On Leniency and Markers in Antitrust: how many informants are enough?

Konstantinos Charistos and Christos Constantatos

Discussion Paper No. 2/2016

Department of Economics, University of Macedonia, 156 Egnatia str, 540 06 Thessaloniki, Greece, Fax: + 30 (0) 2310 891292
On Leniency and Markers in Antitrust: how many informants are enough?

Konstantinos Charistos and Christos Constantatos
Version: October 2018

Abstract
In this paper we investigate the impact of leniency programs on firms’ decision to collude. We depart from previous literature by relaxing the assumption that evidence provided by a single firm suffices to convict an existing cartel with certainty. Assuming the conviction-probability to be increasing in the number of reporting firms, we show first that efficient cartel deterrence requires incentives for all firms to report. Under a regime that secures a marker for the first in line applicant, eligibility for leniency should be extended to at least a second informant. We show that the introduction of the marker system has an ambiguous impact on cartel deterrence. In relation to the manner that the marker is secured and the cartel-related evidence is allocated, we derive the conditions under which allowing the first applicant to secure a marker enhances cartel deterrence.

JEL Classification: K21, L12, L41

Keywords: Antitrust enforcement, Collusion, Leniency programs

Department of Economics, University of Macedonia, e-mails: kostchar@uom.edu.gr and econst@uom.gr
1. Introduction

Leniency Programs (hereafter LPs) aim to improve cartel detection and deterrence by offering fine reductions to cartel members that either voluntarily self-report before there is even an investigation, or report and generally cooperate during investigation (pre- and post-investigation leniency, respectively). Pre-investigation leniency aims at destabilizing a cartel by making deviations from its central policy more attractive. Post-investigation leniency aims to evidence acquisition that is sufficient for an already spotted cartel to be convicted. By sufficiently increasing the probability of cartel conviction, post-investigation leniency has also important impact on the pre-investigation stability of the cartel, affecting the participating firms’ incentives to collude.

A common feature of many LPs is the presence of the marker system. The latter allows a number of reporting parties to reserve the position in queue for a finite period of time before the identities of the eligible for leniency are determined. In other words a marker removes the uncertainty for the leniency applicant about the existence of other informants and its own position in reporting line. According to OECD (2014) the majority of jurisdictions of OECD countries seem to have some kind of marker system. Most of them (including EC) restrict the availability of markers to the first reporting party, while others (including Canada, France, Germany, UK etc) offer this possibility to subsequent applicants as well.

Most of the LPs-related literature assumes that even a single firm’s reporting provides sufficient evidence for conviction with certainty. This implies that up to some details, all firms mainly possess similar evidence, and therefore the usefulness of any additional reporting is simply to strengthen the Antitrust Authority’s (AA) ability to prove the putative infringement. However, some practitioners (see Blatter et al., 2018) observe that firms have incomplete pieces of evidence and total evidence is cumulative, with each single reporting making conviction only more likely. Also, OECD (2012) states that “authorities are likely to find themselves in situations where, while aware of the existence of a cartel as a result of a leniency application by the first applicant, they are not yet in a position to prove the infringement”.

In this paper, we relax the assumption that a single firm’s reporting is sufficient for conviction, assuming instead that the probability of conviction is increasing in the number of reporting firms. This corresponds to assuming that the evidence brought by subsequent informants has added value, rather than being a mere corroboration of the
evidence offered by the first informant. This assumption has the interesting implication that inducing a single firm to report does not imply that all its partners have sufficient incentives to do as well. If the leniency is not sufficiently generous some firms may prefer to remain silent and avoid reporting in an attempt to restrain the conviction likelihood. We show that cartel deterrence requires the LP to provide incentives for universal reporting, \textit{i.e.} the AA should design the LP as to obtain all the available evidence.

The rationale of rewarding subsequent leniency applicants is described in OECD (2012), where it is suggested that when the authorities “are not yet in a position to prove the infringement, the social benefits from cooperation with the second or later applicants may be large compared to the public interest of penalizing the infringers”. These benefits may arise from cost savings in prosecution, increased detection rate and destabilizing effects on cartels.

A basic difference between the US and the EU LP pertains to the number of informants which are eligible for fine reductions. The US Department of Justice (DoJ) allows only the first firm that provides valuable information to receive amnesty from fines. In contrast, the European LP offers milder fine reductions to multiple informants. Parties that reveal information with significant added value can be awarded with reductions up to 50% of the fine that would have otherwise been imposed.

In this paper we compare the two systems by focusing on their impact on firms’ incentive to collude. The relevant theoretical literature usually concludes in favor of the first informant rule. Spagnolo (2004) claims that only the first informant should be rewarded sufficiently. Chen and Rey (2013) notes that allowing additional firms to be eligible for leniency reduces the effectiveness of the program. Harrington (2008) also supports restricting leniency to the first reporting firm, claiming that such a policy induces a “race to the courthouse” effect once an investigation has been launched. Sauvagnat (2014) shows that leniency should be provided when only a single firm reports information; when more than one firm are willing to report, none should receive any fines reduction.

The impact of marker system on the effectiveness of LP is studied in Blatter et al. (2018). Assuming imperfect and asymmetric evidence in a duopoly and restricting the eligibility for leniency to the first informant, they show that under the marker system only one firm reports and the AA obtains only partial evidence. They show that the
marker system increases the deterrence cost unless firms possess sufficiently asymmetric evidence.

Here, we show that if the first informant’s position is protected by a marker, the role of post-investigation leniency in destabilizing cartels in the pre-investigation period is substantially weakened by the first informant rule. Our analysis offers support to the European practice of allowing multiple informants to benefit from lenient treatment.

We also find that the impact of the marker system on the effectiveness of a LP to deter cartel activity crucially depends on the manner that the marker is available. We show that if the marker becomes available to following applicants once its initial holder fails to comply with its information-proving obligation, the use of markers reduces the LP’s power in cartel deterrence. On the contrary, if the marker is permanently lost following a denial to report by its initial recipient, the marker system may facilitate reporting by cartel members.

The rest of the paper is organized as follows: The model is described in the next section. In section 3 we analyze the benchmark case in the absence of marker system. In section 4 we introduce the marker. Section 5 concludes.

2. The model: investigation, conviction probability and fines

Consider an industry with \( N \) symmetric firms producing homogeneous goods and competing in prices for an infinite number of periods. Each firm maximizes the expected sum of future discounted profits using a common discount factor \( \delta \in \left( \frac{1}{2}, 1 \right) \).\(^1\) During each period a competition vs. collusion game takes place. If all firms cooperate setting the collusive price, each one earns an amount of profit, \( \pi \). When one firm unilaterally deviates from the agreed price, it receives \( N\pi \) while the other firms get zero. The competitive gross profits are also zero. In order to lighten the analysis and without loss of generality we normalize \( \pi = 1 \).

Resource limitations allow the AA to investigate the industry with probability \( \alpha \in (0,1) \).\(^2\) Even when the investigated firms are guilty, the start of an investigation

---

\(^1\) For \( \delta < 0.5 \) collusion is not sustainable even in the absence of antitrust policy.

\(^2\) Using data from DoJ price-fixing cases, Bryant and Eckhart (1991) estimated the probability of cartel detection to be between 0.13 and 0.17 in a given year. Combe et al. (2008) estimated the same probability over a European sample to be around 0.13.
does not necessarily imply conviction, unless a sufficient amount of evidence is collected. Concentrating only on cases where the investigated firms have indeed formed a cartel, we assume that, despite the presence of the infringement, the prosecution outcome is uncertain, with the probability of conviction being non-decreasing in the amount of available evidence. We measure evidence by the change in probability of conviction that it induces. Any evidence that is a mere repetition of evidence already in the possession of the AA does not increase the probability of conviction; it is therefore considered as redundant and not taken into account.

