Quid-Pro-Quo or Winner-Takes-It-All? An Analysis of Corporate Leniency Programs *

Eberhard Feess[†]and Markus Walzl[‡]

February 1, 2008

Abstract

We develop a model that allows for a comparison of past and present corporate leniency programs for cartel behavior as enacted e.g. in the USA and the EU. In particular, we analyze whether fine reductions should be independent of the amount of evidence provided by self-reporting cartels (as it is the case in the US leniency program and the new EU program enacted in 2002) or whether fines should depend on evidence provision (as in the old EU program). We find that only firms which provide the (relatively) highest amount of evidence should be granted a substantial fine reduction. We derive sufficient conditions for the optimality and sub-optimality of a full amnesty for this firm. Firms with lower evidence provision should only be indifferent between self-reporting and facing an expected fine due to self-reporting of the high evidence provider. We use these results to review recent developments in the design of leniency programs.

JEL-Classification: D62, D82, H50, K42
Keywords: self-reporting, optimal law enforcement, criminal teams, leniency programs.

[†]Department of Economics, University of Aachen (RWTH), Templergraben 64, D-52062 Aachen, Germany. Email: feess@rwth-aachen.de.

[‡]Department of Economics and METEOR, University of Maastricht, P.O. Box 616, NL-6200 MD Maastricht, The Netherlands. Email: m.walzl@algec.unimaas.nl.

1   Introduction

In February 2002, the Commission of the European Union has substantially revised its corporate leniency program for cartel prosecution from 1996. Although there are still some differences, the new legislation is much closer to the US program originally enacted already in 1978 and revised in 1993. In particular, the EU program now also grants full immunity for first self-reporters if some basic requirements are fulfilled.[1] By contrast, the program valid until 2002 usually granted only partial immunity even to a first self-reporter, and the fine reduction was dependent on the amount of evidence provided to convict other cartel members. Our paper focuses on this institutional change and analyzes the pros and cons of full amnesties vs evidence-dependent fine reductions. Hence, we compare the old European program on the one hand to the US program and the new European program on the other hand.[2]

In our model, two firms form a cartel whenever the expected (aggregated) increase in profits is above expected fines. Our model has infinite time horizon, and the cartel yields a social loss (i.e., a reduction of the sum of consumer- and producer surplus in each period). The authority seeks to minimize this loss by choosing the fine reduction for the first self-reporting firm in an optimal way. To elicit the optimal dependence on the amount of provided evidence, we assume that firms are heterogenous (i.e., that the conviction probabilities for the two firms differ in case the other firm self-reports). Moreover, we assume that there will be a shock to cartel profits (or the respective detection probability) *after* the cartel has been formed. Hence, firms are more uncertain about cartel profits when deciding upon cartel formation than when deciding upon self-reporting.

The resolution of uncertainty leads to an option value of self-reporting as cartels will only be self-reported if profits turn out to be ("surprisingly") small. This option value leads to the following trade-off for the: on the

---

[1] For instance, the applying firm must not be the leader or initiator of the cartel and cartel involvement must be terminated instantaneously. For details see European Community (2002) and US Department of Justice (2004).

[2] The main differences between the *new* European program and the US program are (i) that no fine reduction for subsequent self-reporters is granted under the US program while fine reductions between 20 and 50% apply under the EU program, and (ii) that no fine reduction is granted in the US if investigators have already collected sufficient evidence while a reduction between 20 and 50% is still possible in Europe. We discuss our findings with respect to these institutional differences in a concluding section.

one hand, high fine reductions set high incentives to self-report, thereby terminating the socially detrimental cartel. On the other hand, higher fine reductions enhance the option value of self-reporting and thereby increase the expected profit from forming a cartel in the first place. Analyzing this trade-off, we find the following results:

First, and most important, it is never optimal to grant full immunity to a self-reporter who can only provide little evidence on other firms. It turns out that the high-evidence provider is pivotal for cartel continuation since his incentive to self-report will be higher in the authority's optimal policy. It follows that the authority can choose a higher fine for the low-evidence provider without affecting the overall probability of self-reporting. And since this higher fine increases the expected aggregated fines for the whole cartel, deterrence will be higher than with full amnesty but the number of self-reported cartels will not be reduced. The robustness of this finding lends some support to the old program in the EU.

Second, the optimal self-reporting fine for the firm which provides low evidence can be characterized very easily - the fine should equal the expected fine the firm faces in case the other firm self-reports. Hence, the fine for one firm depends on the evidence provided by the other firm.

Third, we investigate whether a full amnesty should be granted for the firm providing the highest amount of evidence. With respect to this, our results are more subtle as they depend on the probability distribution over the shocks for cartel profits. Intuitively, the larger the number of cartels with small (negative) profit shocks (i.e., the larger the number of additional self-reporters in case of a full amnesty) and the smaller the number of cartels with large (negative) profit shocks (i.e., the smaller the number of overall self-reporters), the more likely a full amnesty for the high evidence provider is optimal. The larger the number of additional self-reporters, the more cartels become instable due to a full for the high-evidence provider. The lower the number of overall self-reporters, the less important is the pro-collusive impact of a lower fine for those who self-report. We show that a full amnesty for the high evidence provider is optimal if the cumulative distribution function of shocks to cartel profits is convex.

