (also because, as we argue below, it is optimal for the Legislator to keep γ low). In the next sections we will focus mainly on this "regular" case, where Becker's result applies.

same time saving on law enforcement costs.[22] And when γ is optimally set at zero Becker's (1968) logic applies: by Lemma 1, fines should then be maximal.

To better focus on the effects of the Leniency Program, in the remainder of the paper we will restrict attention to the most realistic case where $\gamma < \alpha$ (so that maximal fines are optimal), and assume that the Legislator optimally sets $F = F^*$ and $F^r = F^{r*}$ (the optimality of this choice in this model is not affected by changes in RF).

4 The Effects of Leniency

We now let the Legislator introduce a leniency program by choosing a reduced fine $RF < F$ for colluding firms that report hard information sufficient to convict "the rest of the gang". The analysis will focus on the two most important dimensions of leniency programs: the eligibility criteria, i.e. whether all reporting firms or only the first one to report should obtain leniency; and the size of the reduced sanction/reward RF.

Absent Leniency Programs, reporting information was always a strictly dominated action, since it led to the payment of the full sanction F. With a Leniency Program agents may instead find it convenient to report information on the cartel, and this may change both the value of colluding V_i^c and that of defecting V_i^d. To characterize the optimal program we must understand how leniency affects V_i^c and V_i^d.

4.1 "Exploitable" leniency programs

Consider first the effects of a leniency program on the value of colluding. The value V_i^c may change because when the leniency program is sufficiently generous, colluding agents may find it convenient to consensually report their collusive/criminal behavior each period, to avoid facing the risk of being detected and fined by the law enforcing agency. When this is the case we will say that the leniency program is "exploitable", in the sense that agents can use it to reduce the expected cost of misbehavior. Let $V_i^{c\prime}$ denote the value of following the strategy profile that prescribes firms to both collude and report their collusive agreement in each period.

Definition 2 *A leniency program is "exploitable" in industry i when it allows firms to increase the value of a collusive/criminal agreement by reporting it each period to the law enforcing agency (when $V_i^{c\prime} > V_i^c$).*

We assume that when the leniency program restricts eligibility to the first reporting firm only and agents agree to collude and report, either they report simultaneously – so that the reduced fine is allocated randomly – or they take turns to report and each period reallocate the difference between reduced and full fine among cartel members. Under either assumption, when firms agree to both collude and report, a firm's expected fine in each period is $\frac{(N-1)F+RF}{N}$.

Clearly, firms will choose to collude and report only when the leniency program is exploitable, so that $V_i^{c\prime} > V_i^c$. Since colluding agents can always choose not to report, the

[22] Conversely, the worst thing the law enforcement agency can do is looking at unilateral defections and price wars as signals of collusive/criminal behavior and target all cartel members. This policy would stabilze cartels, as a unilateral defection would then not only be punished by partners; it would also increase the probability of being fined for past collusive behavior.

value of colluding cannot decrease with the introduction of a leniency program, it will be $\max\{V_i^c, V_i^{c\prime}\}$. It remains to be checked under which circumstances it is $V_i^{c\prime} > V_i^c$, so that the leniency program can be exploited. The following lemma characterizes such circumstances.

Lemma 2 *Suppose all reporting agents are eligible to leniency. Then:*

1. *When $F^r = F$, the leniency program is exploitable if $RF < \alpha F$;*
2. *When $F^r > F$, the leniency program is exploitable if $RF < \alpha F + \alpha^2 \delta_i \frac{F^r - F}{1-\delta_i}$.*

Suppose only the first reporting agent is eligible to leniency. Then:

3. *When $F^r = F$, the leniency program is exploitable if $RF < \alpha F - (N-1)(1-\alpha)F$;*
4. *When $F^r > F$, the leniency program is exploitable if*

$$RF < \alpha F - (N-1)(1-\alpha)F - \delta_i(F^r - F)(N - 1 - \frac{\alpha^2}{1-\delta_i}).$$

Comparing statements 1-2 with 3-4 one sees that restricting eligibility to the first firm that reports reduces the set of exploitable leniency programs; it restricts the range of reduced fines/rewards that increase the value of colluding. This is of course the case because restricting eligibility to one firm implies that each time firms both collude and report, all but one firms must pay the full fine. Therefore, all else equal, fines' reductions/rewards must be much larger for a restricted leniency program to become exploitable than for an unrestricted one. In the relevant parameter configuration ($\alpha < \frac{1}{2}$) restricting eligibility allows to reward reporting agents – the more generously the smaller α and the larger N – without making the program exploitable. On the contrary, unrestricted programs become exploitable already when the reduced fine equals or falls below the expected fine.

Comparing Lemma 2's statements 1 and 2 one sees that when eligibility is not restricted, the range of exploitable leniency programs is larger when the expected fines for repeated/recidivous offenders are higher. This is because leniency programs do not distinguish between first time and repeated offenders, and reducing all fines to RF they undo the increase in post-conviction deterrence linked to the higher fines for repeated offenders. Finally, comparing statements 3 and 4 one sees that with restricted eligibility, for the interesting range of parameters ($\alpha < \frac{1}{2}$) and N not too small, the range of exploitable leniency programs is smaller when the expected fines for a repeated/recidivous offender are higher. Then even when agents collude and report, after the first period agents face the increased sanctions for recidivous offenders.

4.2 "Effective" leniency programs

We now turn to the effects of a Leniency Program on the value of unilaterally defecting from collusive strategies. Let $V_i^{d\prime}$ denote the value of defecting and simultaneously reporting information to the law enforcing agency. The natural assumption we adopt here is that if an agent decides to both defect unilaterally from collusive strategies and report, he will be able to report first even when collusive strategies already prescribe firms to report (in turn or simultaneously) along the equilibrium path.

The value of defecting cannot decrease with the introduction of the leniency program, it will be $\max\left\{V_i^d, V_i^{d\prime}\right\}$ since when $V_i^{d\prime} < V_i^d$ a defecting agent can always choose not to report. And of course as long as $V_i^{d\prime} < V_i^d$ the leniency program cannot be "effective" in terms of increasing agents' temptation to unilaterally defect and report.

Definition 3 *A leniency program is "effective" in industry i when it allows a firm that unilaterally defects from a collusive agreement to increase its payoff by reporting information (when $V_i^{d\prime} > V_i^d$).*

The next lemma characterizes the circumstances under which the value of defecting increases with the introduction of the leniency program.

Lemma 3 *Independent of how many firms are eligible to leniency:*

1. *When $F^r = F$, the leniency program is effective if $RF < \gamma F$;*
2. *When $F^r > F$, the leniency program is effective if $RF < \gamma F + \frac{\delta_i(1-\beta)}{1-\delta_i}(v^{pr} - v^p)$.*

Lemma 3 tells us that – in our model – there are two ways in which the leniency program may be effective in deterring cartels crime by increasing the value of defecting from an illegal agreement. First, the leniency program protects an agent that defects and reports from the expected fine γF he would otherwise face when defecting. When repeated offenders are not subject to higher fines ($F_i^r = F_i$), only this *protection from fines* effect is at work, and a leniency program increases the value of defecting (and reporting) if the reduced fine RF is smaller than the expected fine γF. When repeated offenders are subject to higher fines ($F_i^r > F_i$) and collusive agreements are sustained by optimal two-phase punishments, a second effect enters the scene. Firms are then willing to bear the costs of a short, harsh punishment phase – say a tough price war – because they expect to go back to collusion right after it. By reporting under the leniency program a defecting firm ensures that the cartel is convicted, and that thereafter further collusion is punished with the higher fines F^r. This reduces the expected profits from further collusion v^{cr}, and the costs firms are willing to incur to punish the initial defection. We name this *protection from punishment* effect: by reporting under the leniency program a defecting firm reduces future collusive profits, thereby softening the punishment it faces for its defection (getting v^{pr} instead of v^p after the defection, where $v^{pr} > v^p$).

5 "Courageous" leniency programs

In this section we use the results just derived to characterize (unconstrained) optimal leniency programs. Unconstrained optimal programs can be labelled "courageous" because it turns out that – for realistic parameters configurations – they prescribe that a substantial reward should be paid to the first agent that reports information to the law enforcing agency. In contrast, leniency programs implemented in reality are often "moderate", in the sense that they only reduce, or at best cancel sanctions for reporting firms. Such constrained-optimal moderate programs will be discussed in Section 6.

5.1 Optimal eligibility

Given other parameters of the law enforcement policy, an optimal leniency program trades off the costs it implies, if any, with the benefits of making the incentive compatibility conditions

$$\max\left\{V_i^c, V_i^{c\prime}\right\} > \max\left\{V_i^d, V_i^{d\prime}\right\} \tag{IC}$$

more stringent thereby increasing the minimum discount factor at which collusion can be sustained (hence reducing the set of industries where collusion is feasible).

We assumed that administering fines and rewards is costless, so that reducing the fines imposed and increasing the rewards paid by extending eligibility to all reporting firms cannot increase welfare.