The total amount of cartel-related evidence is decomposed in two parts: common evidence, denoted by $z$, which is evidence possessed by every participant, and exclusive evidence which represents pieces of evidence in the possession of a subgroup of firms. To keep matters simple, we assume that every exclusive piece is detained by only a single firm, and that the exclusive evidence pieces are distributed symmetrically among firms, each one having in its possession an amount $\Delta \rho = \frac{1-z}{N}$.

We assume that AA’s actions (down-raids etc) uncover only a portion $\rho_0$ of the total evidence, and that this portion contains both, common and exclusive evidence in specific portions, i.e., $\rho_0 = \lambda_1 z + \lambda_2 N \Delta \rho$ with $0 \leq \lambda_h \leq 1$, $h = 1,2$; the value of $\rho_0$ determines the probability of conviction when no firm confesses. For simplicity, we assume that $\lambda_1 = \lambda_2$, i.e. the evidence unveiled by the AA’s efforts consists of equal portions of common and firm-specific evidence.\(^3\)

When a cartelized industry is investigated each convicted firm is forced to pay a fine $\mu$, where $\mu > 1$; $\mu$ is composed of a compensation to injured parties equal to the amount of illegally obtained profits, as well as of a pure fine paid to the authorities.\(^4\) The value of $\mu$ can be neither too low, for in this case the fine cannot induce compliance, nor too high as it may curtail competition in the long run by pushing some competitors out of business.\(^5\)

---

\(^3\) Assuming alternatively that $\lambda_1 \neq \lambda_2$ produces qualitatively similar results.

\(^4\) Harrington (2014) mentions that the standard formula for cartel-related damages is $(p^c - p^n)q^c$, where $p^c$ and $q^c$ are the collusive price and quantity respectively and $p^n$ is the Bertrand-Nash price. Bageri et. al (2013) shows that fines on revenues result in higher collusive prices that fines on illegal gain. In a dynamic context, Katsoulakos et al. (2015) shows that fines based on illegal profits are welfare superior to fines on revenues.

\(^5\) The US (federal) fines correspond to no more than double damages while other jurisdictions allow for up to treble damages, see Harrington (2014). A reasonable assumption for the value of $\mu$ is that
A post-investigation LP allows a cartel participating firm to provide information and/or evidence related to the existence of the cartel after the investigation’s opening, in exchange for a fines reduction. As the probability that an inspected cartel is convicted increases with the amount of the evidence, it is reasonable to assume that it is also non-decreasing in the number of reporting parties. The common share of the additional evidence is provided only once by the first firm that testifies. The first informant increases the probability of conviction by \((1 - \rho_0)z\), in addition to any exclusive piece it may present. When more firms confess, every subsequent informant increases the probability of conviction by \((1 - \rho_0)\Delta \rho\). Hence, when \(n \in \{1, N\}\) firms report, the probability of conviction becomes:

\[
\rho_n = \rho_0 + (1 - \rho_0)(z + n\Delta \rho) = \frac{N(z + \rho_0 - z\rho_0) + n(1 - z)(1 - \rho_0)}{N}
\]

A common feature among LP-related legislation in different countries is that it restricts post-investigation leniency to only a limited number \(m\) of applicants. Usually, the eligible firms are selected on a first-come-first-served basis, subject to the requirement of providing sufficient amount of evidence.\(^7\) Even in jurisdictions where all the applicants are eligible, their treatment is asymmetric, with the “early birds” receiving substantially more generous treatment.

When confessing, those eligible for leniency will receive a fine reduction proportional to their individual contribution to cartel prosecution, being required to pay only a fraction of the full fine. The reduced- to-full-fine ratio is:

\[
\gamma_1(k_f) = \frac{1 - \rho_0 - k_f(1 - \rho_0)(z + \Delta \rho)}{1 - \rho_0} = 1 - k_f(z + \Delta \rho)
\]

for the first informant, and

\(\mu \in [2,3]\). However, as Harrington (2014) points out, in practice firms found guilty by a court of law pay fines that “are probably more on the order of single rather than treble damages”.

\(^6\) Solving \(\rho_n \leq 1\) yields \(n \leq N\). Hence, certain conviction requires reporting by all firms. If instead we assume that one firm’s reporting increases the likelihood of conviction by more than \((1 - \rho_0)\Delta \rho\), e.g. by \((1 - \rho_0)b\Delta \rho\) where \(b > 1\), then confession by less than \(N\) firms suffices to raise the probability of conviction to one. Without any loss of generality we assume here that \(b = 1\), i.e. that certain conviction requires reporting by all firms. Our results hold if we assume that \(b > 1\).

\(^7\) For instance, the US system grants leniency to a single applicant, subject to the condition that it provides substantial evidence. The EU system allows for many applicants, however, it offers them asymmetric treatment, with leniency being more generous for those that come out early and decreasing for subsequent informants.
\[
\gamma(k_s) = \frac{1 - \rho_0 - (1 - \rho_0)z - k_s(1 - \rho_0)\Delta \rho}{1 - \rho_0 - (1 - \rho_0)z} = \frac{N - k_s}{N}
\]
for one subsequent eligible applicant. Both \(\gamma_1\) and \(\gamma\) lie between 0 and 1, i.e. we rule out rewards. Both parameters \(k_f\), \(k_s\) measure how additional information by the first and subsequent informants, respectively, is rewarded by the AA, and thus determine the generosity of the leniency offered to each eligible informant. According to the AA’s specific policy, the parameters \(k_f\), \(k_s\), can be equal or unequal. Unequal \(k\)’s reflect price discrimination, i.e., that the AA awards a difference price per unit of information received by the first, or subsequent informants.\(^8\) As we restrict leniency to non-negative fines the leniency rates are bounded from above: \(\gamma_1(k_f) \geq 0\) and \(\gamma(k_s) \geq 0\) imply \(k_f \leq \bar{k}_f \equiv \frac{N}{1 + z(N - 1)}\) and \(k_s \leq \bar{k}_s \equiv N\) respectively.

A firm’s decision on whether to come forward and provide evidence crucially depends on that firm’s perception about its position on the priority line. The accuracy of this perception depends in turn on the AA’s information-diffusion policy. With respect to the latter, we examine two alternative systems.

The first, that we term “opaque practice” allows no firm to have information about the existence of other confessants; therefore, as the investigation proceeds no firm can be aware of its position on the priority line. Due to this information restriction, cartel participants act as if the decision on whether to come forward must be taken simultaneously by all of them. Combined with the assumption that different positions on the priority line receive asymmetric treatment, simultaneous decision implies that, when making its decision, a cartel participant is unaware of the leniency treatment it will finally receive. If it decides to confess thinking that \(n - 1\) others are going also to confess, it must assign some positive probability to being a) the first informant, b) one of the \(m\) eligible ones, or c) one of the \(n - m\) that report but receive no lenient treatment. Hence, its expected fine must be a fraction \(\hat{\gamma}_n\mu\) of the full fine, where:

\[
\hat{\gamma}_n = \frac{\gamma_1 + \gamma(m - 1) + (n - m)}{n}
\]

\(^8\) It is equally possible to replace \(k_s\) by a sequence \(k_i\), \(i = 2, \ldots, m\), where \(m\) is the number of eligible-for-leniency applicants, implying that the AA accords different importance to the information provided by each different applicant. In order to keep the analysis simple, and as it turns out without loss of generality, we limit the possibility of information-price discrimination only between the first and subsequent applicants.
with \( n \in \{1, \ldots, N\} \) and \( m \leq n \). For the rest of the analysis we assume that the number of eligible for leniency informants can be either one or two, that is \( m \in \{1,2\} \).