Summing up, our model allows for a characterization of optimal fine reductions and suggests their dependence on the amount of evidence provided by the firms.

The basic idea that rewarding self-reporters in criminal teams can increase deterrence is a simple application of the prisoners' dilemma and has been

3

confirmed in many different contexts. Kofman and Lawarrée (1996) analyze double auditing for financial statements and show that high rewards if exactly one auditor is honest eliminate the incentives to collude with the auditee. Garoupa (1999) assumes that arrested members of a criminal team can compromise the whole organization and shows when fine reductions are welfare improving. Feess and Walzl (2004) assume that criminal teams may either act in concert or non-cooperatively when deciding about self-reporting and find that the optimal fine reduction does not depend on the probability that team members behave cooperatively as long as this probability is exogenous.

In recent years, a growing literature on the formal analysis of corporate leniency programs has emerged.[3] As in our framework, Motta and Polo (2003) consider an infinitely repeated collusion game between firms. Collusion may break down since partners may cheat on each other, for instance by setting lower prices than agreed upon. Self-reporting schemes may then even enhance the stability of cartels, because low fines provide credible threats to disclose the cartel in case the partner cheats. However, they also show that the *optimal* leniency programs (weakly) improves social welfare.

Spagnolo (2004) also assumes that partners can cheat on each other and shows that a "courageous" leniency program that rewards the first reporting party with the fines paid by all other parties achieves a first-best.[4] In contrast, any "moderate" program that reduces or cancels fines runs the risk to provide credible threats for cartel members to unravel the deal in case the accomplice cheats. In a similar spirit, Buccirossi and Spagnolo (2005) show that rewarding whistle-blowers and allowing for imprisonment reduces the maximum sanction required to trigger self-reporting.

Whereas Motta and Polo (2003) and Spagnolo (2003) thus focus on environments where team members may cheat on each other, other papers assume that the members of a cartel either stick to the collusion or report to the agency. Harrington (2005a) characterizes optimal leniency programs when cartels have different detection probabilities learned after an investigation has started. As in our paper, the up-date of the team-specific detection probability creates an option value of self-reporting that reduces the deterrence effect of leniency programs. Since confession by one participant is always

---

[3]For an extensive overview on the theoretical literature see Spagnolo (2006).

[4]In the US, already convicted firms may in fact be subsidized for confessing cartels in other markets that have not been detected yet (called "amnesty plus"). See e.g. McElwee (2004).

sufficient to convict all members of the cartel, our question of fine reductions contingent on the amount of evidence provided does not arise in this context. Aubert et al. (2005) investigate the agency problem associated with a whistleblowing program where employees are rewarded if they transmit evidence about their employer's cartel activity. Moreover, they regard hard evidence as a threat to other cartel members and use this to explain why firms do not destroy hard evidence for cartel activity. The amount of evidence that can be provided in case of detection or self-reporting is therefore endogenous. However, since all firms are identical with respect to their costs and revenues, they all hold the same amount of evidence in equilibrium. Our assumption of firms which differ with respect to the amount of evidence they keep could easily be derived in their model by allowing for heterogenous firms (i.e., different production costs or horizontal differentiation).

As usual in the literature on leniency programs, we assume that he probability that a cartel is detected depends only on the authority's effort.[5] In reality, however, the detection probability will also depend on the mark-up agreed upon, and more generally speaking on the dynamics of the price path. This has been considered in Harrington (2004), Harrington (2005b) and in Harrington and Chen (2005). Whereas the two former paper operate with (reasonable) ad-hoc assumptions concerning the influence of the price path on the detection probability, the latter explicitly formalizes the believes of industrial buyers who may inform cartel authorities that something is going wrong, and analyzes the cartel's reaction to these belief formations.

Hinloopen (2006) assumes that the authority can divide its limited resources between different cartels and finds that all industries should at least face positive detection probability. Hence, bundling resources for industries that are more likely to form cartels does not pay off. Festerling (2005) finds that leniency programs on the corporate level are likely to be more effective in deterring cartel formation compared to leniency programs for managers.

Finally, a few empirical and experimental papers deal with the consequences of corporate leniency programs. Connor (2007) demonstrates that the duration of collusion has only a small impact on the fine reduction actually granted by the Department of Justice in the USA and argues that this reduces the positive welfare effects of the program. Direct measurements of

---

[5]Here, we assume that this probability is exogenously given. In an older version of this paper, we have also analyzed a model where this probability is endogenous, which did not offer additional insights.

the deterrence effect of leniency programs, however, is impossible since information about non-detected cartels is not available. Miller (2007) tries to overcome this problem by assuming that discovered cartels are representative for all cartels, and then finds that the US leniency program has lead to a reduction in cartel formation of about 50%. Apesteguia, Dufwenberg and Selten (2006) and Hinloopen and Soetevent (2006) have carried out experimental studies. Their approaches differ with respect to the number of cartel members and in the communication structure, but the bottom line of both studies is a price reduction due to the introduction of leniency programs - lending some support to the most robust result of the various theoretical contributions.