Contrasting Lemma 2's statements 1-2 with 3-4 we have seen that restricting eligibility to the first firm that reports reduces the set of 'exploitable' leniency programs, enlarging the set of fine discounts/rewards that can be awarded without increasing the value of colluding. On the other hand, from Lemma 3 we know that the attractiveness of defecting and reporting is independent of the eligibility criteria, it only depends on the size of the fine discount/reward. Therefore, restricting eligibility to the first reporting firm allows to increase the right hand side of condition (IC) by awarding larger fine discounts/rewards, while leaving the left hand side of (IC) unchanged (i.e. without affecting the value of colluding). This conclusion can be restated as follows.

Proposition 2 *An optimal leniency program is restricted to the first reporting agent.*

In our model there are no gains from extending eligibility to leniency beyond the first reporting firm, while doing it constrains the size of the fine discount/reward that can be awarded to induce a cartel member to report. The result appears consistent with how real world leniency programs are designed, i.e. with a large difference between the amount of leniency obtainable by the first reporting party (automatic complete amnesty) and that available to further reporting parties (discretional, partial reductions of sanctions).[23]

There is a further reason to restrict eligibility to amnesty to the first reporting party, not yet captured by our model but often stressed by DoJ officials, who see it as crucial to the effectiveness of the program. In reality, the first comer restriction appears to generate "races to report" caused by the "fear to arrive second". Were the second, third or forth reporting firms eligible to the same treatment as the first one such races would arguably not occur. Then firms could safely adopt a 'wait and see' strategy ("do not report first, be ready to report if somebody else does it"). We will try to capture this effect in Section 7, where we introduce strategic risk considerations.

Since we are interested in *optimal* leniency programs, the remainder of the paper will focus on programs restricted to the first reporting party (unless otherwise specified). For these programs, the self-financing constraint implies $RF \geq \underline{RF} = -(N-1)F$.

[23] A smaller fine discount for the second reporting firm would become optimal in our model if we relaxed the (standard) assumption that if a firm reports the cartel is convicted with probability one, and assumed that a second report would increase the probability of conviction. Such an extension, however, would increase complexity and length of the paper without affecting any of its central results, so it is left to future work.

5.2 Optimal fine reductions/rewards

Our characterization of the optimal reduced fine/reward RF^* will focus on two cases. The first is the case considered in all the related literature, with $F = F^r = \overline{F}$ (no increase in sanctions for recidivous offenders) and agreements enforced by grim trigger strategies, so that $v^{cr} = \pi^c$ and $v^p = v^{pr} = \pi^n$. The second case allows punishment strategies to be tougher than grim trigger ones, so that $v^p < \pi^n$, and fines for recidivous offenders to be "large", so that after being convicted collusion becomes too risky and $v^{pr} = v^{cr} = \pi^n$. These two cases are sufficient to highlight all effects at play, and can be fully characterized at the current level of generality. Determining the exact RF^* in intermediate cases requires specifying both the underlying stage game and the punishment strategies to calculate $\partial v^p / \partial RF$. Though, there would be little gain to compensate for the loss of generality: it is easy to verify that RF^* would always be included between the levels obtained in the two cases we analyze, so that the conclusions derived for these cases apply to intermediate cases as well.

Assuming that when agents are indifferent between reporting and not reporting they choose to report, one can state the following.

Proposition 3 *When $F = F^r$ and $v^p = \pi^n$, the optimal leniency program (is restricted to the first reporting party and) has*

$$\begin{cases} RF^* = \alpha F - (N-1)(1-\alpha)F & \text{when} \quad F > \pi^c - \pi^n - \frac{\pi^d - \pi^c}{N-1}, \\ RF^* = \underline{RF} = -(N-1)F & \text{otherwise.} \end{cases}$$

When $F < F^r$ and $v^p < \pi^n = v^{pr}$ the optimal leniency program has $RF^* = \underline{RF}$.

This result stands to leniency programs as Becker's (1968) result that optimal fines should be maximal stands to classical law enforcement models.

First, note that in all cases the optimal leniency program establishes a *positive reward* for the first firm that reports information on its cartel. The intuition is of course that rewards improve the effectivity of leniency programs and cartel deterrence by increasing firms' gains from defecting from the collusive agreement and reporting it to the law enforcing agency (the right hand side of condition IC). Of course, too high rewards may make the program exploitable, hence the optimal reward may be smaller than the level at which the self-financing constraint binds. This is what happens in the first case, where $F = F^r$ and $v^p = \pi^n$, as long as fines are not very small (as long as they are close to or larger than the per-period collusive markup). Then the optimal reward is the minimum one that makes the program exploitable (identified by Lemma 2) leaving colluding firms indifferent between reporting and not reporting ($V_i^{c\prime} = V_i^c$; firms then report by our tie-breaking assumption). It is optimal that they report since this allows to save on inspection costs $c_\alpha(\alpha)$. It is not optimal to increase the reward further because it would increase the program's exploitability ($V_i^{c\prime}$) more than its effectivity ($V_i^{d\prime}$), thereby reducing deterrence. When fines are smaller than the per-period's collusive markup, increasing the reward above (decreasing RF below) the level where the program becomes exploitable increases deterrence, as the increase in effectivity dominates on the increase in exploitability, and it is optimal to set the reward maximal. In the second case, where the higher fines for repeated offenders $F^r > F$ ensure that agents cannot sustain collusion after being convicted, agents cannot exploit the program (by colluding and reporting) for more than one period. Then effectivity considerations dominate, and the optimal reward is again the maximal one.

A second thing to note is that in the first case, where fines are not very small, the optimal reward is decreasing in the detection probability α. This means that *investigations and rewards are substitute law enforcement instruments*. When α is large, the optimal reward must be small because a high probability of paying fines when not reporting makes even moderate rewards for reporting attractive and leniency programs exploitable. The optimal self-financing reward is instead in all cases increasing in fines, so that *fines and rewards are complement instruments.* Independent of the inspection probability, heavy fines are good because they allow to award and finance larger rewards without making the leniency program exploitable.

Finally, note that RF^* is increasing in the number of agents involved. More agents make a program (restricted to the first applicant) less exploitable: more fines have to be paid and in more parts must the reward be split when agents choose to collude and report.

5.3 Optimal law enforcement: achieving the first best

Let us now come to law enforcement policy. Since investigations and rewards are substitutes, investigations cost $c_\alpha(\alpha)$ while self-financing rewards are costless, an optimal law enforcement policy should rely as much as possible on self-financing rewards. In fact, we easily obtain the following.

Proposition 4 *There exists a finite level of fines F' such that when $F' \leq \overline{F}$ the optimal law enforcement policy achieves the* first best *– complete and costless deterrence – with $\gamma = \beta = \alpha = 0$, $F' \leq F \leq \overline{F}$, and $-(N-1)F \leq RF \leq -(N-1)F'$.*

This is a simple but remarkable result. It tells us that there is a finite level of fines that – if politically feasible – allows to completely deter collusion in all industries and at no cost. This is done by setting the reward for the first reporting agent equal to the sum of the fines paid by his former partners, interrupting all forms of costly investigations and laying back on the chair waiting for wrongdoers to come forward with information. In other words, in this model the combination of sufficiently high fines and optimally designed, high powered leniency programs make the public enforcement of law – the active investigation of organized crime – redundant, and actually suboptimal.

The result is also remarkable because, to our knowledge, it is the first time that the first best is achieved in a law enforcement model *á la* Becker (1968).[24] Most previous work on optimal law enforcement focused on individual crimes – where nobody has freely available information on the crime besides the criminal – and shares the property of Becker's original model where even infinite fines cannot achieve the first best. A strictly positive probability of detection is necessary for law enforcement to have any effect, and the investigation costs that generate such positive probability are a deadweight loss that keeps society away from the first best. In these models complete deterrence is generally not optimal: the optimal positive amount of residual, undeterred crime equalizes marginal social benefits and costs of deterrence.

[24] Of course there will be costs linked to the court system, who has to evaluate/verify the information reported. These verification costs are usually disregarded in the law and economics literature stemming from Becker (1968), they are considered unavoidable. These costs would be present with and without a leniency program.

On the contrary, in the present, organized crime framework fellows wrongdoers possess information on each other's crime. A sufficiently high, finite reward – financed by the sufficiently high, finite fines the reward generates – can elicit such information. Since self-financing rewards are costless transfers for society, as are Becker's fines, complete deterrence is then optimal and achieved at no cost. Of course, the sharpness of this result is partly due to the relative simplicity of the model. As happened to Becker's conclusion that fines should always be set maximal, our result can probably be softened by complicating the model introducing other aspects of reality. Still, as Becker's result, Propositions 3 and 4 establish a benchmark for future work on the optimal design of deterrence mechanisms against cartels and organized crime.

6 Constrained-optimal leniency programs

Exogenous political and institutional factors may constrain the design of the law enforcement policy. The most obvious way in which the design is usually constrained is in the size of fines and of fine discounts/rewards. In this section we consider the optimal design of the law enforcement policy when institutional restrictions on the size of fines and of fine discounts/rewards are binding.

6.1 Constraints on fines

When exogenous factors constrain fines to be smaller than the level that leads to the first best ($\overline{F} < F'$), as appears to be the case in many countries (particularly in the EU), the second best law enforcement policy implies positive investigation costs and may imply a non maximal reward. The next proposition characterizes the second best law enforcement policy when the upper bound on fines is binding.