In the second system, often termed “marker” system, before providing evidence each firm has secured a position on the priority line, thus knowing exactly the kind of leniency treatment it will receive. In practice, a mixed system is often followed, where the first few applicants for leniency secure their position, while subsequent applicants only know that they will not occupy any of the already reserved positions. In this work we assume that a marker is handed only to the first-to-door applicant. The marker allows its holder some given time period in order to prepare and present the promised evidence. If at the end of that period the marker holder refuses to deliver the evidence, the marker may or may not become available for another potential applicant. As the transferability of the marker has important implications for the LP’s efficiency, we examine both cases.

The timing of the game is as follows. At the beginning of every period each firm decides whether to collude or not; if at least one firm refuses to collude, competition takes place at least up to the end of the period. If a cartel agreement is reached, in all subsequent periods each firm chooses between staying loyal or defecting from it. A deviation from the collusive price implies that the market will be competitive ever after (trigger strategies). At the end of each period, after firms have set their prices and made the current period profit, the AA randomly decides with probability \( a \) whether to investigate the industry. Collusion evidence can be used for only one period, \( i.e. \) firms cannot be convicted for past violations. In case of investigation, each cartel participant chooses between reporting or not. Focusing on the deterring impact of leniency policies, we make the simplifying assumption that, regardless of whether it leads to conviction, an investigation implies the definite dissolution of the collusive agreement.\(^9\) Finally, the cartel is convicted with probability \( \rho_n \), depending on the amount of evidence collected.

3. The opaque practice

\(^9\) We assume that, no matter whether the investigation leads or not to conviction, the AA monitors the investigated market for an infinite number of periods, forcing firms to compete forever after. Assuming instead that convicted firms keep colluding produces qualitatively similar results.
Once an investigation has started each firm faces a multi-person prisoners’ dilemma, and as the AA follows a full-secrecy policy, an investigation node corresponds to a subgame where the reporting decision is taken simultaneously by all firms.\(^{10}\) The superscript “0” indicates hereafter equilibrium values under AA’s opaque policy.

In order to find the equilibrium of the investigation-subgame, we must determine each firm’s best reply function. Consider a node where the AA decides to investigate. If a firm thinks that \(n - 1\) others, \(n \in [2, N]\), are about to report, it expects to pay the full fine with probability \(\rho_{n-1}\), and no fine in case of unsuccessful prosecution. Reporting, on the other hand, reduces that period’s profit by a percentage \(\rho_n \gamma_n\) and is the best reply to \(n - 1\) other firms choosing to report when:\(^{11}\)

\[
1 - \rho_n \gamma_n \mu \geq \rho_{n-1} (1 - \mu) + 1 - \rho_{n-1}
\]

which simplifies to \(\varphi(n) \geq 0\) where

\[
\varphi(n) \equiv \frac{N\rho_n [k_f (1 + (N - 1)z) + k_s] - Nn(1 - z)(1 - \rho_0)}{N^2 n}
\]

where \(N\rho_n = n(1 - z)(1 - \rho_0) + N(z + \rho_0 - z\rho_0)\).

Note that \(\frac{\partial \varphi(n)}{\partial n} = -\frac{[k_f (1 + (N - 1)z) + k_s] \rho_0 + (1 - \rho_0)z}{Nn^2} < 0\). The following lemma links the number of reporting parties with the incentive to denounce the agreement, given the presence of at least one applicant:

**Lemma 1** *In case of investigation, for every \(n \in [2, N]\) the incentive to report is monotonically decreasing in the number of other informants.*

Since the incentive to report is negatively affected by the number of reporting parties, if a firm decides to report as part of a group with \(n - 1\) other informants, it is also willing to report when thinking that there are fewer informants in the group. For any given \(n\), setting \(\varphi(n) = 0\), determines a relation between \(k_f\) and \(k_s\) that allows at most \(n\) firms to come forward. Solving this relation for \(k_f\), obtains

---

\(^{10}\) Simultaneous reporting is meant to represent that before deciding whether to confess, a firm must “guess” how many others are about to report. “Guessing” correctly is equivalent to assuming that during the reporting process each potential whistleblower is notified about the number of the other reporting parties and taking this into account decides its action. At the end of the reporting phase, the names of those eligible for leniency are determined randomly.

\(^{11}\) Observe that allowing an investigated cartel to continue its collusive activity affects both sides of (2) positively.
\[ k_{n,k_s}^0(k_s; n) \equiv \frac{1}{1 + (N - 1)z} \left[ \frac{Nn(1 - z)(1 - \rho_0)}{N\rho_n} - k_s \right] \]  

(3)

For any reward per unit of information offered to the second informant, for the LP to provide sufficient incentive for \( n \) informants to come forward, the per-unit-of-information reward offered to the first informant must be no less than \( k_{n,k_s}^0 \) as defined above. Due to the negative sign of the coefficient of \( k_s \), expression (3) defines a trade-off between rewards to the second and first informant that makes the pair \((k_f, k_s)\) sufficient to induce \( n \) firms “racing to the court,” even if they know that \( n - 2 \) of them will receive no reward for the information they will provide.

We define the value of \( k_{n,k_s}^0 \) when \( k_s = 0 \) \((m = 1)\) as \( k_n^0 \):

\[ k_n^0 = \frac{Nn(1 - z)(1 - \rho_0)}{[1 + (N - 1)z]N\rho_n} \]

**Corollary 1** When \( k_{n+1,k_s}^0 > k_f \geq k_{n,k_s}^0 \), the equilibrium of the post-investigation subgame involves \( n \in [2, N] \) informants.

Note that for every \( n \in [2, N - 1] \), the pair \((k_{n,k_s}^0, k_s)\) determines multiple equilibria of the investigation subgame. These equilibria are qualitatively similar: they contain the same number of informants, differing only with respect to the identity of the reporting firms. More serious is the potential existence of another equilibrium where no firm comes forward, investigated right below.

Lemma 1 establishes that a firm’s incentive to confess is reduced with the number of other firms that this firm thinks they have also decided to confess. This rule applies to firms that think that at least another cartel member is going to report. However, the situation of the unique informant is different and not described by lemma 1: when thinking that no other firm has decided to confess, a firm may decide not to come forward even if it would have done so under the assumption that some another firm has already decided to confess. Because the first informant offers all the common evidence the rewards are larger, but also the consequence in terms of increasing the conviction probability graver. When the latter creates a sufficiently strong disincentive, universal non-reporting is an equilibrium, along with the equilibria mentioned earlier. Compared to them, the non-reporting equilibrium is Pareto dominant, and for this reason it is very important for the AA to design its policy as to eradicate it.
Lemma 2 Universal non-reporting is an equilibrium of the investigation subgame when \( k_f < k_1^0 \), where
\[
k_1^0 \equiv \frac{N(1 - \rho_0)}{1 + (N - 1)(z + \rho_0 - z\rho_0)}
\]

Proof
If a firm reports assuming that no other does so, it expects to pay the reduced fine with probability \( \rho_1 \), and to receive nothing thereafter. If it chooses to remain silent as everybody else, it expects to pay the full fine with probability \( \rho_0 \) and no fine otherwise. Reporting is the best reply to all other firms remaining silent when:
\[
1 - \rho_1 \gamma_1(k_f) \mu \geq 1 - \rho_0 + \rho_0(1 - \mu)
\]
Solving the above for \( k_f \) yields the value \( k_1^0 \) stated in (4).■

Subject to the constraint that the resulting fine is nonnegative, the value of \( k_1^0 \) in (4) represents the minimal implicit price per piece of information at which the AA must purchase the first informant’s evidence—both, common and exclusive—in order to induce at least a single firm to come forward when thinking that all the others will remain silent. Substituting \( k_1^0 \) into the definition of \( \gamma_1 \) yields:
\[
\gamma_1(k_1^0) \equiv \frac{N\rho_0}{1 + (N - 1)(z + \rho_0 - z\rho_0)}
\]
Note that \( \gamma_1(k_1^0) \) always decreases with \( z \), thus, when the unique informant possesses a large amount of the common share of evidence, it requires more generous fine reductions in order to come forward.