The remainder of our paper is organized as follows. In Section 2, we present the model and the timing of the game. We analyze the game by backwards induction in Section 3. In particular, Section 3.4 contains a characterization of the authority's optimal policy. We conclude with some remarks on the robustness of our findings and a respective critique of current policy designs in Section 4.

2 The Model

In our model, there are two profit maximizing risk-neutral firms and an authority minimizing the expected social loss from cartel formation. The game has infinite horizon and firms discount time with discount factor δ. In each period, a cartel is detected with probability p_G, and the maximum fine on a convicted cartel member is s. The authority's choice parameter is the fine reduction granted for a self-reporting firm. The fine reduction may depend on the amount of evidence provided in case of self-reporting (see below).

Firms are ex ante symmetric and have a private benefit $B \in [0, \overline{B}]$ from cartel behavior in each period. B is distributed with density $f(B)$ and has full support. For simplicity, we assume that the benefits within a cartel are perfectly correlated and hence identical.[6] To avoid trivial results we assume $\overline{B} > p_G s$ as cartels would otherwise never be formed.

After firms have formed a cartel, a shock to cartel profits $\theta \in [0, 1]$ distributed with density $g(\theta)$ occurs, leading to cartel profits of θB in each

---

[6]Assuming independently distributed benefits is more tedious but has no qualitative impact on our results.

subsequent period. This expresses the realistic assumption that firms do not perfectly know their benefit from cartel behavior in advance. In particular, the assumption ensures the existence of cartels that subsequently self-report. [7]

Firms differ with respect to the amount of evidence they can provide in case of self-reporting, and we assume that this amount is also only learned after cartel formation. Specifically, we assume that there is a firm H holding a large amount of evidence and a firm L holding a small amount of evidence. If firm $F \in \{L, H\}$ self-reports, the other cartel member will be convicted with probability

$$p_{CF} = p_G + p_F - p_G p_F, \ F = H, L. \tag{1}$$

with $1 > p_H > p_L > 0$. p_F denotes the probability with which the other firm will be convicted on the basis of firm F's report and p_{CF} is the overall probability of conviction if firm F self-reports. Hence, the conviction probability is always higher if the partner self-reports, but firm L is more likely to be convicted if firm H self-reports (because $p_H > p_L$). Before cartel formation, each firm expects to be each of the two types with probability $1/2$, so that the two firms are indeed identical ex ante.[8]

To allow fines to depend on evidence provision we denote the self-reporting fine for firm F as

$$r_F = \lambda(1 - \mu p_F)s \tag{2}$$

where $\lambda \in [0, 1]$[9] depicts the overall size of the fine reduction, and $\mu \in [0, 1]$ captures the impact of evidence. $\lambda = 0$ expresses an amnesty, and $\mu = 0$ are fine reductions independent of the amount of evidence provided. Thus, the lower λ the higher is ceteris paribus the fine reduction, and the higher μ, the more important is evidence provision. Note that the definition of r_F implies that the fine is non-increasing in the amount of evidence provided ($r_H \leq r_L$)

---

[7]Of course, the uncertainty could also refer to other aspects of the cartel like the detection probability or preferences for not cheating on a "partner". Restricting attention to negative shocks is thereby without loss of generality.

[8]To see that this is without loss of generality, assume that firms know their type ex ante. Then, the incentive of the L-firm to form a cartel is lower as it can more easily be convicted if the partner self-reports. With side payments, nothing would change. Without side payments, firm L will be pivotal for cartel formation. Qualitatively, our results regarding optimal evidence dependence of fines will stay the same.

[9]Of course, the model could be extended to "courageous" leniency programs by admitting $\lambda < 0$. However, in this paper we aim at a discussion of existing schemes, which all exhibit $\lambda \geq 0$.

which holds in all leniency programs we are aware of, and which would also occur endogenously in our model.

If both firms decide to self-report, each of them wins the "race to the courtroom" with probability $1/2$ and pays r_F whereas the other firm pays s if it is convicted (i.e., with probability p_{CF}). Modelling a simultaneous self-reporting decision as a race to the courtroom is often realistic, but our results would not change qualitatively if we assumed instead that a second self-reporter gets also a fine reduction.

The timing of the game can now be summarized as follows:

1. Authority commits to policy $\Omega \equiv (\lambda, \mu)$.

2. Firms learn B and decide upon forming a cartel. If not, the game ends.

3. If a cartel is formed, firms learn θ and their identity (H or L).

4. Firms decide upon self-reporting and cartel continuation. The game ends after self-reporting, after conviction without self-reporting, and when firms decide not to continue the cartel. Otherwise Stage 4 repeats itself.[10]

Note that all periods after firms have learned θ in stage 3 are identical such that self-reporting or cartel termination will occur only in the first period after stage 3 or never. We will specify the social welfare function when we turn to the authority's optimal policy in stage 1. In the following we solve the game via backwards induction for a subgame perfect equilibrium.

3 Equilibrium Analysis

3.1 Cartel Continuation

Let us first consider the continuation payoff in case firms do not self-report in a given period. In the next round, each firm earns θB and has expected fine

---

[10]Hence, we exclude the case that cartel activity continuous after (unsuccessful) self-reporting or after conviction. This assumption seems reasonable as the risk of being detected and convicted after the case has come to the authority's attention is too high. Relaxing this assumption would clearly reduce the overall social benefit of the leniency program, but leaves the structure of optimal fine reductions unchanged.