Proposition 5 *When the first best cannot be achieved because of a too low upper bound on fines ($\overline{F} < F'$), the optimal law enforcing policy has $\gamma = 0$, $F = \overline{F}$, $F^r = \overline{F}^r$, $RF = RF^*$ and:*

1. *$\alpha > 0$ and such that $c'_\alpha(\alpha)$ equals the marginal social benefit of deterrence, when $F = F^r$, $v^p = \pi^n$ and $F > \pi^c - \pi^n - \frac{\pi^d - \pi^c}{N-1}$;*
2. *$\alpha = 0$, when $F = F^r$, $v^p = \pi^n$ and $F < \pi^c - \pi^n - \frac{\pi^d - \pi^c}{N-1}$;*
3. *$\alpha > 0$ and such that $c'_\alpha(\alpha)$ equals the marginal social benefit of deterrence, when $F^r > F$ and $v^{cr} = \pi^n$.*

The proposition tells us that when maximal fines are too small to achieve the first best through self-financing rewards, it may be optimal to couple rewards with active investigations (statements 1 and 3). Note that in the first case – where $F = F^r$, $v^p = \pi^n$ and fines are not too small – investigations and rewards are substitute instruments, so that since the second best implies a positive α, it also implies less than maximal rewards (statement 1). Only in the less plausible case where $F = F^r$, $v^p = \pi^n$ and fines are smaller than the per-period collusive markup it is never optimal to investigate (statement 2).

6.2 Constraints on rewards: "moderate" leniency programs

Offering rewards to wrongdoers that cooperate with prosecutors against their former partners is a consolidated practice in the US (see the discussion of misreporting in Section 8). However, all *codified* leniency programs we are aware of are "moderate", in the sense that they do not explicitly allow to reward a wrongdoer that reports information and cooperates with the law enforcing agency. They only allow to reduce, or at best cancel sanctions against agents that spontaneously self-report. For this reason in this section we consider the optimal design of moderate leniency programs, constrained to non-negative reduced sanctions for wrongdoers who self report (the optimization is constrained by $RF \geq \underline{RF} = 0$); and in the next section we consider their deterrence effects.

The following proposition characterizes the optimal moderate leniency programs and the correspondent optimal law enforcement policy.

Proposition 6 *Suppose leniency programs are constrained to be moderate ($\underline{RF} = 0$). Then:*

1. *The constrained-optimal leniency program is restricted to the first reporting party and has $RF^{**} = \max\{RF^*, 0\}$, where RF^* is defined in Proposition 3;*
2. *The optimal law enforcement policy has $\gamma = 0$, $F = \overline{F}$, $F^r = \overline{F}^r$, $RF = RF^{**}$ and $\alpha > 0$ and such that $c'_\alpha(\alpha)$ equals the marginal social benefit of deterrence.*

The first statement obtains because the reasoning behind Propositions 2 and Proposition 3 continues to apply when the leniency programs are constrained to non-negative reduced fines (not to pay rewards). The constraint simply determines a corner solution ($RF = 0$) whenever the unconstrained optimal leniency program would require a reward. Since in the relevant parameter space ($\alpha < \frac{1}{2}$) it is $RF^* < 0$, in this region it will be $RF^{**} = 0$. The second part tells us that, as expected, with a moderate leniency program it will be generally optimal to spend resources to actively investigate cartels.

6.3 Deterrence effects of moderate leniency programs

Restricting focus to grim trigger strategies and assuming $\gamma = 0$ would immediately lead to conclude that a moderate leniency program restricted to the first, spontaneously reporting party cannot have deterrence effects. This is because the incentive to defect, the left hand side of condition (IC), is not reinforced by such a program: a defecting agent does no better by reporting under a moderate leniency program than by just not reporting, which is possible with or without the leniency program.[25]

However, we cannot take this irrelevance result too seriously. First, because the restriction to $\gamma = 0$ and grim trigger strategies is not empirically warranted: as already mentioned, defecting from a cartel today does not usually guarantee not to be convicted for yesterday's wrongdoing (although we showed it would be optimal if it did); and grim trigger strategies are suboptimal for many oligopoly models, while real world punishment phases are often short and tough. Second, because the result is inconsistent with the experience of the US's

[25] A previous version of this paper assumed $\gamma = 0$ and emphasized this "irrelevance result" for the case where firms use grim trigger strategies; analogous results are derived by Motta and Polo (2001) and Rey (2001), who also assume $\gamma = 0$ and grim trigger strategies.

DoJ where in the last years about half of the applications for leniency have been falling under Section A of the program, suggesting that moderate programs do have direct destabilizing effects on cartels (lead many agents to spontaneously self-report).

Allowing γ to be positive and firms to use other strategies than grim trigger ones break up the irrelevance result and unveil some reasons why moderate leniency programs may have direct deterrence effects. Note first that the constraint $\underline{RF} = 0$ and Lemma 2 together imply that moderate leniency programs are never exploitable, so that the left hand side of condition (IC) does not change with their introduction ($V_i^c = V_i^{c\prime}$). Then, only effectivity considerations matter, and we can state the following.

Proposition 7 *Moderate leniency programs have deterrence effects if and only if*

1. $0 < RF < \gamma F, \qquad$ *when* $F = F^r$;
2. $0 < RF < \gamma F + \frac{\delta_i(1-\beta)}{1-\delta_i}(v^{pr} - v^p), \qquad$ *when* $F^r > F$.

The proposition highlights two reasons why moderate leniency programs may have deterrence effects.

First, as long as there is a positive probability γ of being convicted for past collusive activities when one defects from collusive strategies, the *protection from fines* effect is at work: a moderate leniency program with $RF < \gamma F$ increases the value of defecting and reporting by reducing the fine expected by a defecting agent from γF to RF, while leaving the value of colluding unaffected. This effect is independent of which punishment strategies sustain collusion and of whether repeated offenders are punished more severely than first time offenders.

Second, even if the Legislator optimally sets $\gamma = 0$, as long as firms use optimal two-phase punishment strategies and recidivous wrongdoers are punished harder ($F^r > F$), there is the *protection from punishment* effect that encourages defections: a report raises fines for further collusion from F to F^r, which limits the costs firms are willing to incur to punish the agent that defected and reported, so that $(v^{pr} - v^p) > 0$.

When both $\gamma = 0$ and $v^{pr} = v^p$ (either because $F = F^r$ or because firms use grim trigger strategies) the irrelevance result obtains: the conditions in Proposition 7 are necessary ones, so that when no protection from fine nor from punishment takes place moderate leniency programs have no deterrence effects. However, there is at least one additional reason why moderate leniency programs may have direct deterrence effects: they may increase the perceived *riskiness* of entering or maintaining a collusive agreement.

7 Risk dominance

In previous sections we assumed that as long as the IC condition was satisfied a collusive/criminal agreement could be sustained; i.e. that coordination on the collusive agreement was not a serious problem. Under this assumption, to deter a cartel a leniency program had to ensure that the correspondent IC condition was violated. In this section we recognize that in reality to set up an effective collusive/criminal agreement agents must also establish "trust", they must be sufficiently confident that all agents will indeed stick to the agreement. With the multiple equilibria typical of dynamic strategic situations, coordination problems, and in particular *risk dominance* considerations (in the sense of Harsanyi and Selten (1988)) may play an important role.[26] We will here take the view that besides being sustainable in equilibrium, to be viable a collusive/criminal agreement must be sufficiently "safe" from coordination problems: it must not be risk dominated by defecting at any stage. If agents do not play risk dominated equilibria, to deter a collusive/criminal agreement a leniency programs need not go all the way to ensure that the correspondent IC condition is violated: it is sufficient that it makes the corresponding equilibrium risk dominated.[27]

To give a formal account of the potential effects of leniency programs on the riskiness of organized crime and keep what we are doing intuitive for the reader, we will work with a simplified version of our model where industries are duopolistic ($N = 2$) and agreements are supported by grim trigger strategies.[28] We continue assuming that establishing and maintaining coordination on a collusive agreement generates each period hard information that agents could report. Other features of the model remain also unchanged, including the timing: first the Legislation sets policy parameters, then agents interact in the collusive/criminal supergame. To simplify exposition and facilitate comparison with Spagnolo and Blonski (2001) we let here Γ^i denote the supergame beginning in industry i after the Legislator has set the law enforcement parameters; φ denote a collusive equilibrium supported by grim trigger strategies φ_j: "stick to the agreement until a defection is observed, defect forever thereafter", $j = 1, 2$; and ω denote the defection equilibrium where each agent plays the best response of the one-shot game forever.

7.1 Absent or ineffective leniency programs

When no leniency program is present ($RF = F$), or when the program is ineffective because it is not sufficiently generous ($RF > \gamma F$), for an agent that unilaterally defects from a collusive/criminal agreement it is a dominant strategy not to report information to the law enforcing agency. Then, in the defection equilibrium ω each agent plays strategy ω_j : "defect and don't report (in case there is something to report), forever". As in Blonski and Spagnolo (2001), to evaluate the riskiness of a generic collusive/criminal agreement φ we can then focus on a substructure of the supergame, the $2 \times 2-$game $\Gamma^{i\,no}_{\varphi\omega}$ defined by the strategy space $\{\varphi_j, \omega_j\}$ called the $\varphi\omega-$*formation of* Γ^i (the superscript *no* stands for *no* leniency programs). The bimatrix-form of this game when agents establish a collusive/criminal

[26] Harsanyi and Selten (1988) favored payoff-dominance over risk dominance as selection criterion, but the theoretical and experimental support for risk dominance increased since then. Theoretical support has been offered by evolutionary game theory (Michihiro Kandori, George Mailath and Rafael Rob, 1993; Peyton Young 1993) and global games (Hans Carlsson and Eric van Damme, 1993), and experiments showed that agents priviledge risk and security considerations (John van Huyck, Raymond Battalio, and Richard Beil, 1990). Moreover, Harsanyi (1995) proposed later an alternative selection theory where he favoured risk dominance over payoff dominance.