If the LP provides sufficient incentives for a single informant to come forward, the race to the court is not guaranteed: other firms may not follow suit, and to the extent that they possess pieces of evidence not available to the first informant, conviction is not 100% certain. As we will see later, this may also have serious implications for the effectiveness of the LP in deterring cartel formation.

Lemma 3 For \( k_f = \max\{k_1^0, k_n^0, k_s^0\} \) at least \( n \) firms reveal under investigation.

Proof
When \( k_1^0 \geq k_n^0 \Leftrightarrow \rho_0 \leq \frac{[1+(N-1)z]z}{(1-z)(n-1-(N-1)z)} \) (for positive \( k_s \) the latter threshold increases rendering \( k_1^0 \geq k_n^0, k_s^0 \) easier to hold) offering \( k_f = k_n^0 \) to the first informant implies that either \( n \) or no firms confess. Note that the payoff when \( n \) informants exist
is $1 - \rho_n \mu \hat{\gamma} \left( k_{n,k,s}^0, k_s^0 \right)$ for each reporting firm whereas if all firms remain silent each one earns $1 - \rho_0 \mu$. As

$$\rho_n \hat{\gamma} \left( k_{n,k,s}^0, k_s^0 \right) = \frac{(n - 1)(1 - z)(1 - \rho_0) + N(z + \rho_0 - z\rho_0)}{N} \geq \rho_0$$

holds for every $\rho_0 \leq 1$ and assuming that firms coordinate on the most profitable equilibrium, reporting by $n$ parties requires $k_f = k_1^0$ to be offered. When $k_1^0 < k_n^0$, offering $k_f = k_n^0$ is enough for the same outcome ($n$ informants) to be achieved.

Solving (3) for $k_s$ and setting $k_f = k_1^0$ yields that $k_s(k_1^0) > 0$ holds for $\rho_0 > \frac{[1 + (N - 1)z]}{(1 - z)(n - 1 - (N - 1)z)}$. Therefore, offering $k_1^0$ to the first and $k_s(k_1^0)$ to the second informant is not possible when $k_1^0 > k_n^0$. In the latter case $k_1^0$ should be offered to only one informant.

Only by imposing fines $\gamma_1 \leq \gamma_1(k_1^0)$ to the first informant the AA can be certain that at least one firm will come forward. Had a single firm’s reporting been able to raise the conviction probability to 1—as is commonly assumed in the literature—a fine $\gamma_1 \leq \gamma_1(k_1^0)$ would have been sufficient to induce universal reporting. However, if the probability of conviction increases monotonically with the number of informants, offering sufficient incentive for the first informant to come forward may not always induce a universal reporting.

Note that if it is feasible to offer leniency rates contingent on the number of informants, setting $k_f = k_{n,k,s}^0$ would be sufficient to induce reporting by $n \geq 2$ firms even if $k_1^0 \geq k_{n,k,s}^0$; offering $k_f = k_1^0$ or $k_f = k_{n,k,s}^0$ if the number of eligible informants is either one or $n \geq 2$ respectively would be enough to induce reporting by $n \geq 2$ informants as every firm would be motivated to confess as the unique informant. For the rest of the analysis assume that leniency rates are not possible to

---

12 That is $\rho_1 = 1$ which under the assumption of firms' symmetry with respect to the evidence they possess implies $z = 1$. In the latter case $k_1^0 = 1 - \rho_0 \geq k_{n,k,s}^0 = -k_s \frac{N}{n}$.

13 In case where $z = 1$, promising $k_f = 0$ if $n \geq 2$ and $k_f = \frac{N(1 - \rho_0)}{1 + (N - 1)}$ if $n = 1$ would be enough to induce universal reporting with no fine reductions in equilibrium, as in Sauvagnat (2014). For $0 \leq z < 1$ some leniency is necessary, even if leniency rates are contingent on the number of informants.
be contingent on the number of informants, therefore, the condition described in lemma 3 must hold.

Both $k_1^0$ and $k_{N,k_s}^0$ depend on the values of $(N, z, \rho_0)$, whereas $k_{N,k_s}^0$ depends also on the value of $k_s$ promised to the second informant if $m = 2$. For every value of $k_s$, there exists a constellation $(N, z, \rho_0)$ such that $k_1^0 \equiv k_{N,k_s}^0$. Solving the latter for $z$ obtains

$$z_a \equiv \frac{\sqrt{[k_s+N(1-k_s)]^2+4N^2\rho_0(N-1)-(k_s+2N\rho_0)(N-1)-N}}{2(N-1)N(1-\rho_0)} \quad (5)$$

It can be shown that $z > z_a$ is equivalent to $k_1^0 > k_{N,k_s}^0$, hence when $z > z_a$ offering $k_f = k_{N,k_s}^0$ to the first informant is not enough to induce universal reporting. In such case offering $k_f = k_1^0$ for the first in line and no leniency for any subsequent informant induces reporting by every cartel participant. This implies that when the evidence brought-in by the first comer increases significantly the conviction rate, inducing the unique informant to come forward suffices to trigger a race to report by all firms. Otherwise, even if some firms report, others may find it preferable to hold back.

**Corollary 2** *Inducing the unique informant to report triggers reporting by every firm when $z \geq z_a$. Otherwise the same outcome requires $k_f = k_{N,k_s}^0$ and $k_s$ to be offered.*

Regarding the equilibrium selection, we consider some random mechanism (perhaps focal points) determining which firm belongs to each group. We assume that in case of multiple equilibria, equilibrium selection takes place at the beginning of the investigation. Thus, at the beginning of the game where firms must adopt an open loop strategy, all firms know that there will be equilibrium with $n$ firms reporting and $N - n$ remaining silent, but no firm knows which equilibrium will be selected and therefore to which group it will belong in the occurrence of an investigation. Instead, if a typical investigation-subgame has multiple equilibria involving both reporting, and non-reporting firms, each firm assigns a probability for being in the reporting group.

**Equilibrium of the entire game**

The previous analysis implies that depending of the characteristics of the LP (generosity and number of eligible firms) the equilibrium of the subgame starting
from a node where the AA decides investigation, none, many, or all the firms involved in the cartel may report information and/or evidence.

When firms decide whether to cheat or to remain loyal to the agreement, they compare the value of collusion to the gain from unilateral deviation. We assume that the firm that unilaterally defects is absolved from any fine imposition, therefore the value of cheating is $N$. The purpose of the AA is, with the use of LP, to minimize the cartel value and consequently to increase the minimum $\delta$ above which the collusive agreement is sustainable.