$p_G s$. Since the probability to reach the next period is $1 - p_G$, the expected discounted pay-off stream of each firm $F \in \{H, L\}$ is

$$\pi_F = (1-p_G)\delta(\theta B - p_G s) + (1-p_G)^2 \delta^2(\theta B - p_G s) + \ldots = \frac{\delta(1 - p_G)(\theta B - p_G s)}{1 - \delta(1 - p_G)}.$$

π_F is identical for firm H and L as it captures the case without self-reporting. Continuation is preferred to termination if and only if $\pi_F > 0$, i.e. if $\theta > \bar{\theta} \equiv \frac{p_G s}{B}$.

3.2 Self-Reporting

The following table summarizes the payoff matrix in case one or two firms self-report (row player is firm H and column-player firm L; S denotes the action "self-report" and N "no self-report").

H/L	S	N
S	$-\frac{1}{2}(r_H + p_{CL}s) /$ $-\frac{1}{2}(r_L + p_{CH}s)$	$-r_H/ - p_{CH}s$
N	$-p_{CL}s/ - r_L$	$\theta B - p_G s + \max(\pi_F, 0)/$ $\theta B - p_G s + \max(\pi_F, 0)$

Table 1: Payoffs with and without self-reporting

If no firm self-reports, it receives the benefits θB and is convicted with probability p_G. Afterwards, each firm gets the maximum from termination (i.e. zero profit) and continuation (i.e., π_F)– $\max(\pi_F, 0)$. If only one firm self-reports, it pays the reduced fine r_F whereas the respective co-cartelist is convicted with p_{CF} and pays s in this case. If both firms want to be the first one to come forward, each of them wins the race to the courtroom with probability 1/2 which explains the payoffs in case both firms self-report. Inspecting Table 1 yields the following insights:

Observation 1. *1. If $\theta \geq \bar{\theta} \equiv \frac{p_G s}{B}$, then $\mathbf{N} \equiv (N, N)$ is a pure strategy Nash-equilibrium of stage 3(PNE) in which firms continue the cartel. If $\theta < \bar{\theta}$ and $\theta \geq \underline{\theta} \equiv \frac{p_G s - r_H}{B}$, $\mathbf{N}$ is a PNE, but firms terminate the cartel after this period because they have learned that θ is too low. As $r_H \leq r_L$, firm H has a (weakly) lower incentive to stick to the $\mathbf{N}$-equilibrium. It receives a (weakly) lower fine when it self-reports. In other words, firm H is pivotal as for $\mathbf{N}$ being a PNE.*

2. (S, N) *or* (N, S) *can never be PNE as a necessary condition for an asymmetric equilibrium is that the inequalities* $r_H > p_{CL}s$ *and* $r_H < p_G s - \theta B$ *hold simultaneously. This, however, is impossible as* $\theta \geq 0$ *and* $p_{CL} \geq p_G$. *Hence, if a firm anticipates that the partner self-reports, it will always try to win the race to the courtroom.*

3. $\mathbf{S} \equiv (S, S)$ *is a PNE whenever* $r_H \leq p_{CL}s$ *and* $r_L \leq p_{CH}s$ *hold simultaneously.*

From Observations (1) and (3), it follows that $\mathbf{N}$ and $\mathbf{S}$ may coexist as PNE. Intuitively, this will be the case if the self-reporting fines are sufficiently low (this makes $\mathbf{S}$ a PNE) and when the cartel benefit θB is sufficiently high (this makes $\mathbf{N}$ a PNE). To select among the (possibly) coexisting pure strategy Nash-equilibria $\mathbf{S}$ and $\mathbf{N}$, we apply pay-off dominance and find

Lemma 1. *Suppose firms have formed a cartel. Then,* $\mathbf{S}$ *is the pay-off dominant PNE if and only if* $r_H < p_G s$ *and* $\theta < \underline{\theta}$. *Otherwise,* $\mathbf{N}$ *is the pay-off dominant PNE.*

Proof. All proofs can be found in the Appendix. $\square$

Comparing Lemma 1 to Observation 1 about the existence of a PNE without self-reporting shows that the $\mathbf{N}$-equilibrium is payoff dominant whenever it exists. The intuition is that $\mathbf{N}$ can only be an equilibrium if even the high evidence provider (firm H) does not benefit from being the *only* self-reporter. But since being the only self-reporter leads to a higher payoff than the one in the $\mathbf{S}$-equilibrium, $\mathbf{N}$ is preferred to $\mathbf{S}$ if it exists. In what follows, we assume that firms play the pay-off dominant equilibrium. Hence the authority can induce self-reporting only by destroying the $\mathbf{N}$ as an equilibrium.

3.3 Cartel Formation

Given the termination-, continuation- and self-reporting-decisions as derived above, we can write the expected payoff of firm F when *having formed* a cartel as[11]

$$\Pi_F = -\int_0^{\underline{\theta}} dG(\theta) 1/2(r_F + p_{C-F}s) + \int_{\underline{\theta}}^{\bar{\theta}} dG(\theta)(\theta B - p_G s) + \int_{\bar{\theta}}^1 dG(\theta) \frac{\theta B - p_G s}{1 - \delta(1 - p_G)} \tag{3}$$

---

[11] p_{C-F} denotes the probability that firm F is convicted when the other firm (expressed by "$-F$") self-reports.