[27] Matthias Blonski and Spagnolo (2001) extended Harsanyi and Selten's definition of risk dominance to fit dynamic games with the strategic features of the repeated Prisoner's Dilemma, and we will heavily rely on their paradigm. Among other things, they show that the critical discount factor below which all cooperation equilibria are risk dominated by defection is strictly higher than the discount factor at which cooperation is supportable in equilibrium. This imples that if risk dominance matters we do not have to care anymore about the IC condition.

[28] Besides ensuring that no "protection from punishment" effect is at work, grim trigger strategies have been shown to be the "safest" strategies with respect to the risk dominance concept: if a collusive/criminal agreement is risk dominated when supported by grim trigger strategies, it is also risk dominated when it is supported by any other punishment strategies (Blonski and Spagnolo, 2001).

agreement is

$\Gamma^{ino}_{\varphi\omega}$	φ_2	ω_2
φ_1	V_i^c, V_i^c	B_i, D_i
ω_1	D_i, B_i	$V_i^n - \gamma F$, $V_i^n - \gamma F$

,

where the values in the matrix are the discounted flows of payoffs agents expect, respectively:
when they both stick to the agreement, V_i^c;
when they unilaterally defect, $D_i = \pi^d - \gamma F + \delta_i V_i^n$, where $V_i^n = \frac{\pi^n}{1-\delta_i}$;
when their opponent unilaterally defects, $B_i = \pi^b - \beta F + \delta_i V_i^n$;
and when they both defect simultaneously, $V_i^n - \gamma F$.

One can then apply Harsanyi and Selten's (1988) original definition of risk dominance to this normal form game, by first transforming it into the best response-equivalent "unanimity game"

$\Gamma^{ino\prime}_{\varphi\omega}$	φ_2	ω_2
φ_1	$V_i^c - D_i$, $V_i^c - D_i$	0, 0
ω_1	0, 0	$V_i^n - \gamma F - B_i$, $V_i^n - \gamma F - B_i$

,

and then calculating the "Nash products" of its two pure strategy equilibria

$$\begin{aligned} u_1(\varphi)u_2(\varphi) &= (V_i^c - \pi^d + \gamma F - \delta_i V_i^n)^2, \\ v_1(\omega)v_2(\omega) &= ((1-\delta_i)V_i^n - \gamma F - \pi^b + \beta F)^2 = \left(\pi^n - \pi^b + (\beta - \gamma)F\right)^2. \end{aligned}$$

A generic criminal/collusive equilibrium φ is then risk dominated by ω when $u_1(\varphi)u_2(\varphi) < v_1(\omega)v_2(\omega)$, and its riskiness is measured by $\rho(\varphi) = v_1(\omega)v_2(\omega) - u_1(\varphi)u_2(\varphi)$.[29] With ineffective or absent leniency programs ($RF > \gamma F$) the riskiness of a generic collusive agreement φ in industry i is then

$$\rho_i^{no} = \left(\pi^n - \pi^b + (\beta - \gamma)F\right)^2 - \left(V_i^c - \pi^d + \gamma F - \delta_i V_i^n\right)^2.$$

Inspecting ρ_i^{no} one sees that also for risk dominance considerations, with poor or absent leniency programs it is optimal to set $\gamma = 0$. Increasing γ is costly and stabilizes criminal/collusive agreements, not only by making the IC condition less stringent (Proposition 1), but also by reducing the riskiness of criminal/collusive equilibria. Increasing α increases riskiness and deters criminal/collusive agreements, and since $c_\alpha(0) = 0$ setting $\alpha > 0$ is optimal. With $\gamma = 0$ and $\alpha > 0$ increases in fines increase the riskiness of collusion, and since higher fines imply no additional costs fines should be set maximal.

[29] From Blonski and Spagnolo (2001) we know that if a cooperation equilibrium supported by grim trigger strategies is not risk dominated by defecting at the beginning, then it is *risk perfect*, i.e. it is not risk dominated by defecting at any later stage.

7.2 Leniency programs and risk

Let us now consider how effective leniency programs (with $RF \leq \gamma F$) affect the riskiness of collusive/criminal agreements. If an agent defects from a collusive agreement it is a dominant strategy for him to report information to the law enforcing agency, hence in the defection equilibrium ω agents play strategy ω_j : "defect and report (in case there is something to report), forever".

Consider leniency programs restricted to the first reporting party. When agents establish a collusive/criminal agreement the bimatrix-form of the $\varphi\omega-$formation game $\Gamma^{i\,rlp}_{\varphi\omega}$ (where rlp stands for restricted leniency programs) is

$\Gamma^{i\,rlp}_{\varphi\omega}$	φ_2	ω_2
φ_1	$\max\{V_i^c, V_i^{c\prime}\}$, $\max\{V_i^c, V_i^{c\prime}\}$	B_i', D_i'
ω_1	D_i', B_i'	$V_i^n - L$, $V_i^n - L$

,

where $D_i' = \pi^d - RF + \delta_i V_i^n$, $B_i' = \pi^b - F + \delta_i V_i^n$, and $L = \frac{RF+F}{2}$, and the riskiness of a generic collusive equilibrium φ in industry i is

$$\rho_i^{rlp} = \left(\pi^n - \pi^b + \frac{F - RF}{2}\right)^2 - \left(\max\left\{V_i^c, V_i^{c\prime}\right\} - \pi^d + RF - \delta_i V_i^n\right)^2.$$

When eligibility to the leniency program is no restricted, i.e. when the program offers the same reduced fine RF to first and second reporting agent, $\Gamma^{i\,ulp}_{\varphi\omega}$ (where ulp stands for unrestricted leniency programs) has instead $D_i' = \pi^d - RF + \delta_i V_i^n$, $B_i' = \pi^b - RF + \delta_i V_i^n$, and $L = RF$, and the riskiness of a generic collusive equilibrium φ in industry i is

$$\rho_i^{ulp} = \left(\pi^n - \pi^b\right)^2 - \left(\max\left\{V^c, V^{c\prime}\right\} - \pi^d + RF - \delta_i V^n\right)^2.$$

Comparing the three measures of riskiness ρ_i^{no}, ρ_i^{rlp} and ρ_i^{ulp} we obtain the following.

Proposition 8 *Let $RF \leq \gamma F$. Then: (1) $\rho_i^{rlp} > \rho_i^{ulp}$; and (2) for non-exploitable leniency programs $\rho_i^{rlp} > \rho_i^{no}$.*

According to (1), collusive/criminal agreements are strictly more risky when eligibility to the leniency policy is restricted to the first reporting agent, than when it is open to all reporting agents. When the deterrence effects of leniency programs are due to the increase in the riskiness of collusion they generate, extending eligibility to full leniency to other reporting agents than the first strictly reduces deterrence. The result is intuitively appealing, since when eligibility is not restricted a colluding agent is "safer" in the sense that he can always enjoy the fine discounts offered by the leniency policy by reporting, whatever other agents do. It reinforces Proposition 2 offering further theoretical support to DoJ official's assertion that the restriction to the first applicant is a crucial feature of the leniency program, generating falls in trust and "rushes to report" among cartel members (e.g. Hammond, 2000).

According to (2), there are always restricted leniency programs that *strictly* increase the riskiness of collusive/criminal agreements (this is not true for unrestricted programs). Note that a restricted leniency program strictly increases the riskiness of collusion even when $RF = \gamma F$, i.e. when the leniency program is "moderate" and there is no "protection from punishment" (so that the "irrelevance result" would obtain with respect to the IC condition). The effect on the riskiness of collusion is the third, important reason identified in this paper why even the current moderate leniency programs may have the direct deterrence effects they appear to have.

To conclude, we characterize the optimal law enforcement policy when risk dominance considerations matter.

Proposition 9 *Suppose agents do not play risk dominated equilibria. Then an optimal law enforcement policy has:*

1. *A leniency program that restricts eligibility to the first applicant;*
2. *Maximal fines* $F = \overline{F}$*;*
3. *Maximal reward* $RF = -F$ *if* $\delta_i < 2/3$*;*
   Reward $RF = -F(1-2\alpha)$ *if* $\delta_i > 2/3$*;*
4. $\gamma, \beta = 0$ *and* $\alpha \geq 0$, *with* $\alpha > 0$ *iff at* $\alpha = 0$ *deterrence is not complete and* $\delta_i > 2/3$.

The first statement is the immediate implication of Proposition 8 and needs no further discussion. The second statement obtains because with restricted leniency programs the riskiness of collusive/criminal agreements ρ_i^{rlp} is monotonously increasing with fines. This is the case also for moderate leniency programs, a result capturing the intuition that these strategic risk by making the event of being fined due to another agent's reporting information more salient. The fact that riskiness increases with the severity of fines suggests that – unless heavier sanctions against cartels are introduced – the new leniency program of the European Commission will likely not be as successful as the DoJ's one, even though the two programs are similar in most aspects.