Now we turn to the case where the AA promises $k_f = k^0_{n,k_s} > k^1_0$ to the first eligible firm and $k_s$ to one subsequent firm in order to induce post-investigation reporting by $n$ parties. Initially, each participant expects to earn the collusive profits and, in case of successful -with probability $\rho_n$- investigation, to pay the full fine with probability $\frac{N-2}{N}$ or a reduced fine with probability $\frac{2}{N}$. Therefore, the value of collusion where $n$ firms report under investigation is:

$$V_n^0 = \frac{1 - a\rho_n \gamma(k^0_{n,k_s})\mu}{1 - \delta(1-a)}$$

with $\gamma(k^0_{n,k_s}) = \gamma_1(k^0_{n,k_s}) + \gamma(k_s) + (N-2)$, $\psi = \rho_n N^2$ and $\psi = N^2(z + \rho_0 - z\rho_0) + n(N - 1)(1 - z)(1 - \rho_0)$

The number of informants affects the value of the cartel through the conviction rate as well as through the level of the expected fine. The following proposition determines the optimal number of firms induced to report under investigation:

**Proposition 1** Offering the minimum amount of leniency that induces every firm to report when an investigation is underway, always dominates in terms of ex-ante deterrence any asymmetric case where $n \leq N - 1$ firms report.

**Proof**

Substituting for $\gamma_1(k^0_{n,k_s})$ and $\gamma(k_s)$ into $V_n^0$ yields

$$V_n^0 = \frac{N^2[1 - a\mu(z + \rho_0 - z\rho_0)] - a\mu(1 - z)(1 - \rho_0)(N - 1)}{N^2[1 - \delta(1-a)]}$$

Note that $\frac{\partial V_n^0}{\partial n} = - \frac{a\mu(N-1)(1-z)(1-\rho_0)}{N^2[1-\delta(1-a)]} < 0$, i.e. the value of the cartel is decreasing in the number of reporting firms. Hence, it is always optimal to set $n = N$, i.e. to induce reporting by every cartel participant.
Observe that the cartel value is higher as the overall level of fine decreases:

\[
\frac{\partial V_0^n}{\partial \gamma} = -\frac{a\mu\rho_n}{1 - \delta(1 - a)} < 0
\]

while \( \frac{\partial \tilde{V}(k_{n,k_s})}{\partial n} = -\frac{(1-z)(1-\rho_0)(z+\rho_0-z\rho_0)}{\rho_n^2N^2} < 0 \), \text{i.e. more lenient treatment is required for more informants to be attracted. Thus, more informants have a positive impact on cartel sustainability as the value of the cartel increases when fines are lower, while the latter have to be low in order to attract more informants. In the contrary, as

\[
\frac{\partial V_0^n}{\partial \rho_n} = \frac{a\mu_n[(n(N-1)(1-z)(1-\rho_0) + N^2(z+\rho_0-z\rho_0)]}{N[1 - \delta(1 - a)][n(1-z)(1-\rho_0) + N(z+\rho_0-z\rho_0)] < 0}
\]

the effect of the likelihood of conviction on the value of cartel is negative, as expected. At the same time the probability of conviction increases with the number of informants: \( \frac{\partial \rho_n}{\partial n} = \frac{(1-z)(1-\rho_0)}{N} \). Therefore, \( n \) has a parallel impact on cartel stability through \( \rho_n \): more informants increase the likelihood of conviction which in turn decreases the value of collusion, \( \frac{\partial \rho_n}{\partial n} < 0 \).

Proposition 1 states that \( \frac{\partial V_0^n}{\partial \gamma} + \frac{\partial V_0^n}{\partial \rho_n} < 0 \) always holds: the benefit that the increased likelihood of conviction has on deterrence outweighs the adverse impact that the reduced level of overall fines generates. Hence, offering substantial post-investigation leniency improves both the deterrence and the prosecution, as the set of the created cartels is minimized and every investigated cartel is condemned.

Finally observe that \( V_0^n \) is independent of \( k_s \). As \( k_{n,k_s}^0 = \frac{N(1-z)(1-\rho_0)-k_s}{1+z(N-1)} \) increases when \( k_s \) lowers and \( \gamma_1(k_{n,k_s}^0) + \gamma(k_s(k_{N,k_s}^0)) = 1 + \rho_0 + z(1 - \rho_0) \), \text{i.e. the expected fine is unaffected by \( k_s \), we can hereafter assume that also when \( z < z_a \) the number of eligible informants is one (\( m = 1 \)), that is \( k_s = 0 \).

Assumption When no marker is available only the first informant is eligible for leniency. The only eligible for leniency firm receives \( k_f = k_N^0 = \frac{N(1-z)(1-\rho_0)}{1+z(N-1)} \) if \( k_N^0 > k_1^0 \) and \( k_f = k_1^0 \) otherwise.

Let us now define strategy C of the entire game as the usual trigger strategy with the additional feature of dictating to remain silent in case of investigation. A firm that plays C expects with probability \( 1 - a \) to keep receiving the collusive profits, and
with probability $a\rho_0$ to pay the fine $\mu$, and keep receiving the competitive profit for an infinite number of periods. The value of $C$ is therefore:

$$V^C = (1 - a)(1 + \delta V^C) + \alpha[(1 - \rho_0) + \rho_0(1 - \mu)]$$

Solving for $V^C$ yields:

$$V^C = \frac{1 - a\rho_0\mu}{1 - \delta(1 - a)}$$

As mentioned before, when $k_1^0 \geq k_N^0$ promising $k_f = k_N^0$ to one firm, is not sufficient to induce any reporting: as $V^C > \tilde{V}_N^0$ holds for every $n \in [1, N]$, firms select to coordinate on the most profitable $C$ which entails that no one confesses under investigation. Hence, offering $k_f = k_1^0 > k_N^0$ to the first informant induces reporting by every participant. In this case the value of the cartel becomes:

$$\tilde{V}_N^0 = \frac{1 - a\mu_1(k_1^0) + (N - 1)}{N}$$

For the rest of the analysis consider that the number of eligible for leniency firms under the no marker regime is $m = 1$, i.e. that only the first in line reporting firm receives fine reduction. The value of collusion becomes:\footnote{Observe again that if offering leniency rate contingent on the number of informants was possible, promising $k_1^0$ to the unique informant and $k_n^0$ if $n \geq 2$ would be enough to induce universal reporting with $k_f = k_1^0$, regardless of the level of $z$. In such case the cartel value would be always equal to $\tilde{V}_N^0$.}

$$V_N^0 = \begin{cases} \tilde{V}_N^0 = \frac{N(1 - a\mu) + a\mu(1 - z)(1 - \rho_0)}{N[1 - \delta(1 - a)]} & \text{if } z < z_a \\ \tilde{V}_N^0 & \text{if } z \geq z_a \end{cases}$$

where from (5) $z \geq z_a \equiv \frac{\sqrt{1 + 4N\rho_0(N-1)-2\rho_0(N-1)^{-1}}}{2(N-1)(1-\rho_0)}$. 

4. Marker

Now consider that the first-to-door applicant can secure its position and the AA announces that the privileged first position is no longer available. Besides this piece of information, potential subsequent informants remain unaware of the total number of informants that may have shown up already as well as of their precise position on the informants’ queue. Usually the marker is secured only for a specific time period considered necessary for its holder to organize and present the promised evidence. If at the expiry date the holder has failed to deliver the evidence, its position ceases to be secured, and we assume that this is common knowledge. We also assume that if a firm
has denied confession as marker holder, it shows no interest in confessing as subsequent applicant, and this is common knowledge as well.\footnote{Using very mild restrictions on the fines structure it can be shown that confessing as marker holder dominates confessing as subsequent informant. We state it as assumption in order to avoid burdening the analysis.} We distinguish two types of marker system according to the way the AA may treat a confession denial by the marker holder: the marker is either transferred to the next firm in the priority line, or permanently lost. As all firms know whether the marker has been already attributed (although they may not know the identity of the holder) both marker systems are unable to bring in more than one informant unless they offer leniency to at least one more applicant from the lot.