For $\theta < \underline{\theta}$, **N** is no equilibrium such that we get a race to the courtroom. This explains the first term. For $\underline{\theta} \leq \theta < \overline{\theta}$, **N** is a (payoff dominant) equilibrium and hence played, but the cartel is terminated after the shock. Finally, for $\theta \geq \overline{\theta}$, the cartel lasts unless it is detected.

Since firms assume to be firm H and firm L with equal probabilities, the expected payoff *before* cartel formation is

$$\overline{\Pi} = \frac{1}{2}\left(\Pi_H + \Pi_L\right). \tag{4}$$

Note that the cartel will be formed iff $\overline{\Pi} \geq 0$. For any given policy Ω of the authority, $\overline{\Pi}$ depends on B. For further reference, let us define $\widetilde{B}$ as the critical B where the cartel will only just be formed, i.e. $\overline{\Pi}\left(\Omega, \widetilde{B}\right) = 0$.

3.4   The Authority's Optimal Policy

To derive the authority's objective function, we need to specify the social loss from cartel formation. For each cartel period, there is a reduction in the sum of producers' and consumers' surplus denoted τ. Furthermore, let us define $\gamma(\lambda, \mu)$ as the probability that a cartel is formed depending on the authority's policy $\Omega = (\lambda, \mu)$. Making use of $\underline{\theta}$ and $\overline{\theta}$, it follows immediately that expected social losses from cartels are

$$
\begin{aligned}
SC &= \gamma(\lambda, \mu)\left[\int_0^{\underline{\theta}} dG(\theta)\tau + \int_{\underline{\theta}}^{\overline{\theta}} dG(\theta)\tau + \int_{\overline{\theta}}^1 dG(\theta)\frac{\tau}{1 - \delta(1 - p_G)}\right] \\
&= \gamma(\lambda, \mu)\left(\tau + \int_{\overline{\theta}}^1 dG(\theta)\frac{\tau\delta(1 - p_G)}{1 - \delta(1 - p_G)}\right). \tag{5}
\end{aligned}
$$

If a cartel is formed, it is self-reported with probability $\int_0^{\underline{\theta}} dG(\theta)$, such that the social loss τ occurs only once. With probability $\int_{\underline{\theta}}^{\overline{\theta}} dG(\theta)$, the cartel is not self-reported but terminated, and again τ happens only once. Finally, with probability $\int_{\overline{\theta}}^1 dG(\theta)$, the cartel lasts until it is detected.

Note that neglecting the authority's effort costs in the social cost function is without loss of generality. It is straightforward to assume that there are some costs $C(p_G)$ arising in each period, and we could also allow for lower effort costs after a cartel has been detected once. In a working paper, we

have analyzed an extended version where p_G is endogenous, which did not show any qualitative impact on our results.

The Authority minimizes SC by choosing λ and μ. We find[12]

Proposition 1. *(i) The optimal self-reporting fine is always positive for at least one firm: $\lambda^* > 0$.*
(ii) It is always optimal to differentiate between the two firms: $\mu^* > 0$.
(iii) The self-reporting fine for the high evidence provider is always below the expected fine without self-reporting: $r_H^* < p_G s$.
(iv) The self-reporting fine for the low evidence provider is equal to the expected fine when the high evidence provider self-reports: $r_L^* = p_{CH} s$.
(v) The self-reporting fine for the high evidence provider is positive $(r_H^* > 0)$ if $G(\theta) > g(\theta) \frac{(p_{CL} + 2p_{CH})s}{B}$ for $\theta = \frac{p_G s}{B}$.
(vi) The self-reporting fine for the high evidence provider is zero $(r_H^* = 0)$ if $G(\theta)$ is strictly convex for all $\theta \in \left[0, \frac{p_G s}{B}\right]$.

From a policy perspective, part (i) may be considered as our main result: it says that it is never optimal to grant full immunity regardless of the amount of evidence provided. The reason is that the low evidence provider should always pay a positive fine. To see this, recall that the high-evidence provider is pivotal for the existence of an equilibrium without self-reporting. It follows that, for *any* fine for the high-evidence provider (and hence also for $r_H^* = 0$), one can choose a higher fine for the low-evidence provider *without affecting the overall probability of self-reporting.* And since this higher fine for firm L increases the overall expected fine (and thereby deterrence), the result follows. Part (ii) follows exactly the same logic.

Given parts (i) and (ii) of the Proposition, it is easy to make more precise statements about the self-reporting fines for the two types. Part (iii) says that it is always optimal to reduce the high-evidence provider's fine below the punishment he would expect without self-reporting. This resembles well-known results from the self-reporting literature and says that is optimal to give incentives to come forward. Technically speaking, the reason is that the marginal number of self-reporters is non-zero (due to the full-support assumption of $g(\theta)$) while the number of self-reporters obviously vanishes at $\underline{\theta} = 0$. Hence, a marginal fine reduction below $p_G s$ induces self-reporting and thereby higher expected fines in the self-reporting equilibrium **S** but does not reduce deterrence due to lower fines for self-reporters.

