Are Leniency Programs too Generous?

Julien SAUVAGNAT*

Abstract

I present a simple model of collusion in which the competition authority offers leniency rates contingent on the number of firms that report information. The optimal leniency policy involves what I refer to as a *single informant* rule - that is, leniency should be given only when a single firm reports information. The single informant rule allows to increase expected sanctions compared to the first informant rule, which overall improves cartel deterrence.

Keywords: Antitrust law and policy; Cartels; Leniency; Informant Rules.

JEL classification codes: K42, K21, L41.

*Corresponding address: Sauvagnat Julien, ENSAE, 3 Avenue Pierre Larousse, 92245 Malakoff, France. Phone: +33(0)607087049. E-mail: julien.sauvagnat@ensae.fr

1 Introduction

I investigate in a simple model of collusion whether the competition authority (CA) can improve welfare by offering leniency rates contingent on the number of firms that report information.

Since its revision in 1993, the corporate leniency program has been the US Department of Justice's most effective investigative tool.[1] The commonly used argument to explain its success is the implementation of first informant rules combined with large expected fines. This "stick and carrot" logic motivated the recent revisions (2002 and 2006) of the European leniency program which turns out to be also effective in fighting cartels.

Comparing the findings of the recent empirical literature confirms that the design of leniency schemes is a key element of success. Using information reports issued by the Department of Justice between 1985 and 2005, Miller (2009) shows that the pattern of cartel discoveries around the revision in 1993 of the US leniency program is consistent with enhanced cartel detection and deterrence capabilities. By contrast, Brenner (2009)'s work on European data shows that the introduction of a leniency policy in Europe in 1996 (without a first informant rule and full leniency) had no clear effect on deterrence.

The theoretical literature generally finds that restricting leniency to the first informant is strictly better that granting leniency to all informants (with few exceptions such as Motta and Polo, 2003). Spagnolo (2004) notes for instance that offering the first informant a reward equal to the sum of the fines imposed on the other conspirators can achieve full deterrence. Harrington (2008) shows that the first informant rule generates a "race to the courthouse effect" in case of investigation, which overall increases expected sanctions inflicted on cartel members.[2]

The novelty in this paper is to allow the CA to offer leniency rates contingent on the number of informants. The first informant rule is no longer optimal in that case. Instead, it is shown that the optimal policy involves what I refer to as a *single informant* rule –

[1]See e.g. Hammond (2010): *"These revisions made the program more transparent and raised the incentives for companies to report criminal activity and cooperate with the Antitrust Division. As a result of these changes, the Antitrust Division has seen a nearly twenty-fold increase in the leniency application rate, making the Leniency Program the Antitrust Division's most effective investigative tool."*

[2]See also Chen and Rey (forth.) and Sauvagnat (forth.).

that is, leniency should be given only when a single firm reports.

2 The Model

Firms. There is a continuum of industries with unit mass. In each industry, $N \in \{2, \ldots, \overline{N}\}$ symmetric firms play an infinitely repeated game and maximize the discounted sum of their profits using the same discount factor $\delta \in (0, 1)$. In each period, each firm decides whether to compete or collude. The gross profit of a firm is normalized to 0 if firms compete, equal to $\Pi > 0$ if firms collude, and equal to $\Pi^D \geq \Pi$ if the firm deviates from collusion – that is, if it competes when the other firm(s) collude.[3]

In order to analyze the impact of leniency on cartel formation, we follow Harrington and Chang (2009) and assume that industries are heterogenous with respect to the gains from deviating, Π^D, which are distributed according to a function G defined over the support $[\Pi; +\infty[$.

Competition Authority. The CA minimizes the social cost of collusion using the same discount factor δ as firms. Society incurs a per-period deadweight loss $L > 0$ when firms collude.