4.1 Scrolling marker

First, we analyze the case where the marker is transferred to the next applicant when a previous holder chooses to remain silent. If the marker holder indeed confesses, the subsequent $N-1$ firms take their reporting decision simultaneously, knowing that the first-informant position is not available. If the first marker holder decides to remain silent the marker is transferred to the second in line and in case of the second marker holder’s reporting the $N-2$ subsequent firms take the reporting decision simultaneously, etc.

Since further reporting cannot be induced without making sure that some leniency is also offered to at least one applicant on top of the marker holder, we assume that $m = 2$. If a subsequent firm thinks that $n - 1$ others are going to report, it deduces that by remaining silent it pays the full fine with probability $\rho_{n-1}$, whereas, by reporting it takes a leniency-winning stake with probability $\frac{1}{n-1}$ at the price of increasing the probability of conviction by $\Delta \rho$. Hence it will report if:

$$\rho_{n-1} \geq \rho_{n} \frac{y(k_s) + (n-2)}{n-1}$$

which yields

$$k_s \geq k_n \equiv \frac{N(n-1)(1-z)}{[n(1-z) + N(z + \theta)]}$$

where $\theta = \frac{\rho_0}{1-\rho_0}$. Observe first that $k_n$ is decreasing in $z$, since a larger common evidence reduces the impact of additional reporting on the conviction rate, and
therefore the price of the reporting lottery. Second, note that since any $k_s \geq k_n$ is able to bring forward $n$ informants, the rule of offering only fine reductions and no positive rewards, i.e. $\gamma(k_s) \geq 0$, requires that $k_s \leq N$. As $k_n \leq N \Leftrightarrow (1 - z) + N(z + \theta) \geq 0$ always holds, just offering $k_s = k_n$ to one subsequent applicant suffices to attract any exogenously determined number $n$ of informants (given that the marker recipient confesses).

The following lemma defines the necessary treatment for the marker recipient to report:

**Lemma 4** Assuming a regime that offers a transferrable marker to the first-to-door applicant, the latter requires at least $k_f = k_1^0$ in order to come forward, where $k_1^0$ is defined in (4).

**Proof**

See the Appendix.

When the LP offers sufficient incentive for $n$ firms to report when the cartel is under investigation, when making the decision of whether to join the cartel and respect the agreement, each firm anticipates that in case of investigation it will be the marker holder with probability $\frac{1}{N}$, and a subsequent leniency recipient with probability $\frac{m-1}{N} = \frac{1}{N}$ for the case of $m = 2$ assumed here. The value of collusion where $n$ firms report under investigation is:

$$V_{n}^{1} = \frac{1 - a\rho_n \mu \gamma(k_1^0, k_n)}{1 - \delta(1 - a)}$$  \hspace{1cm} (8)

where $\gamma(k_1^0, k_n) = \frac{\gamma_1(k_1^0) + \gamma(k_n) + (N-2)}{N}$

The next proposition states that eligibility for leniency should be extended in order to allow the AA to obtain maximum evidence:

**Proposition 2** Consider that the AA offers a transferable marker to the first in line applicant:

i. **Maximum efficiency in cartel-formation deterrence requires a LP design offering incentives that induce reporting by all firms, i.e. the AA should obtain all the available evidence. Necessary for the latter is to offer sufficient leniency to at least one subsequent applicant.**

ii. **It is always superior in terms of cartel deterrence to maintain the uncertainty among cartel participants, regarding their position in reporting line.**
Proof

See the Appendix.

Proposition 2 shows that the DoJ’s leniency regime, where i) only one informant receives leniency, ii) its position is reserved (marker), and iii) in case of marker holder’s withdrawal the marker can be transferred to another applicant, may not attain maximum efficiency in deterring cartel formation. While the amount of leniency offered may be sufficient in order to attract the (important) first informant, as leniency is restricted to one firm, no other party has incentive to increase the likelihood of conviction by reporting, therefore the number of eligible and the number of actual informants coincide. According to proposition 2, leniency should be offered to at least one more applicant, and according to proposition 1 this additional leniency should be offered without marker.

The above provides an argument that supports the European system’s practice to extend the eligibility to subsequent applicants. As mentioned before, the DoJ’s LP restricts the eligibility to the first informant which implies that when firms possess imperfect evidence only a single firm’s testimony is obtained. On the other side, the European system manages to extract evidence from multiple firms, a fact that seems to improve not only the cartel detection and the collection of fines but cartel deterrence as well.

4.2 Non-scrolling marker

Let us now analyze the case where the marker is withdrawn once the marker holder denies confession. In this case if the marker recipient confesses all subsequent firms take the reporting decision simultaneously, as in the transferred marker case. If the holder remains silent the marker is lost and the \(N - 1\) firms take their decision as in the no marker regime.

Consider that the AA provides incentive to the \(n - 1\) subsequent parties to come forward, following a confession by the first in line, i.e. \(k_s = k_n\) is offered to one additional applicant. Consequently, if the marker holder confesses the total number of informants is \(n\).

Let us keep the assumption from the previous section that without marker (opaque system) leniency is offered only to the first-to-door applicant. Incentive for universal reporting (race to the court) is provided if the value of \(k_f\) is given by (3) for \(k_s = 0\):
\[ k_n^0 = \frac{Nn(1 - z)}{[1 + (n - 1)z][n(1 - z) + N(z + \theta)]} \]

**Lemma 5** In a regime that offers a non-transferable marker to the first-to-door applicant; inducing universal reporting is always optimal.

**Proof**

Reporting by the marker holder implies a conviction rate equal to \( \rho_n \), while if remaining silent conviction takes place with probability \( \rho_{n-v} \), with \( v \in [0, n] \). Therefore the marker holder has sufficient incentive to confess when:

\[
1 - \rho_n \gamma_1(k_f) \mu \geq (1 - \rho_{n-v} + \rho_{n-v}(1 - \mu)
\]

which is equivalent to \[
\frac{[N-k(1+(N-1)z)[n(1-z)\rho_n+N(z+\rho_0 z \rho_0)]}{N^2} \leq z + \rho_0 - z \rho_0 + \frac{(1-z)(1-\rho_0)(n-v)}{N}
\]

or \( k_f \geq k_{1v}^n \), where

\[ k_{1v}^n = \frac{\nu N (1 - z)}{[1 + z(N - 1)][n(1 - z) + N(z + \theta)]]}
\]

Substituting \( k_f = k_{1v}^n \) into \( \gamma_1(k_f) \), we obtain the corresponding value of the cartel:

\[
V_{1v}^n = \frac{1 - a \rho_n \mu \gamma_1(k_{1v}^n) + \gamma(k_n) + (N - 2)}{1 - \delta(1 - a)}
\]

or \( V_{1v}^n = \frac{N^2[1-a \mu(z+\rho_0 z \rho_0)]-a \mu(1-z)(1-\rho_0)[n(N-1)+1-\nu]}{N^2[1-\delta(1-a)]} \).