---

[12]Subgame perfect policy choices by the authority are denoted by $()^*$.

More interestingly, part (iv) says that the self-reporting fine of the low-evidence provider should equal the expected fine he faces when his accomplice self-reports. The only reason to reward the low-evidence provider at all is to avoid destruction of the full self-reporting equilibrium **S**. Hence, firm L pays the same expected fine as without self-reporting (given that the high evidence provider self-reports). Summing up, our results (i)-(iv) show that the amount of evidence provided should play an important rule for the fine of the first self-reporter (the winner of the race to the courtroom), and that guaranteeing a fine of zero is inefficient.

Parts (v) and (vi) analyze whether full immunity should be granted for the high-evidence provider. Here, the intuition is less straightforward than for parts (i) to (iv). As we present in more detail in the proof of Proposition 1 (see Appendix), the impact of a fine increase on social costs can be divided into two contributions. $\frac{\partial SC}{\partial r_F} > 0$ displays marginal social benefits of a fine increase – higher fines for those who self-report reduce cartel profits and thereby enhance deterrence. In contrast, $\frac{dSC}{d\theta} \frac{d\theta}{dr_H} < 0$ represents marginal social costs of a fine increase. For higher fines, firm H is less tempted to self-report (and thereby to destroy **N** as an equilibrium) which leads to higher cartel stability and thereby to higher expected cartel profits and reduced deterrence. Marginal benefits are proportional to $G(\underline{\theta})$ (the number of self-reporters), while marginal costs are proportional to $g(\underline{\theta})$ (the marginal number of self-reporters). Part (v) shows that positive fines for firm H are optimal if the number of self-reporters in case of a full amnesty (i.e. at $\theta = \frac{p_G s}{B}$) is sufficiently large. Intuitively, the higher the number of self-reporters, the more important is a higher fine for those who self-report compared to the respective increase in cartel stability. Analogously, part (vi) shows that a full amnesty for the high-evidence provider is optimal if the number of self-reporters for full immunity is sufficiently low. This is the case, for instance, if $G(\theta)$ is convex such that shocks are likely to be mild and the option value of self-reporting is small.

4   Conclusion

One of the features in which (past and present) corporate leniency programs differ is whether the fine reduction depends on the amount of evidence provided or whether full immunity is granted regardless of this evidence. We have found that it is *never* optimal to grant full immunity to firms providing

low evidence on other cartel members. Our result seems to be robust as it is based on the insight that the high evidence provider will be pivotal for whether cartel continuation is an equilibrium or not. It follows that fines for low evidence providers can be positive at no cost since these fines have no impact on the overall probability of self-reporting as long at hey are (weakly) below the expected fine in case the other firm self-reports. Self-reporting fines for low-evidence providers should hence be based on the regulator's estimation of the evidence other firms have to offer. Out result lends considerable support to the old system in the European Union and hence argues that following the US-procedure with respect to the independence of the amount of evidence creates social costs not yet recognized in the literature. Whether these additional costs of a leniency program with full amnesty are balanced by the recognized benefits with respect to transparency and predictability of fine reduction is beyond the scope of our paper but should be analyzed on the basis of the history of cartel cases.

Even though we restrict our model to the case of two firms, our analysis should have convinced the reader that it is straightforward to generalize our findings to larger cartels. All that matters is the existence of a "highest evidence provider" who is pivotal for the destruction of an equilibrium in which no firm self-reports. Keeping the fine reduction for this firm fixed, the authority benefits from higher fines for firms with lower evidence while the number of self-reported cartels remains unaltered. This leads again to no fine reduction beyond expected fines without self-reporting for all but the highest evidence provider. The optimal fine reduction for the latter firm, however, again depends on the distribution of cartel shocks.

In our analysis, we have confined attention to one element of corporate leniency programs, the impact of the evidence provided. Other important elements include (i) whether fine reductions should also be granted for subsequent self-reporters which is still possible under the new EU program but not under the US program, and (ii) whether fine reductions are sensible after an investigation has started and when the authority has already collected substantial evidence. Again, these fine reductions are feasible under the EU program, but not under the US legislation.

To analyze these questions, we have developed an extended model which allow for self-reporting before and after an investigation has started, and where firms can self-report subsequently (see Feess and Walzl (2003)). Although this leads inevitably to a more complex model, some clear-cut results can be derived: First, we show that fine reductions should also be granted to second

self-reporters if conviction requires some investigation effort. Second, in the conviction stage (i.e. after the cartel has already been detected), both fines should not depend on the own evidence provided, but only on the probability of being convicted by the partner's testimony. Third, the fine for the low-evidence provider should be the same in both self-reporting stages, i.e. independent of whether the case has already been detected or not.

References

Apesteguia, J, M. Dufwenberg and R. Selten (2006). Blowing the whistle, Economic Theory 31, 133-166.

Aubert, C., P. Rey, and W. Kovacic (2005). The Impact of Leniency and Whistleblowing Programs on Cartels, International Journal of Industrial Organization, 24, 1241-1266.

Connor, J.M. (2007). A critique of cartel fine discounting by the U.S. Department of Justice, Purdue University.