As for reduced fines/rewards, according to (3) they should be maximal when the discount factor is low, but may have to be kept less than maximal when it is high. This is because high rewards make the program exploitable (when $\alpha > 0$), increasing $V_i^{c\prime}$ and pushing down ρ_i^{rlp}. Since the gains from exploiting the program are distributed in time, this force is stronger the higher the discount factor. Finally, (4) obtains for the same reason why analogous results obtained in Propositions 4, 5 and 6: with leniency programs γ and β are irrelevant, while if deterrence is not complete at $\alpha = 0$ it may be worthwhile to increase deterrence by accompanying the leniency program with public investigations.

8 Concluding remarks

In this final section we briefly discuss some important aspects of the real world that for reason of space could not be incorporated in the model.

Misreporting. Our stylized model with no mistakes in law enforcement highlighted the potential benefits of "high-powered incentives" in law enforcement policy. Of course, if one allows for more realism by introducing asymmetric information and mistakes, will find that there may be drawbacks in offering high rewards to law violators that spontaneously self-report. One potential drawback often put forward as a reason why (nowadays) rewards are seldom used is that these may induce agents to distort/fabricate information. Indeed, the US prosecutors' practice of awarding reductions in sanctions and monetary rewards in exchange for testimony – "soft" information easy to fabricate/distort – is a dangerous and highly debated one.[30] However, this potential drawback can be dealt with directly – restricting eligibility to agents reporting "hard" information, not allowing for testimony, and substantially increasing sanctions against information fabrication/distortion – rather that indirectly, by giving up the potentially large social gains from high-powered leniency programs.

Further research could better clarify the issue. Though, we see at least two reasons why information fabrication/distortion may not be a serious obstacle to the implementation of the optimal schemes proposed in this paper. First, the incentives to distort/fabricate information created by rewards are fully analogous to those generated by the possibility to obtain damage settlements in private law suites. Nobody claims that damage payments should not be allowed for in private suites because they create incentives to fabricate information. In fact, distinguishing reliable from unreliable information – and deciding on the base of the first – is the normal task of courts of justice. Second, fabricating information in a trial is subject to severe criminal sanctions. Innocent parties accused by an agent who fabricated information will have all the incentives – and in the case of cartels the financial resources – to fight back and demonstrate their innocence and the first agent's wrongdoing. Fabricating information to cash rewards appears therefore an extremely risky activity.

Treble damages. When a cartel is successfully prosecuted, *all* former cartel members, including firms that self-reported and cooperated with the Antitrust Authority, are exposed to suits for damages from their customers. How does this feature of reality affects our analysis?

It is easy to realize that taking damages into account does not alter any of our conclusions. Let $E[D]$, with $E[D] > 0$, denote the damages a firm expects to pay if convicted for collusive behavior, and $E[RD]$, with $E[RD] \geq 0$, the damages a firm that spontaneously self-reports expects to pay, with $E[RD] \leq E[D]$ (at present both in the EU and US it is $RD = D$, but in light of our previous results it is not hard to see why one may wish to protect reporting firms by setting $E[RD] < E[D]$). Now let us redefine variables in the previous sections so that $F = MF + E[D]$ and $RF = RMF + E[RD]$, where MF and RMF stand for fines and reduced fines respectively. It is immediate to verify that all our results continue to apply, whit the only difference that fine reductions/rewards must be increased to compensate for expected damage payments. Moreover, if we allow $E[RD]$ and $E[D]$ to differ, it becomes clear that as long as increasing rewards for self-reporting firms is more costly than modifying the law to protect them from damage suits, the Legislator's optimal policy is to set $E[RD] = 0$. In practice, this implies that present leniency programs are not

[30]Some debated cases where US prosecutors exchanged rewards against testimony are discussed at *http://www.pbs.org/wgbh/pages/frontline/shows/snitch/readings/paying.html*paying.html Reporting agents were asked to testify even though the provisions of 18 USC Section 201(c)(2) explicitely makes it an offense to pay a witness for testifying. As mentioned, we regard as a mistake to ask reporting agents to testify when they receive leniency. In this paper we excluded testimony assuming that to obtain leniency an agent must provide "hard" information (videotapes, documents, etc.).

even constrained-optimal: they would be optimal moderate programs only if they would protect a reporting firm from being sued for damages.

Restitution. According to the US Corporate Leniency Program, to obtain leniency self-reporting firms are required to pay back collusive profits to customers ("if they can," that is, if this do not drive them bankrupt; see Spratling, 1998). It is easy to verify that when self-reporting firms must pay back realized collusive profits to customers the attractiveness of defecting and reporting is reduced. As for damages, then higher fine discounts/rewards are needed to compensate for these additional losses if one wishes to deter cartels by inducing firms to spontaneously self-report. The restitution requirement is unambiguously counterproductive and should be eliminated.

Individuals vs. Organizations. In the paper we focused on generic agents or firms. In reality, agents of criminal organizations are sometimes themselves organizations composed of many individuals. This is the case for firms within a cartel. Allowing individual employees that report information on a cartel in which their firm is involved to cash the optimal rewards discussed in this paper (the sum of all fines paid by convicted firms) exponentially increases agents' incentives to report, and the deterrence effects of the program. However, individual leniency programs undermine trust not only between, but also within each colluding firms, which may be costly. Aubert et al. (in progress) focus of the additional deterrence effects generated by individual leniency programs, and on the potential costs of the internal lack of trust they bring with.

Violence. Criminal organizations often arrange for credible, violent sanctions against members that turn them down. This may even be true for cartels. For example, Gambetta and Reuter (1995) argue that, in Sicily, Mafia has met the enforcement demand of oligopolistic firms with a supply of coordination and enforcement services, particularly in procurement auction markets. In these situations, firms (or executives) that self-report risk their life, and to be effective leniency programs must try to compensate for this risk by providing effective protection, besides sufficiently high rewards.

9 Appendix: Proofs

Proof of Lemma 1. Fines are costless to administer and, being transfers, do not directly affect social welfare, hence an optimal law enforcement policy sets them to maximize crime/cartel deterrence by making inequality (2) as stringent as possible. The conditions in the statements obtain by differentiating (2). Suppose first that F and F^r can be set independently. Bringing all terms of (2) on the left hand side we obtain net gains from collusion

$$V_i^c - V_i^d = \frac{\pi^c + \alpha(\frac{\delta_i v^{cr}}{1-\delta_i} - F)}{1 - \delta_i(1-\alpha)} - \pi^d + \gamma F - \frac{\delta_i(\beta v^{pr} + (1-\beta)v^p)}{1-\delta_i}.$$

Differentiating we obtain

$$\frac{\partial(V_i^c - V_i^d)}{\partial F} = -\frac{\alpha}{1-\delta_i(1-\alpha)} + \gamma - \frac{\delta_i(1-\beta)}{1-\delta_i}\frac{\partial v^p}{\partial F},$$

which is negative when

$$\gamma < \frac{\alpha}{1-\delta_i(1-\alpha)} + \frac{\delta_i(1-\beta)}{1-\delta_i}\frac{\partial v^p}{\partial F},$$

from which statements 1*i*) follows.

Consider now the effect of changes in F^r on net gain from (first time) collusion. We have

$$\frac{\partial(V_i^c - V_i^d)}{\partial F^r} = \frac{\delta_i}{1-\delta_i}\left[\frac{1}{(1-\delta_i(1-\alpha))}\alpha\frac{\partial v^{cr}}{\partial F^r} - \beta\frac{\partial v^{pr}}{\partial F^r}\right],$$

negative when $\frac{\alpha}{(1-\delta_i(1-\alpha))\beta}\frac{\partial v^{cr}}{\partial F^r} \leq \frac{\partial v^{pr}}{\partial F^r}$. To understand the sign of $\frac{\partial v^{cr}}{\partial F^r}$, note first that if F^r is so high that collusion cannot be sustained after the first conviction, then $V_i^{cr} = \frac{\pi^n}{1-\delta_i}$ and $\frac{\partial v^{cr}}{\partial F^r} = 0$. If collusion can still be sustained then expected gains from further collusion are $V_i^{cr} = \frac{\pi^c - \alpha F^r}{1-\delta_i}$, so that $\frac{\partial v^{cr}}{\partial F^r} = -\alpha < 0$. Since $\frac{\partial v^{pr}}{\partial F^r} \geq 0$ statement 1*ii*) follows.