As \( \frac{dV_{1v}^n}{dn} = - \frac{a \mu(N-1)(1-z)(1-\rho_0)-a \mu(1-z)(1-\rho_0)[n(N-1)+1-\nu]}{N^2[1-\delta(1-a)]} < 0 \), it is optimal to provide \( k_f = k_{1v}^n = \frac{\nu (1-z)(1-\rho_0)}{1+z(N-1)} \)

for the marker holder and \( k_s = k_N \) to one subsequent applicant. \(\blacksquare\)

Assume now that regardless of the marker recipient’s reporting decision \( k_N = (N - 1)(1 - z)(1 - \rho_0) \) is offered to one of the subsequent reporting firms. Therefore, following the marker holder’s denial to confess, one of the \( N - 1 \) others receive \( k_f = k_N \). Note that when the marker holder fails to confess and \( k_N \geq k_1^0 \) holds, offering \( k_f = k_N \) to one informant induces reporting by \( n \) firms provided that \( k_N \geq k_n^0 \). If \( k_N < k_1^0 \) no firm reports for \( k_f = k_N \). Defining the level of \( z \) below which reporting is possible in case of marker holder’s denial to report, \( k_N \geq k_1^0 \) holds when \( z \leq z_b \) with

\[
z_b \equiv \frac{N - 2 + \sqrt{N\sqrt{N - 4(1 - \rho_0) - 2(N - 1)\rho_0}}}{2(N - 1)(1 - \rho_0)}
\]
Note further that if $k_N < k_{N-1}^0$, less than $N - 1$ subsequent firms report once the marker holder failed to confess. $k_N < k_{N-1}^0$ holds for $z < z_c$ where

$$z_c \equiv \sqrt{N^3 - 2(2 - \rho_0)N^2 + N[8 - (8 - \rho_0)\rho_0] - 4(1 - \rho_0) - 2(1 - \rho_0) - N(N - 2 + \rho_0)}$$

The following proposition compares the deterrent impact of the marker and the no marker systems once the marker is lost after its holder’s decision to withhold the evidence and given that $k_N$ is provided for the subsequent firms regardless of the marker holder’s reporting decision:

**Proposition 3** Consider a non-transferable marker system:

i. When $z_c < z_a < z_b \iff \rho_0 > \frac{1}{N(N-2)^2}$ for $z_c < z < z_a$, the opaque and marker systems produce similar results in terms of ex-ante deterrence. For $z_a < z < z_b$ the marker system enhances ex-ante deterrence compared to the opaque. For $z > z_b$ and for $z < z_c$ the opaque system is superior in terms of ex-ante deterrence compared to the marker system.

ii. When $z_b < z_a \iff \rho_0 < \frac{1}{N(N-2)^2}$ the opaque and marker system produce similar results in terms of ex-ante deterrence if further $z_c < z < z_b$. Otherwise the opaque system produce better deterrence compared to the marker system.

**Proof**

See the Appendix.

Name $k_N^{1'} = \frac{(1-z)(1-\rho_0)}{1+z(N-1)} \left( k_N^{1''} = \frac{N(1-\rho_0)}{1+z(N-1)} \right)$ the minimum $k_f$ that makes the marker holder to come forward when every other (no) firm confesses following the marker holder’s denies to confess, that is when $z_c < z < z_b \left( z > z_b \right)$. It can be easily verified that $k_N^{1'} < k_N^{1''}$ and that $k_N > k_N^{1'}$. The expected fine for any subsequent applicant is equal to the marker holder’s actual fine:

$$\gamma_1(k_N^{1'}) = \frac{\gamma(k_N) + (N - 2)}{N - 1} = \frac{N - 1 + z + \rho_0 - z \rho_0}{N} = \gamma(k_N^{1'}, k_N) > \gamma(k_N)$$

Therefore, the actual reduced fine that one eligible subsequent applicant pays is always lower than that of the marker holder, when $z_c < z < z_b$. The certainty that the marker creates for the latter renders this applicant less demanding in terms of the leniency requested.

The previous analysis implies that the marker holder has strong incentive to come forward when all the others are going to do the same regardless of what the first in
line decides. If further $z > z_a$, which implies that under the no marker regime the first-to-door requires $k_0^0 (> k_0^0)$ to come forward, the marker acts as a mechanism that induces firms to compromise with a lower level of leniency, that is $\hat{\gamma}(k_{N}^1, k_{N})$. This results in a value of collusion which is lower under the marker system. When $z > z_b$ the marker holder recognizes the gravity of its reporting, as remaining silent implies universal non-reporting: confession by this firm needs a more generous treatment to be offered, a fact that reduces the overall lever of expected fines raising the value of collusion, and finally stabilizing the collusive agreement.

In the graph below (figure 1) the black, dashed and dotted lines represent the value of the cartel under the opaque, the non-scrolling and the scrolling marker system respectively, for the following values of the parameters: $\alpha = .15$, $\mu = 2$, $\delta = .9$, $N = n = 4$ and $\rho_0 = .7$. The common share of evidence $z$ is on the horizontal axis.

For $z_c = .026 < z < z_a = .38$ the value of collusion under the no marker, the non-scrolling marker systems coincide. For $.38 < z < z_b = .64$ the value of the cartel under the non-scrolling marker is lower, while for $z > .64$ the no marker system is superior. Comparing the two marker systems, the non-scrolling results in lower deterrence when $z > .64$. For every $z$ the scrolling marker produces worse deterrent results compared to the no marker regime.

Furthermore, we consider useful to discuss the robustness of proposition 3 with respect to the assumption that $k_N$ is offered to one subsequent applicant, regardless of
the marker holder’s reporting decision. If instead we suppose that $k_{N-1}^0$ or $k_1^0$, depending on the level of $z$, is offered to one informant, all $N-1$ subsequent firms are induced to confess following the marker recipient’s denial to come forward. When $z > (\leq) z_d \equiv \frac{1+\rho_0[2+4(N-2)+\rho_0-(2N-3)\rho_0^{-1}]}{2(N-1)(1-\rho_0)}$, $k_1^0 > (\leq) k_{N-1}^0$ holds and promising $k_f = k_1^0$ ($k_f = k_{N-1}^0$) to one subsequent applicant in case of the holder’s denial to confess, is enough to trigger universal reporting if additionally $k_f = k_N^{1'}$ and $k_s = k_N$ are offered to the marker recipient and to one subsequent informant respectively. The marker holder recognizes that remaining silent (confessing) implies a conviction likelihood equal to $\rho_{N-1}$ ($\rho_N = 1$). Thus, $k_f = k_N^{1'}$ is enough to induce reporting by the marker holder and $k_s = k_N$ secures that all subsequent firms also come forward. In both cases the resulting collusive value is $V_N^0$:  

\textbf{Proposition 4} If $k_f = k_1^0$ ($k_f = k_{N-1}^0$) is offered to one subsequent applicant when $z > (\leq) z_d$, the non-transferable marker regime produces at least equal deterrent outcome compared to the opaque system: if $z > z_a$ the non-transferable marker system has better deterrent results while if $z \leq z_a$ both systems are equivalent in terms of deterrence.

---

In equilibrium $k_f = k_N^{1'}$ and $k_s = k_N$ are offered to the marker holder and to one additional applicant respectively. Offering $k_1^0$ or $k_{N-1}^0$ is just a credible threat that induces the marker recipient to compromise with the least possible reward.
Figure 2 depicts the collusive values of the *opaque* (black line) and the non-scrolling (dashed line) *marker* systems under the parameter values of figure 1. In this case if $z > z_a = .38$ the non-transferrable marker induces the marker holder to compromise with the minimum possible reward resulting in a lower cartel value and consequently in better deterrent outcome.

Thus, persuading the marker recipient that remaining silent implies that all the others are going to confess is the key in order to induce confession with the lowest possible reward. In such case the non-transferrable marker regime minimizes the value of the cartel and as a result enhances deterrence compared to the *opaque* system.

5. **Concluding remarks**

As OECD (2012) states “it is often the case that co-operation from the second applicant is of particular value because its testimony and other evidence it presents can be used to corroborate the evidence submitted by the first applicant. Co-operation of subsequent applicants may contribute to proving additional facts either in terms of duration, product or geographic scope or the composition of the cartel”. In this paper we show that it is always optimal in terms of both deterrence and detection to induce every firm to reveal the evidence when being under inspection. This occurs because the impact of reduced overall fines is always lower than the effect that the higher likelihood of conviction has on the profitability of collusion and consequently on cartel deterrence.