European Union (2002), Commission notice on immunity from fines and reduction of fines in cartel cases, Official Journal of the European Communities, 2002/C 45/03.

Feess, E and M. Walzl (2004). Self-reporting in optimal law enforcement when there are criminal teams, Economica 71, 333-348.

Feess, E. and M. Walzl (2003). Corporate Leniency Programs in the USA and the EU, German Working Papers in Law and Economics 2003-1-1077, Berkeley Electronic Press .

Hammond, S.(2000). Detecting and Deterring Cartel Activity through an effective Leniency Program, Speach of the Director of Criminal Enforcement, US Departement of Justice, presented at the International Workshop on Cartels, Brighton, November 2000.

Hinloopen, J. (2006): Internal cartel stability with time-dependent detection probabilities, International Journal of Industrial Organization, 24, 1213-1229.

Hinloopen, J. and A.R. Soetevent (2006): Trust and recidivism; the partial success of corporate leniency programs in the laboratory, University of Amsterdam.

Harrington, Jr., J.E. (2004). Post cartel pricing dynamics in the presence of an antitrust athority, RAND Journal of Economics 35, 651-673.

Harrington, Jr., J.E. (2005a). Optimal cartel pricing in the presence of an antitrust authority, International Economic Review 46, 145-169.

Harrington, Jr., J.E. (2005b). Optimal corporate leniency programs, John Hopkins University working paper No. 527.

Harrington, Jr., J.E. and J. Chen (2006). Cartel pricing dynamics with cost variability and endogenous buyer detection, International Journal of Industrial Organization 24, 1185-1212.

Levenstein, M. and V. Suslow (2002). What Determines Cartel Success, University of Michigan Working Paper.

McElwee, D. (2004). Should the European Commission Adopt "Amnesty Plus" in its Fight Against Hard-Core Cartels?, European Competition Law Review, 25, 558-564.

Miller, N.H. (2007). Strategic leniency and cartel enforcement, University of California, Berkeley.

Motta, M. and M. Polo. (2003). Leniency Programs and Cartel Prosecution, International Journal of Industrial Organization 21, 347-379.

Motchenkova, E. (2004). Effects of Leniency Programs on Cartel Stability. CENTER discussion paper No. 2004-98.

Kaplow, L., and S. Shavell (1994). Optimal law enforcement with self-reporting of behavior. Journal of Political Economy 102(3), 583-605.

Livernois, J., and C.J. McKenna (1999). Truth or consequences: Enforcing pollution standards with self-reporting. Journal of Public Economics 71, 415-440.

Spagnolo, G. (2003). Divide et Impera: Optimal Leniency Programmes CEPR discussion paper No. 4840.

Spagnolo, G. (2006). Leniency and Whistleblowers in Antitrust Economics, in P. Buccirossi (Ed.), Handbook of Antitrust Economics.

US Department of Justice (2004), Official Giudelines
(http://www.usdoj.gov/atr/public/guidelines/0091.htm).

A Proofs

Proof of Lemma 1

1. Suppose either $r_H > p_{CL}s$ or $r_L > p_{CH}s$. Then, **N** is the pay-off dominant PNE - in fact it is the unique PNE of stage 3.

2. Suppose that $\theta B > p_G s$. Then, **N** is the pay-off dominant PNE.

3. Suppose that $\theta B < p_G s$, $r_H < p_{CL}s$, $r_L < p_{CH}s$, and $\theta < \underline{\theta}$. Then, **S** is the unique PNE. Recall that $\underline{\theta} = 0$ for $r_H \geq p_G s$. Hence, it is equivalent to say that **S** is the unique PNE if $\theta B < p_G s$, $r_H < p_G s$ and $\theta < \underline{\theta}$.

4. Suppose that $\theta B < p_G s$, $r_H < p_{CL}s$, $r_L < p_{CH}s$, and $\theta > \underline{\theta}$. Then, both **N** and **S** are PNE. **N**, however, is payoff-dominant if $-1/2(r_{H(L)} + p_{CL(H)}s) < \theta B - p_G s$ or $\theta > \widetilde{\theta} \equiv \max(\frac{p_G s - 1/2(r_H + p_{CL}s)}{B}, \frac{p_G s - 1/2(r_L + p_{CH}s)}{B})$. But as $r_H < p_{CL}s$, $r_L < p_{CH}s$, and $r_L \geq r_H$, it follows that $\widetilde{\theta} < \underline{\theta}$. Therefore, if **N** is indeed a PNE (next to **S**) (a necessary condition being $\theta > \underline{\theta}$) it is also payoff dominant (a sufficient condition being $\theta > \widetilde{\theta}$.

Proof of Proposition 1 The authority chooses $\omega = (\lambda, \mu)$ as to minimize social costs SC (see Eqn. 5). Now observe that $\frac{dSC}{d\lambda} = \frac{d\gamma(\lambda,\mu)}{d\lambda}(\tau +$

$\int_{\underline{\theta}}^{1} dG(\theta) \frac{\tau\delta(1-p_G)}{1-\delta(1-p_G)}$). The term in brackets is strictly positive and independent of λ. Minimizing social costs is therefore equivalent to minimize cartel frequency $\gamma(\lambda, \mu)$, *i.e.,* the authorties objective is to minimize the firms expected benefits from forming a cartel $\overline{\overline{\Pi}}$.