Suppose now that it must be $F = F^r$. Then $V^c = \frac{\pi^c - \alpha F}{1-\delta_i}$, $V^d = \pi^d - \gamma F + \frac{\delta_i v^p(F)}{1-\delta_i}$, $V_i^c - V_i^d = \frac{\pi^c - \alpha F}{1-\delta_i} - \pi^d + \gamma F - \frac{\delta_i v^p}{1-\delta_i}$

$$\frac{\partial(V_i^c - V_i^d)}{\partial F} = \frac{-\alpha}{1-\delta_i} + \gamma - \frac{\delta_i}{1-\delta_i}\frac{\partial v^p}{\partial F},$$

which is negative as long as

$$\gamma < \frac{\alpha}{1-\delta_i} + \frac{\delta_i}{1-\delta_i}\frac{\partial v^p}{\partial F}$$

and positive otherwise. Statement *2* follows. ■

Proof of Proposition 1. An optimal law enforcement policy maximizes deterrence by making condition (2) as stringent as possible while minimizing enforcement costs (c_k). Since $\frac{\partial(V_i^c - V_i^d)}{\partial\gamma} = F \geq 0$, when $F > 0$ condition (2) is more stringent the smaller is γ. This and $c'_\gamma > 0$ imply that setting $\gamma = 0$ is optimal. Lemma 1 and $\gamma = 0$ together imply that the optimal fines are $F = \overline{F}$, and $F^r = \overline{F^r}$. ■

Proof of Lemma 2. If firms agree to collude and report, each firm's expected gains from collusion are

$$V_i^{c\prime} = \pi^c - RF + \frac{\delta_i v^{cr}}{1-\delta_i}.$$

1) If all firms are eligible and $F_i^r = F_i$ then $v^{cr} = (1-\delta_i)V_i^c$ and $V_i^c = \frac{\pi^c - \alpha F}{1-\delta_i}$, $V_i^{c\prime} = \frac{\pi^c - RF}{1-\delta_i}$ and $V_i^{c\prime} > V_i^c$ when $RF < \alpha F$.

2) If all firms are eligible but $F_i^r > F_i$, it is still $V_i^{c\prime} = \frac{\pi^c - RF}{1-\delta_i}$, but after being convicted $v^{cr} = \pi^c - \alpha F^r < \pi^c - \alpha F = v_i^c$, hence

$$\begin{aligned} V_i^c &= \frac{\pi^c + \alpha(\frac{\delta_i v^{cr}}{1-\delta_i} - F)}{1-\delta_i(1-\alpha)} = \frac{\pi^c - \alpha F + \alpha\delta_i V_i^{cr}}{1-\delta_i(1-\alpha)} \\ &= \frac{\pi^c - \alpha F - \alpha^2\delta_i\frac{F^r - F}{1-\delta_i}}{1-\delta_i}. \end{aligned}$$

Then

$$\begin{aligned} V_i^{c\prime} - V_i^c &= \frac{\pi^c - RF}{1-\delta_i} - \frac{\pi^c - \alpha F - \alpha^2 \delta_i \frac{F^r - F}{1-\delta_i}}{1-\delta_i} > 0 \text{ iff} \\ &= \alpha F + \alpha^2 \delta_i \frac{F^r - F}{1-\delta_i} > RF. \end{aligned}$$

3) If only one firm is eligible and $F^r = F$ then again $v^{cr} = (1-\delta_i)V_i^c$ and $V_i^c = \frac{\pi^c - \alpha F}{1-\delta_i}$, but now $V_i^{c\prime} = \frac{\pi^c - \frac{(N-1)F + RF}{N}}{1-\delta_i}$,

$$\begin{aligned} V_i^{c\prime} - V_i^c &= \frac{\alpha F - \frac{(N-1)F+RF}{N}}{1-\delta_i} = \frac{\alpha F - (N-1)(1-\alpha)F - RF}{N(1-\delta_i)} > 0 \text{ when} \\ RF &< N\alpha F - (N-1)F \text{ or } RF < \alpha F - (N-1)(1-\alpha)F. \end{aligned}$$

4) If only the first reporting firm is eligible and $F_i^r > F_i$, so that $v^{cr} < v^c$, then $V_i^c = \frac{\pi^c + \alpha(\frac{\delta_i v^{cr}}{1-\delta_i} - F)}{1-\delta_i(1-\alpha)} = \frac{\pi^c - \alpha F - \alpha\delta_i(V_i^c - V_i^{cr})}{1-\delta_i} = \frac{\pi^c - \alpha F - \alpha^2 \frac{\delta_i}{1-\delta_i}(F^r - F)}{1-\delta_i}$,

$$\begin{aligned} V_i^{c\prime} &= \pi^c - \frac{(N-1)F + RF}{N} + \frac{\delta_i}{1-\delta_i}\left(\pi^c - \frac{(N-1)F^r + RF}{N}\right) \\ &= \frac{\pi^c - \left[\frac{RF}{N} + \frac{N-1}{N}\left((1-\delta_i)F + \delta_i F^r\right)\right]}{1-\delta_i}, \end{aligned}$$

and $V_i^{c\prime} - V_i^c > 0$ when

$$\begin{aligned} \frac{\pi^c - \left[\frac{RF}{N} + \frac{N-1}{N}\left((1-\delta_i)F + \delta_i F^r\right)\right]}{1-\delta_i} - \frac{\pi^c - \alpha F - \alpha^2\delta_i \frac{F^r - F}{1-\delta_i}}{1-\delta_i} &> 0 \Leftrightarrow \\ \alpha F - (N-1)(1-\alpha)F - \delta_i(F^r - F)(N - \frac{\alpha^2}{1-\delta_i} - 1) &> RF. \end{aligned}$$

■

Proof of Lemma 3. Without leniency programs $V_i^d = \pi^d - \gamma F + \frac{\delta_i(\beta v^{pr} + (1-\beta)v^p)}{1-\delta_i}$. With the leniency program, the value of defecting is $V_i^{d\prime} = \pi^d - RF + \frac{\delta_i v^{pr}}{1-\delta_i}$.

1) Independent of how many firms are eligible or the defecting only, with $F^r = F$, $v^p = v^{pr}$ and $V_i^{d\prime} > V_i^d$ when

$$\begin{aligned} V_i^{d\prime} &= \pi^d - RF + \frac{\delta_i v_i^p}{1-\delta_i} > \pi^d - \gamma F + \frac{\delta_i v_i^p}{1-\delta_i} = V_i^d \\ &\Rightarrow RF < \gamma F. \end{aligned}$$

2) Independent of whether only the first reporting firm or all firms are eligible, when $F_i^r > F_i$ so that $v_i^p < v_i^{pr}$, $V_i^{d\prime} > V_i^d$ when

$$\begin{aligned} V_i^{d\prime} &= \pi^d - RF + \frac{\delta_i v^{pr}}{1-\delta_i} > \pi^d - \gamma F + \frac{\delta_i(\beta v^{pr} + (1-\beta)v^p)}{1-\delta_i} = V_i^d \\ &\Rightarrow RF < \gamma F + \frac{\delta_i}{1-\delta_i}(1-\beta)(v^{pr} - v^p). \end{aligned}$$

■

Proof of Proposition 2. Follows directly from Lemma 2 (restricting eligibility reduces $V_i^{c\prime}$) and Lemma 3 (restricting eligibility does not affect $V_i^{d\prime}$). ■

Proof of Proposition 3. An optimal leniency program makes the inequality $\max\{V_i^c, V_i^{c\prime}\} > \max\{V_i^d, V_i^{d\prime}\}$ as stringent as possible.

Consider first the case $F_i^r = F_i$ and $v^p = \pi^n$. Then the leniency program is exploitable if $RF < \alpha F - (1-\alpha)(N-1)F$, and is effective if $RF < \gamma F$. Although $\gamma < \alpha$, for the relevant parameter range ($\alpha < \frac{1}{2}$ and $N \geq 2$) it is $\alpha F - (N-1)(1-\alpha)F < 0 < \gamma F$. Then:

In the region $RF > \gamma F$ the leniency program is irrelevant, it is neither exploitable nor effective.

In the region $\alpha F - (N-1)(1-\alpha)F < RF < \gamma F$ the program is effective ($V_i^{d\prime} \geq V_i^d$) and is not exploitable ($V_i^{c\prime} < V_i^c$). From Lemma 3 we know that $V_i^{d\prime}$ is decreasing in RF, so that in this region it is optimal to decrease RF to make the IC condition more stringent by increasing its right hand side.

In the region $RF \leq \alpha F - (N-1)(1-\alpha)F$ the program is both effective and exploitable, and a reduction of RF increases both the left and the right hand side of condition IC. When $RF \leq \alpha F - (N-1)(1-\alpha)F$ condition IC can be written as

$$\begin{aligned} V_i^{c\prime} &\geq V_i^{d\prime} \Leftrightarrow \frac{\pi^c - \frac{(N-1)F+RF}{N}}{1-\delta_i} \geq \pi^d - RF + \frac{\delta_i \pi^n}{1-\delta_i} \\ &\Leftrightarrow \delta_i \geq \underline{\delta} = \frac{\pi^d - \pi^c + \frac{N-1}{N}(F - RF)}{\pi^d - \pi^n - RF} = \frac{\pi^d - \pi^c + \frac{N-1}{N}F - \frac{N-1}{N}RF}{\pi^d - \pi^n - RF}, \end{aligned}$$

and

$$\begin{aligned} \frac{\partial \underline{\delta}}{\partial RF} &= \frac{-\frac{N-1}{N}(\pi^d - RF - \pi^n) + (\pi^d - \pi^c + \frac{N-1}{N}(F - RF))}{(\pi^d - RF - \pi^n)^2} \\ sign\left(\frac{\partial \underline{\delta}}{\partial RF}\right) &= sign\left\{-\frac{N-1}{N}(\pi^d - \pi^n) + \pi^d - \pi^c + \frac{N-1}{N}F\right\} \\ &= sign\left\{\pi^d + (N-1)(\pi^n + F) - N\pi^c\right\} > (<)0 \\ \text{if } F &> (<)\pi^c - \pi^n - \frac{\pi^d - \pi^c}{N-1}. \end{aligned}$$