We assume that collusion leaves evidence that is needed to condemn the cartel. In each period, the CA opens an investigation with probability α.[4] The probability of finding evidence during an investigation is p, in which case the cartel is condemned and each member must pay a fine F. We assume that firms compete forever following a condemnation.[5]

Leniency. For the sake of exposition, we assume that a firm which deviates from collusion faces no risk of being convicted; this rules out any role for *pre-investigation* leniency.[6]

[3]For example, in a standard Bertrand oligopoly, static price competition drives profits to 0, the benefit from collusion is equal to a share $\frac{1}{N}$ of the monopoly profits ($\Pi = \frac{\Pi^M}{N}$) whereas a deviation brings the whole monopoly profits ($\Pi^D = \Pi^M$).

[4]α might represent the probability of receiving initial incriminating evidence from third parties such as internal employees, buyers or local agencies.

[5]Our results are qualitatively unchanged if we assume instead that, following a condemnation, firms compete only for a finite length of time. Enforcing competition can be achieved through either close monitoring of the industry or higher fines for repeat offenders.

[6]When deviating firms face a risk of being convicted, offering full amnesty in case of pre-investigation

The CA can however offer leniency during investigations, in which case each firm decides simultaneously whether or not to report. If (at least) one firm reports, the cartel is condemned with probability one.

We allow the leniency rate to be fully flexible with respect to the number of informants. A leniency policy is defined as a $\overline{N}$-tuple $Q = (q_1, \ldots, q_{\overline{N}})$ where $q_l \geq 0$ for $l \in \{1, \ldots, \overline{N}\}$. If the CA receives l leniency applications, each informant is eligible to the leniency rate q_l, and thus pays only a reduced fine equal to $(1 - q_l)F$. We rule out rewards, so that $q_l \leq 1$ for all $l \in \{1, \ldots, \overline{N}\}$.

Observe that adopting a first informant rule (hereafter, FI rule) boils down to offer the leniency policy $(q, \frac{q}{2}, \ldots, \frac{q}{N})$.[7]

Timing. At the beginning of the game, the CA announces the leniency policy $Q = (q_1, \ldots, q_{\overline{N}})$. Then, in each period: (1) firms decide whether to collude or not; (2) if firms decide to collude, each firm chooses whether to respect the agreement or to deviate and compete on the market, and the game proceeds to stage 3; (3) the CA opens an investigation with probability α, in which case each firm decides simultaneously whether or not to report. If at least one firm reports, the cartel is condemned with probability 1; otherwise, the CA finds evidence that allows to condemn the cartel with probability p.

3 Optimal Leniency and the Single Informant Rule

Following most papers in the literature, we consider two modes of (symmetric) collusive equilibrium enforced by trigger strategies,[8] in which firms either "collude and remain silent" or "collude and report in case of investigation",[9] and assume that firms coordinate on the most profitable equilibrium when both are sustainable. In order to be sustainable,

reports is optimal because it allows defecting cartel members to report and avoid paying the fine (which is referred to as the "protection from fines effect" in Spagnolo (2004) and the "deviator amnestor effect" in Harrington (2008)).

[7]The literature generally assumes that under a FI rule, if m firms simultaneously apply for leniency, each firm is equally likely to be the first informant and thus faces an expected fine equal to $(1 - \frac{q}{m})F$.

[8]That is, any deviation from collusion is punished by reverting forever to competition, which is here the minmax and thus constitutes the most severe punishment.

[9]This is without loss of generality. If the CA offers positive leniency, the equilibrium reporting strategies are necessarily symmetric since a firm is better off reporting whenever at least one other firm reports.

both collusive strategies must resist unilateral deviations on the market – i.e., the expected value of future collusion must exceed the gains from deviating on the market, Π^D. The strategy "collude and remain silent" must moreover be robust to unilateral reporting deviations: no firm should gain by reporting in case of investigation when the other firms remain silent.

Below, we first characterize firms' decisions given the CA's leniency policy; then, we solve for the optimal policy.