Part (i). Suppose $\lambda^* = 0$ (and, hence, $r_L^* = r_H^* = 0$). Then, there is always a $\widetilde{\lambda} > 0$ and a $\widetilde{\mu} = 1/p_H$ such that $r_H = 0$ and $r_L > 0$. If the authority commits to $\widetilde{\lambda}$ and $\widetilde{\mu}$, it yields $r_H = r_H^*$ and $r_L > r^*$. Hence, $\underline{\theta} = \frac{p_G s - r_H}{B}$ (recall that player H is pivotal for $\mathbf{N}$ to be an equilibrium as he has a higher incentive to self-report because of $r_H < r_L$). Therefore, for $(\widetilde{\lambda}, \widetilde{\mu})$ the same fraction of cartels is self-reported as under (λ^*, μ^*) and player H faces the same expected fine in case of self-reporting (payoffs in $\mathbf{N}$ are not effected at all). However, player L faces a higher fine if he successfully self-reports ($r_L > r^*$) which happens with probability $1/2$ in $\mathbf{S}$ and therefore increases the total expected fine and thereby reduces $\overline{\overline{\Pi}}$. A contradiction to the optimality of (λ^*, μ^*).

Part(ii). By Part (i), $\lambda^* > 0$. Now suppose that $\mu^* = 0$ such that $r^* \equiv r_H^* = r_L^* = \lambda^* s > 0$. By construction there is always a $\widetilde{\lambda} > \lambda^*$ and a $\widetilde{\mu} > 0$ such that $\widetilde{\lambda}(1 - \widetilde{\mu} p_H)s = r^*$. If the authority commits to $\widetilde{\lambda}$ and $\widetilde{\mu}$, it yields $r_H = r^*$ and $r_L > r^*$ (recall that $p_L < p_H$). Hence, $\underline{\theta} = \frac{p_G s - r_H}{B}$ (recall that player H is pivotal for $\mathbf{N}$ to be an equilibrium as he has a higher incentive to self-report because of $r_H < r_L$). Therefore, for $(\widetilde{\lambda}, \widetilde{\mu})$ the same fraction of cartels is self-reported as under (λ^*, μ^*) and player H faces the same expected fine in case of self-reporting (payoffs in $\mathbf{N}$ are not effected at all). However, player L faces a higher fine if he successfully self-reports ($r_L > r^*$) which happens with probability $1/2$ in $\mathbf{S}$ and therefore increases the total expected fine and thereby reduces $\overline{\overline{\Pi}}$. A contradiction to the optimality of (λ^*, μ^*).

Part (iii). $\frac{d\overline{\overline{\Pi}}}{dr_H} = \frac{\partial\overline{\overline{\Pi}}}{\partial r_H} + \frac{\partial\overline{\overline{\Pi}}}{\partial\underline{\theta}}\frac{d\underline{\theta}}{dr_H}$ with $\frac{\partial\overline{\overline{\Pi}}}{\partial r_H} = -\frac{1}{4}G(\underline{\theta})$, $\frac{d\underline{\theta}}{dr_H} = -\frac{1}{B}$, and $\frac{\partial\overline{\overline{\Pi}}}{\partial\underline{\theta}} = \frac{1}{4}g(\underline{\theta})(-r_L + 3r_H - p_{CL}s - p_{CH}s)$. Therefore, $\frac{d\overline{\overline{\Pi}}}{dr_H}|_{r_H = p_G s} > 0$ (recall that $\underline{\theta}|_{r_H = p_G s} = 0$).

Part(iv). By Part (iii), $r_H^* < p_G s$. Then, $G(\underline{\theta}) > 0$ and $\frac{d\overline{\overline{\Pi}}}{dr_L} = -\frac{1}{4}G(\underline{\theta}) < 0$ for all $r_L \in (0, p_{CH}s)$.

Part(v). With

$$\frac{d\overline{\overline{\Pi}}}{dr_H} = -\frac{1}{4}G(\underline{\theta}) + \frac{1}{4}\frac{1}{B}g(\underline{\theta})(r_L - 3r_H + p_{CL}s + p_{CH}s) \tag{6}$$

(see Part (iii)) and $r_L^* = p_{CH}s$ (see Part (iv)), it follows by Part (iii) that the first term in Eqn.(6) is strictly negative while the second term is strictly

positive. If $G(\theta) > g(\theta)\frac{(p_{CL}+2p_{CH})s}{B}$ in $\theta = \frac{p_G s}{B}$ it follows that $\frac{d\overline{\Pi}}{dr_H}|_{r_H=0} < 0$ such that $r_H^* > 0$.

Part (vi). Observe that $\frac{1}{B}g(\underline{\theta})(p_{CL}s + 2p_{CH}s - 3r_H) > 3g(\underline{\theta})\underline{\theta}$ (recall that $p_{CF} > p_G$). If $G(\theta)$ is strictly convex (*i.e.*, $G(\theta) < \theta g(\theta)$)), this implies that $\frac{d\overline{\Pi}}{dr_H} > 0$ for all $r_H \in [0, p_G s]$.