Consider now the other case, where $F^r > F$ and such that $v^p < v^{pr} = v^{cr} = \pi^n$. Then colluding and reporting $V_i^{c\prime}$ is not sustainable for more than one period, since after the first period that firms collude and report, they become repeated offenders, fines grow up to F^r, and collusion is no more sustainable. This implies that the leniency program is never exploitable: at each point in time agents strictly prefer to defect and report, which delivers $\pi^d - RF + \frac{\delta_i \pi^n}{1-\delta_i}$, than to collude and report, which delivers at best $\pi^c - \frac{(N-1)F+RF}{N} + \frac{\delta_i \pi^n}{1-\delta_i}$. Since V_i^c is not affected by RF while $V_i^{d\prime}$ is decreasing in RF, the optimal program minimizes RF; i.e. $RF^* = \underline{RF} = -(N-1)F$. ■

Proof of Proposition 4. Parameters γ and β do not affect the IC condition and are costly to increase, hence their optimal level is $\gamma = \beta = 0$. From Proposition 3 we know that when $\alpha = 0$ the optimal program has $RF = -(N-1)F$ in all cases (including intermediate ones not characterized by the proposition). Keeping $RF = -(N-1)F$ and letting F grow, the IC condition for each industry becomes more stringent (the left hand side V_i^c is independent of RF and F, while the right hand side $V_i^{d\prime}$ increases with F). Hence there must be a finite level of the fine F' such that for fines higher than this level the IC condition is not satisfied in any industry (the critical discount factor at which collusion is supportable becomes higher than $\delta_{\max}$).■

Proof of Proposition 5. Again, setting $\gamma = \beta = 0$ is optimal because increasing them is costly and does not make the IC condition more stringent, and setting $F = \overline{F}$, $F^r = \overline{F}^r$ is optimal because it makes the IC more stringent than at any other level of fines while raising fines is not costly. In addition:

1) When $F = F^r$, $v^p = \pi^n$ and $F > \pi^c - \pi^n - \frac{\pi^d - \pi^c}{N-1}$, $RF^* = \alpha F - (N-1)(1-\alpha)F$ and the IC condition is

$$IC = \frac{\pi^c - \alpha F}{1-\delta_i} \geq \pi^d - \alpha F + (N-1)(1-\alpha)F + \frac{\delta_i \pi^n}{1-\delta_i}$$
$$\Leftrightarrow \quad \delta_i \geq \underline{\delta} = \frac{\pi^d - \pi^c + (N-1)(1-\alpha)F}{\pi^d - \pi^n - \alpha F + (N-1)(1-\alpha)F}.$$

Since

$$\frac{d\,(\underline{\delta})}{d\alpha} = \frac{-(N-1)F(\pi^d - \pi^n - \alpha F + (N-1)(1-\alpha)F) + NF(\pi^d - \pi^c + (N-1)(1-\alpha)F)}{(\pi^d - \pi^n - \alpha F + (N-1)(1-\alpha)F)^2}$$
$$= F\frac{\pi^d - F - \pi^n - N(\pi^c - F + \pi^n)}{(\pi^d - \pi^n - \alpha F + (N-1)(1-\alpha)F)^2} > 0 \text{ when } F > \pi^c - \pi^n - \frac{\pi^d - \pi^c}{N-1},$$

increasing α increases deterrence by making the IC condition more stringent. Hence, if complete deterrence is not achieved at $\alpha = 0$ and $RF^* = -(N-1)F$, and if $c'_\alpha(0)$ is smaller than the marginal social benefit of further deterrence, it is optimal to set $\alpha > 0$ and $RF^* = \alpha F - (N-1)(1-\alpha)F < -(N-1)F$.

2) When $F = F^r$, $v^p = \pi^n$ but $F < \pi^c - \pi^n - \frac{\pi^d - \pi^c}{N-1}$, $RF^* = -(N-1)F$ and the IC condition is

$$\frac{\pi^c - \frac{(N-1)F + RF}{N}}{1-\delta_i} \geq \pi^d - RF + \frac{\delta_i \pi^n}{1-\delta_i},$$
$$\Leftrightarrow \quad \delta_i \geq \underline{\delta} = \frac{\pi^d + (N-1)F - \pi^c}{\pi^d + (N-1)F - \pi^n}.$$

This is the best exploitable program. The best non-exploitable one (with $RF > \alpha F - (N-1)(1-\alpha)F$) delivers

$$\delta_i \geq \underline{\delta}' = \frac{\pi^d + (N-1)(1-\alpha)F - \pi^c}{\pi^d - \alpha F + (N-1)(1-\alpha)F - \pi^n},$$
$$\text{where } \frac{\partial \underline{\delta}'}{\partial \alpha} < 0 \text{ since } F < \pi^c - \pi^n - \frac{\pi^d - \pi^c}{N-1}.$$

Hence $\alpha = 0$ and $RF^* = \underline{RF} = -(N-1)F$ is the global optimum.

3) When $F^r > F$ and $v^p = v^{pr} = v^{cr} = \pi^n$ the leniency program is never exploitable since after a report collusion is no more sustainable, in which case it is a dominant strategy for each agent to immediately report and defect from any agreement. Then the relevant IC condition is $V_i^c > V_i^{d\prime}$. Since $\frac{\partial V_i^c}{\partial \alpha} < 0$ and $\frac{\partial V_i^d}{\partial \alpha} = 0$, an increase in α does make the IC condition more stringent. It follows that when complete deterrence is not costlessly achieved by setting $\alpha = 0$ and $RF = \underline{RF} = -(N-1)F$, if $c_\alpha'(0)$ is smaller than the marginal social benefit of additional deterrence the second best law enforcement policy has $\alpha > 0$ (and $RF^* = \underline{RF} = -(N-1)F$).■

Proof of Proposition 6. The first part follows immediately from Propositions 3 and 4. For the second part, setting $\gamma = \beta = 0$ is optimal because increasing them is costly and does not affect the IC condition, and setting $F = \overline{F}$, $F^r = \overline{F}^r$ is optimal because it makes the IC more stringent at no cost.

For the relevant parameters range ($\alpha < \frac{1}{2}$) it is

$$RF^{**} = 0 > \alpha F - (N-1)(1-\alpha)F > -(N-1)F,$$

so that the program is never exploitable and the IC condition is $V_i^c > V_i^{d\prime}$. Since $\frac{\partial V_i^c}{\partial \alpha} < 0$ and $\frac{\partial V_i^{d\prime}}{\partial \alpha} = 0$, increasing α increases deterrence by making the IC condition more stringent. Hence if complete deterrence is not achieved at $\alpha = 0$ and if $c_\alpha'(0)$ is smaller than the marginal social benefit of further deterrence, it is optimal to set $\alpha > 0$. ■

Proof of Proposition 7. Sufficiency follows immediately from Lemma 3. As for necessity, the only other case is $RF \geq \gamma F$ $(+\frac{\delta_i(1-\beta)}{1-\delta_i}(\pi^n - v^p))$. Assume this is the case. Then $V_i^d \geq V_i^{d\prime}$ – a defecting agent that reports does weakly worse than one who does not – hence $\max\left\{V_i^d, V_i^{d\prime}\right\} = V_i^d$, and given $V_i^c = V_i^{c\prime}$ condition (IC) remains unchanged. ■

Proof of Proposition 8. Restricting eligibility to the leniency program is optimal if $\rho_i^{rlp} > \rho_i^{ulp}$, or

$$\begin{aligned} \left(\pi^n - \pi^b + \frac{F - RF}{2}\right)^2 &> \left(\pi^n - \pi^b\right)^2 \Leftrightarrow \\ \frac{F - RF}{2} &> 0, \end{aligned}$$

which is always satisfied when $RF \leq \gamma F$. Introducing a restricted leniency program is optimal if it increase riskiness without raising costs $\rho_i^{rlp} > \rho_i^{no}$. Considering a non-exploitable leniency program

$$\begin{aligned} \rho_i^{rlp} &> \rho_i^{no} \Leftrightarrow \\ & \left(\pi^n - \pi^b + \frac{F - RF}{2}\right)^2 - \left(V_i^c - \pi^d + RF - \delta_i V_i^n\right)^2 \\ &> \left(\pi^n - \pi^b + (\beta - \gamma)F\right)^2 - \left(V_i^c - \pi^d + \gamma F - \delta_i V_i^n\right)^2 \end{aligned}$$

which is always satisfied because $\beta < \frac{1}{2}$ and $RF \leq \gamma F$. ■

Proof of Proposition 9. Statement 1) follows immediately from Lemma 4.

2) Differentiating we obtain

$$\frac{\partial \rho_i^{rlp}}{\partial F} = 1 - 2\frac{\partial \max\{V_i^c, V_i^{c\prime}\}}{\partial F}.$$

When the leniency program is not exploitable $\max\{V_i^c, V_i^{c\prime}\} = V_i^c$ and since $\partial V_i^c/\partial F = -\frac{\alpha}{1-\delta_i} < 0$ it is $\partial \rho_i^{rlp}/\partial F > 0$. When the program is exploitable $\max\{V_i^c, V_i^{c\prime}\} = V_i^{c\prime}$ and $\partial V_i^{c\prime}/\partial F = -\frac{1}{2(1-\delta_i)} < 0$, and again $\partial \rho_i^{rlp}/\partial F > 0$.