"Collude and report in case of investigation" (R). If firms report when the CA opens an investigation, each firm pays a reduced fine equal to $(1 - q_N)F$ and the cartel is condemned with probability 1. When condemned, firms compete forever. Therefore, the value of collusion, V^R, solves $V^R = \Pi - \alpha(1 - q_N)F + (1 - \alpha)\delta V^R$, that is:

$$V^R(q_N) = \frac{\Pi - \alpha(1 - q_N)F}{1 - \delta(1 - \alpha)}$$

"Collude and remain silent" (S). If firms remain silent when the CA opens an investigation, the probability of finding evidence of collusion is p, in which case firms pay F and compete forever. The value of collusion, V^S, is thus such that $V^S = \Pi - \alpha pF + (1 - \alpha p)\delta V^S$, that is:

$$V^S = \frac{\Pi - \alpha pF}{1 - \delta(1 - \alpha p)}$$

As already mentioned, S should resist unilateral reporting deviations – i.e., no firm should gain by reporting during an investigation when the other firms remain silent; in such a case, the reporting firm would pay only a reduced fine $(1 - q_1)F$. If instead all firms remain silent, with probability p, the investigation is successful and the cartel is condemned, and with probability $(1 - p)$, the investigation fails and firms' discounted continuation payoffs are δV^S. This gives the following "incentive-compatibility" (IC) constraint:[10]

[10] For the sake of exposition, we assume that a firm decides to report whenever it is indifferent between reporting and remaining silent.

$$-pF + (1-p)\delta V^S > -(1-q_1)F \tag{IC}$$

As already mentioned, when both collusive strategies are sustainable, we assume that firms coordinate themselves on the most profitable collusive equilibrium. Lemma 1 derives the firms' decisions as a function of q_1 and q_N.

Lemma 1. *There exists a threshold,* $\hat{q} \equiv (1-p)\frac{\delta\Pi+(1-\delta)F}{(1-\delta+\alpha\delta p)F}$ *such that:*

- *for* $\max(q_1, q_N) < \hat{q}$*, firms collude and remain silent if* $V^S \geq \Pi^D$*; otherwise, they compete.*
- *for* $\max(q_1, q_N) \geq \hat{q}$*, firms collude and report in case of investigation if* $V^R(q_N) \geq \Pi^D$*; otherwise, they compete.*

Proof. Firms choose S instead of R if S is both incentive-compatible and more profitable than R. After some computations, we obtain that i) the (IC) constraint is satisfied if and only if $q_1 < \hat{q}$ with $\hat{q} = (1-p)\frac{\delta\Pi+(1-\delta)F}{(1-\delta+\alpha\delta p)F}$; and ii) $V^S > V^R$ is satisfied if and only if $q_N < \hat{q}$. □

The threshold $\hat{q}$ is lower than 1 only if p is higher than a threshold $\underline{p} < 1$. As we rule out rewards, it is impossible to trigger reporting when $p < \underline{p}$. For the rest of the analysis, we assume $p \geq \overline{p}$.

To trigger reporting, the CA can either i) offer a large reduction in the fine, q_1, for a single informant - in that case, S is not incentive-compatible and cartel members are forced to play R; or ii) offer a large reduction in the fine, q_N, when all firms report information – in that case, cartel members play R because it is more profitable than S.

Note that the choice between offering a large q_1 or a large q_N has very different effects in terms of deterrence. Increasing q_N raises the value of R and thus dilutes deterrence. By contrast, increasing q_1 has no negative effect *in equilibrium* on expected sanctions. Intuitively, this explains why the CA will use q_1, rather than q_N, as an instrument to trigger reporting.

To see this formally, let us compute the total social costs of collusion, which are equal to the proportion of collusive industries in the economy multiplied by the social cost of a cartel. As the gains from deviating, Π_D, are distributed across industries according to the function G, $G(V)$ denotes the initial proportion of collusive industries.

When firms collude, society incurs the loss L. The CA opens an investigation with probability α, which allows to condemn the cartel with probability p if firms remain silent, and with probability 1 if firms report. When condemned, firms compete forever; otherwise, society incurs again the loss L in the following period. The social cost of a cartel, denoted C, thus satisfies $C^S = L + \delta(1-\alpha p)C^S$ when firms play S and $C^R = L + \delta(1-\alpha)C^R$ when firms play R.

It follows from Lemma 1 that the total social costs of collusion are equal to $G(V^S)\frac{L}{1-\delta(1-\alpha p)}$ for $\max(q_1, q_N) < \hat{q}$, and equal to $G(V^R(q_N))\frac{L}{1-\delta(1-\alpha)}$ for $\max(q_1, q_N) \geq \hat{q}$.

We are now able to characterize the optimal leniency policy, which is summarized in the following proposition:

Proposition 1. *When the CA offers leniency rates contingent on the number of informants:*

- *It is optimal to offer leniency $q_1^* \geq \hat{q}$ only when a single firm reports, and no leniency otherwise ($q_l^* = 0$ for $l \in \{2, \ldots, \overline{N}\}$).*
- *Welfare is enhanced compared with a first informant rule: the initial proportion of collusive industries drops to $G(V^R(0))$ and the social cost of a cartel is reduced to $\frac{L}{1-\delta(1-\alpha)}$.*

Proof. In an industry with $N \in \{2, \ldots, \overline{N}\}$ firms, the social costs of collusion are equal to $G(V^S)\frac{L}{1-\delta(1-\alpha p)}$ for $\max(q_1, q_N) < \hat{q}$ and equal to $G(V^R(q_N))\frac{L}{1-\delta(1-\alpha)}$ for $\max(q_1, q_N) \geq \hat{q}$, and are thus minimized for $q_1^* \geq \hat{q}$ and $q_N^* = 0$. Hence, the optimal leniency policy prescribes $q_1^* \geq \hat{q}$ and $q_l^* = 0$ for $l \in \{2, \ldots, \overline{N}\}$.

□

Note first that $q_l^* = 0$ for $l \geq 2$ because we assume that a firm reports whenever it is indifferent between reporting and remaining silent (see footnote **??**). More generally, q_l^*

for $l \geq 2$ must be positive (equal to ϵ with ϵ arbitrarily low) to ensure that reporting is a (strictly) dominant strategy.

When designing the optimal leniency policy, the CA faces two *a priori* conflicting objectives – that is, giving cartel members incentives to report while minimizing the reduction in the fines that are ultimately granted. Offering leniency rates contingent on the number of informants allows to solve both objectives. Specifically, offering $q_1^* \geq \hat{q}$ in case of a single informant suffices to make S not incentive-compatible and $q_N^* = 0$ minimizes the value of R.[11] In other words, the optimal leniency scheme involves a *single informant* rule (hereafter, SI rule) – that is, leniency should be given only if a single firm reports information.[12]

Comparison with the FI rule. When the CA adopts a SI rule, it wins on both counts: firms are forced to play R (S is not robust to single reporting deviations when $q_1 \geq \hat{q}$) and they end up paying the full fine. It follows that adopting a SI rule instead of a FI rule enhances deterrence (both rules have the same effect on firms' incentives to betray the cartel): the initial proportion of collusive industries equals $G(V^R(0))$ under the SI rule and $G(V^R(\frac{\hat{q}}{N}))$ under the FI rule.[13]

4 Concluding Remarks

Spagnolo (2008, p293) discusses the objective of an optimal leniency program:

"*This means that a well-designed program must maximize incentives to betray the cartel by reporting important information to the Antitrust Authority, while at the same time limiting as much as possible the reduction in fines imposed on the whole cartel. This*

[11] As mentioned above, q_1 has no negative effect in equilibrium on expected sanctions. That is why, *any* q_1 above $\hat{q}$ is optimal.

[12] For a given industry with N firms, q_l^* for $l \in \{2, \ldots, N-1\}$ is irrelevant as firms, in a collusive equilibrium, either all remain silent or all report in case of investigation. The effective reduction in the fine is zero (or arbitrarily close to 0) for $l \in \{2, \ldots, \overline{N}\}$ because the number of firms may vary across industries.

[13] The analysis of the optimal leniency under a FI rule is thus the same as above with $q_1 = q$ and $q_N = \frac{q}{N}$. It follows that for $q < \hat{q}$, firms collude and remain silent if $V^S \geq \Pi^D$ and compete otherwise; and for $q \geq \hat{q}$, firms collude and report in case of investigation if $V^R(\frac{q}{N}) \geq \Pi^D$ and compete otherwise. Offering $q = \hat{q}$ then suffices to trigger reporting in case of investigation. As leniency is granted only to the first informant, expected penalties increase and the value of collusion drops from V^S to $V^R(\frac{\hat{q}}{N})$.

objective can be achieved by maximizing the benefits an individual cartel member can receive from reporting under the leniency program, but restricting such maximal benefit to one and only one reporting party, the first comer."