3) Differentiating we obtain

$$\frac{\partial \rho_i^{rlp}}{\partial RF} = -3 - 2\frac{\partial \max\{V_i^c, V_i^{c\prime}\}}{\partial RF}.$$

When the program is not exploitable $\max\{V_i^c, V_i^{c\prime}\} = V_i^c$, and since $\partial V_i^c/\partial RF = 0$ it is $\partial \rho_i^{rlp}/\partial RF = -3 < 0$. Hence it is always optimal to reduce RF (increase fine discounts/rewards) until $V_i^c = V_i^{c\prime}$. Once the program becomes exploitable and $\max\{V_i^c, V_i^{c\prime}\} = V_i^{c\prime}$, $\frac{\partial V_i^{c\prime}}{\partial RF} = -\frac{1}{2(1-\delta_i)}$, and $\partial \rho_i^{rlp}/\partial RF < 0$ iff $\frac{1}{1-\delta_i} < 3$. Hence if $\delta_i < 2/3$ the optimal reward is the maximal one $RF = -F$. If $\delta_i > 2/3$ the optimal reward is such that $V_i^c = V_i^{c\prime}$, i.e. $\alpha F = \frac{F+RF}{2} \Leftrightarrow (2\alpha - 1)F = RF$.

4) With $RF \leq \gamma F$ parameters β and γ do not affect riskiness, and they are costly to increase; hence it is optimal to set them equal to 0. If the leniency program does not achieve complete deterrence at $\alpha = 0$ additional deterrence is socially beneficial; since $c_\alpha'(0) = 0$, at the optimum it must be $\alpha > 0$ if deterrence increases with α. This is always the case when $\delta_i > 2/3$, since then the optimal reward is $RF = (2\alpha - 1)F$,

$$\rho_i^{rlp} = \left(\pi^n - \pi^b + \frac{F - (2\alpha - 1)F}{2}\right)^2 - \left(\frac{\pi^c - (2\alpha - 1)F}{1 - \delta_i} - \pi^d + (2\alpha - 1)F - \delta_i V_i^n\right)^2$$

and

$$\frac{\partial \rho_i^{rlp}}{\partial \alpha} = -2F - 2(\frac{-2F}{1-\delta_i} + 2F) > 0 \text{ if } \delta_i > 1/3.$$

With $\delta_i < 2/3$, by 3) the optimal reward remains maximal $RF^* = -F$ when α grows from 0,

$$\rho_i^{rlp} = \left(\pi^n - \pi^b + \frac{F - RF}{2}\right)^2 - \left(\frac{\pi^c - \frac{F+RF}{2}}{1 - \delta_i} - \pi^d + RF - \delta_i V_i^n\right)^2$$

and $\frac{\partial \rho_i^{rlp}}{\partial \alpha} = 0$. Hence it is optimal to set $\alpha = 0$. ■

References

[1] Abreu, Dilip. "Extremal Equilibria of Oligopolistic Supergames," *Journal of Economic Theory* 39(1), June 1986, 191-225.

[2] Abreu, Dilip. "On the Theory of Infinitely Repeated Games with Discounting", *Econometrica 56,* 1988, 383-396.

[3] Aubert, Cecile; William Kovacic and Patrick Rey. "The Impact of Leniency Programs on Cartels," in progress, IDEI Toulouse.

[4] Becker, Gary. "Crime and Punishment: An Economic Approach," *Journal of Political Economy* 76(2), Mar./Apr. 1968, 169-217.

[5] Blonski, Matthias and Spagnolo, Giancarlo. "Prisoners' Other Dilemma," Working Paper, 2001, University of Mannheim (available from *http//www.ssrn.com*).

[6] Carlsson, Hans, and van Damme, Eric. "Global Games and Equilibrium Selection", *Econometrica* 61, 1993, 989-1018.

[7] Ellis, Christopher and Wilson, Wesley. "Cartels, Price-Fixing, and Corporate Leniency Policy: What Doesn't Kill Us Makes Us Stronger," manuscript, 2002, University of Oregon.

[8] Farrell, Joseph and Makin, Eric. "Renegotiation in Repeated Games", *Games and Economic Behavior* 1, 1989, 327-360.

[9] Fiorentini, Gianluca, and Pelzman, Sam, Eds. *The Economics of Organized Crime*, 1995, Cambridge: C.U.P.

[10] Friedman, James. "A Noncooperative Equilibrium for Supergames," *Review of Economic Studies* 38(113), Jan. 1971, 1-12.

[11] Gambetta, Diego, and Reuter, Peter, "Conspiracy Among The Many: The Mafia in Legitimate Industries," in Fiorentini, Gianluca, and Pelzman, Sam, Eds., 1995, 116-136.

[12] Garoupa, Nuno. "The Theory of Optimal Law Enforcement," *Journal of Economic Surveys* 11(3), Sept. 1997, 267-295.

[13] Garoupa, Nuno, "The Economics of Organized Crime and Optimal Law Enforcement," *Economic Inquiry* 38(2), Apr. 2000, 278-288.

[14] Hammond, Scott D. "Fighting Cartels – Why and How? Lessons Common to Detecting and Deterring Cartel Activity," 2000a, available for download from *http://www.usdoj.gov/atr/public/speeches/6487.htm*

[15] Hammond, Scott D. "Detecting and Deterring Cartel Activity Through an Effective Leniency Program," 2000b, available for download from *http://www.usdoj.gov/atr/public/speeches/speech_ hammond.htm.*

[16] Hammond, Scott D. "When Calculating the Costs and Benefits of Applying for Corporate amnesty, How do You Put a Price Tag on an Individual Freedom?" 2001, available for download from *http://www.usdoj.gov/atr/public/speeches/speech_ hammond.htm*.

[17] Harsanyi, John. "A New Theory of Equilibrium Selection for Games with Complete Information", *Games and Economic Behavior 8,* 1995, 91-122

[18] Harsanyi, John, and Selten, Reinhard. *A General Theory of Equilibrium Selection in Games,* 1988, Cambridge MA: MIT Press.

[19] Innes, Robert. "Remediation and Self-Reporting in Optimal Law Enforcement," *Journal of Public Economics* 72(3), June 1999a, 379-393.

[20] Innes, Robert. "Self-Policing and Optimal Law Enforcement When Violator Remediation is Valuable," *Journal of Political Economy* 107(6), December 1999b, pp. 1305-1325.

[21] Kaplow, Louis, and Shavell, Steven. "Optimal Law Enforcement with Self-Reporting of Behavior," *Journal of Political Economy* 102(3), June 1994, 583-606.

[22] Kandori, Michihiro; Mailath, George, and Rafael Rob. "Learning, Mutation, and Long Run Equilibria in Games", *Econometrica* 61, 1993, 29-56

[23] Konrad, Kai and Skaperdas, Stergios. "Credible Threats in Extortion," *Journal of Economic Behavior and Organization, 33, 1997, 23-29.*

[24] Konrad, Kai and Skaperdas, Stergios. "Extortion," *Economica 65, 1998, 23-39.*

[25] Malik, Arun. "Self-Reporting and the Design of Policies for Regulating Stochastic Pollution," *Journal of Environmental Economics and Management*, 24(3), May 1993, 241-257.

[26] Malik, Arun and Schwab, Robert. "The Economics of Tax Amnesties," *Journal of Public Economics*, 46, 1991, 29-49.

[27] Motta, Massimo, and Polo, Michele. "Leniency Programs and Cartel Prosecution," *International Journal of Industrial Organization,* 21(3), March 2003, 347-379.

[28] Polinsky, Mitchell, and Shavell, Steven. "The Economic Theory of Public Enforcement of Law," *Journal of Economic Literature*, March 2000, 45-76.

[29] Polo, Michele. "Internal Cohesion and Competition among Criminal Organisations," in Fiorentini, Gianluca, and Pelzman, Sam, Eds., 1995, 87-109.

[30] Rey, Patrick. "Towards a Theory of Competition Policy," Invited Lecture at the 5th World Congress of the Econometric Society in Seattle, August 2000 (revised, 2001), IDEI Toulouse.

[31] Spagnolo, Giancarlo. "Self-Defeating Antitrust Laws: How Leniency Programs Solve Bertrand's Paradox and Enforce Collusion in Auctions," June 2000, Nota di Lavoro No. 52.00, Fondazione ENI "Enrico Mattei," Milano (Available from *http://www.ssrn.com/*).

[32] Spratling, Gary R. (Deputy Assistant Attorney, Antitrust Division, U.S. Dept. of Justice). "The Corporate Leniency Policy: Answers to Recurring Questions," presented at the Spring 1998 ABA Meeting (Antitrust Section), available for download at *http://www.usdoj.gov/atr/public/speeches/1626.htm.*

[33] Spratling, Gary R. "Making Companies an Offer they Shoudn't Refuse", 1999, available for download at *http://www.usdoj.gov/atr/public/speeches/2247.htm*

[34] van Huyck, John; Battalio, Raymond, and Richard Beil. "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure", *American Economic Review 80, 1990,* 234-248

[35] Young, Peyton. "The Evolution of Conventions", *Econometrica,* 61, 1993, 57-84.