We fully agree with the diagnosis. However, our analysis challenges the optimal response: in order to minimize the reduction in fines imposed on the whole cartel, the SI rule outperforms the FI rule.

The SI rule may however be risky to implement in practice; in particular, any cartel members might then threaten to report if another firm reports first, thereby eliminating any gain from reporting in the first place.[14] To minimize this risk, the SI rule should be applied during a period over which the privacy of leniency applications could be guaranteed.

Note also that the success of the SI rule rests on the assumption that cartel members compete once condemned. When instead firms can go on colluding after being condemned (as e.g. in Chen and Rey, forth.; Spagnolo, 2004), firms will switch to a collusive equilibrium in which they take turns in reporting. In that case, the value of collusion is the same under the FI rule and the SI rule. The SI rule is thus likely to perform better than the FI rule only if industries in which cartels have been condemned are closely monitored (or if fines for repeat offenders are high enough to deter firms from colluding again).

Following Apesteguia et al. (2007), Hinloopen and Soetevent (2008) and Bigoni et al. (2012), an interesting avenue for future research could be to analyze the effects of the SI rule on cartel formation and collusive prices in laboratory experiments of leniency.

Acknowledgments

I thank Patrick Rey for his guidance, and Yassine Lefouili for very useful comments.

[14]Observe however that such a retaliation might not be credible. To see this, suppose that the probability of conviction is lower than one when one firm applies for leniency. In that case, in response to another firm being the first to apply, a cartel member would not want to retaliate because doing so delivers no reduction in penalty and raises its probability of conviction (as it now pleads guilty as part of its leniency application). I thank the referee for suggesting to me this point.

References

[1] Apesteguia, J., Dufwenberg, M., and Selten, R., 2007. Blowing the whistle. Economic Theory 31, 143–166.

[2] Aubert, C., Rey, P., and Kovacic, W., 2006. The Impact of Leniency and Whistleblowing Program on Cartels. International Journal of Industrial Organization 24, 1241–1266.

[3] Bigoni, M., Le Coq, C., Fridolfsson, S., and Spagnolo, G., 2012. Fines, Leniency and Rewards in Antitrust. Rand Journal of Economics 43, 368-390.

[4] Brenner, S., 2009. An Empirical Study of the European Corporate Leniency Program. International Journal of Industrial Organization 27, 639–645.

[5] Chen, Z., and Rey, P., forth. On the Design of Leniency Programs. Journal of Law and Economics.

[6] Hammond, S. D., 2010. The Evolution of Criminal Antitrust Enforcement Over the Last Two Decades. The 24th Annual National Institute on White Collar Crime, February 25, Miami.

[7] Harrington, J. E. Jr., 2008. Optimal Corporate Leniency Programs. Journal of Industrial Economics 56, 215–246.

[8] Harrington, J. E. Jr., and Chang, M.H., 2009. Modeling the Birth and Death of Cartels with an Application to Evaluating Antitrust Policy. Journal of the European Economic Association 7, 1400–1435.

[9] Hinloopen, J., and Soetevent, A.R., 2008. Laboratory Evidence on the Effectiveness of Corporate Leniency Programs. Rand Journal of Economics 39, 607–616.

[10] Miller, N.H., 2009. Strategic Leniency and Cartel Enforcement. American Economic Review 99, 750–768.

[11] Motta, M., and Polo, M., 2003. Leniency Programs and Cartel Prosecution. International Journal of Industrial Organization 21, 347–379.

[12] Sauvagnat, J., forth. Prosecution and Leniency Programs: The Role of Bluffing in Opening Investigations. Journal of Industrial Economics.

[13] Spagnolo, G., 2004. Divide et Impera: Optimal Leniency Programmes. Center for Economic Policy Research, Discussion Paper N.4840.