It has been highlighted by leniency applicants’ representatives and practitioners that transparency and certainty are crucial parameters that should be taken into consideration for the implementation of the LP. ICN (2014) mentions that “a leniency applicant needs to be able to foresee with a high degree of certainty how it will be treated if it reports anticompetitive conduct and what the consequences will be if it does not come forward”.

The adoption of a marker system succeeds to eliminate the uncertainty, at least for the first-to-come applicant, reserving for the latter a position in line that secures its eligibility for a given lenient treatment. Here we show that offering information to applicants about the availability of leniency affects the effectiveness of the LP, depending on the way that the marker is secured: if the latter is repeatedly obtainable, regardless of the reporting decision by the marker holder, the marker system requires
higher overall level of fine reductions. This increases the profitability of collusion and consequently it hurts deterrence. Otherwise, a marker that is not transferable in case of failed-reporting by the marker holder could induce universal reporting with the first applicant to be less demanding in terms of leniency. In such case the marker acts as a mechanism that induces confession in instances where in the absence of it reporting would require more generous fine reductions to be offered.

Admitting that a marker is provided to the first applicant, as applied in major jurisdictions like the DoJ and EC, we show that it is always preferable to extend the eligibility for leniency to -at least- a second applicant in order to achieve universal reporting. This occurs because the disincentive to collude that emanates from the secured conviction surpasses the cost of lower overall fines needed for that. The latter may suggest an argument that justifies the EC’s (among others) practice to offer temperate fine reductions to subsequent informants.

Indeed, our assumptions do not allow evaluating other aspects of the marker system. The latter offers a possibility to infringers to come forward in an early stage, likely before the opening of an investigation, and to reserve a position in queue, providing the necessary time to gather sufficient information. This aspect is out of the present paper’s scope and requires additional model specifications in order to be examined.

References
European Commission, 2006. Commission notice on immunity from fines and reduction of fines in cartel cases
Appendix

Proof of lemma 4

The incentive to report for a marker holder always increases with the number of other reporting parties:

\[ 1 - \rho_1 \mu - (1 - \rho_{n-1} \mu) = (\rho_{n-1} - \rho_n \mu) \mu \]

\[ \frac{\partial (\rho_{n-1} - \rho_n \mu)}{\partial n} = \frac{k_f (1 - z) (1 - \rho_0) [1 + (N - 1)z]}{N^2} \]

Assume that \( N - 1 \) firms have previously claimed and subsequently denied the marker, and now the \( N^{th} \) firm contemplates whether receiving it or not. As in case of confession that firm will be the first and unique informant, it will confess iff \( k_f \geq k_1^0 \) is satisfied. If \( k_f < k_1^0 \), the \((N - 1)^{th}\) firm will face the same dilemma knowing that, on the one hand none of the previous firms has confessed, and on the other hand that in case it denies confession and the marker goes to the \( N^{th} \) firm, the latter will find optimal not to confess. Thus, the \((N - 1)^{th}\) firm also considers itself in the position of the first and unique informant, and since \( k_f < k_1^0 \) has been assumed not to hold, the \((N - 1)^{th}\) marker recipient will not confess either. Backwards induction yields that if


OECD, 2014. The Use of Marker in Leniency Programs. Working Party No. 3 on Co-operation and Enforcement


$k_f < k_1^0$ does not hold no firm will honor the marker: every previous marker recipient knows that its reporting implies a conviction rate at least equal to $\rho_1$ while remaining silent entails a conviction rate equal to $\rho_0$. Hence, if the last in line is not to report as the unique informant, no other has any incentive to come forward as a marker holder.

If $k_f \geq k_1^0$ is offered to this applicant, the $(N - 1)^{th}$ recipient recognizes that reporting entails conviction with probability $\rho_1$ or $\rho_2$. On the other side, if remaining silent the $N^{th}$ marker recipient confesses and conviction is the investigation outcome with probability $\rho_1$. The previous implies that the $(N - 1)^{th}$ marker recipient confesses if at least $k_f = k_1^0$ is offered. Under the same rationale, the first in line knows that its reporting implies that the cartel will be convicted with probability $\rho_n$ while remaining silent results in conviction with probability, at least, equal to $\rho_{n-1}$. Consequently the first marker holder has no incentive to remain silent, even if the action of its reporting entails a probability of conviction equal to $\rho_n$. Therefore, $k_f = k_1^0$ is necessary for at least one informant to exist and sufficient to persuade any marker holder to confess.

Consider the following $N = 4$ paradigm where $k_s = k_4 = 3(1 - z)(1 - \rho_0)$ is offered to the first subsequent applicant. If $k_f < k_1^0 = \frac{4(1-\rho_0)}{1+3(z+\rho_0-\rho_0 z)}$ the forth marker receiver denies confessing as the unique informant. The third recognizes that reporting entails conviction likelihood equal to $\rho_2$ (the forth confesses as subsequent if $k_s \geq k_2$, which is the case) while remaining silent implies that conviction occurs with probability $\rho_0$ and thus denies to confess. The second receiver knows that remaining silent implies that conviction occurs with probability $\rho_0$ while reporting implies conviction with probability $\rho_3$. Similarly the first in line knows that remaining silent implies that conviction occurs with probability $\rho_0$ while its reporting convicts the cartel with probability $1$. Consequently no one reports if $k_f < k_1^0$.

Consider now that $k_f \geq k_1^0 = \frac{4(1-\rho_0)}{1+3(z+\rho_0-\rho_0 z)}$ is offered to the marker holder. In this case the last receiver confesses as the unique informant. At the same time every firm confesses given that one other firm confesses ($k_s = k_4$). The third receiver knows that reporting implies conviction probability equal to $\rho_2$ while remaining silent implies that conviction occurs with probability $\rho_1$ (the last confesses). The second receiver knows that remaining silent implies that conviction takes place with probability $\rho_2$ while reporting convicts the cartel with probability $\rho_3$. Similarly the first in line
knows that remaining silent implies that conviction occurs with probability $\rho_3$ while reporting implies certain conviction. Therefore, for $k_f = k_1^0 = \frac{4(1-\rho_0)}{1+3(z+\rho_0-z\rho_0)}$ and $k_s \geq k_2$, the first marker holder confesses and all three subsequent firms do the same.

**Proof of proposition 2**

Substituting for $\gamma_1(k_n^1)$ and $\gamma(k_n)$ into $V_n^1$ and taking the derivative with respect to $n$ yields

$$\frac{\partial V_n^1}{\partial n} = -a\mu(1-z)(1-\rho_0)\left[\frac{N-2+(1-\rho_0)z(N-2)(N-1)+\rho_0[2+N(N-2)]}{N^2[1-\delta(1-a)][1+(N-1)(z+\rho_0-z\rho_0)]}\right]$$

The numerator’s expression in brackets is positive if $z > -\frac{N-2+\rho_0[2+N(N-2)]}{(N-2)(N-1)(1-\rho_0)}$ which always holds. Therefore, $\frac{\partial V_n^1}{\partial n} < 0$ always holds. As the value of the cartel reduces with the number of reporting parties, it is always optimal to set $N = n$, i.e. to induce reporting by all firms.

Using (4), (6), (7) and (8) $\hat{V}_N^0 < V_N^1$ holds for $\rho_0 \leq 1$. At the same time $\hat{V}_N^0 < V_N^1$ always holds as apart from $k_1^0$ which is offered to the first informant in the marker case (the same if offered only to the first informant in the no marker case) some additional leniency is required for one subsequent applicant for additional informants to exist.
Proof of proposition 3

First, \( z_a < (>) z_b \) holds for \( \rho_0 > (<) \frac{1}{N(N-2)^2} \). Also the \( N-1 \) subsequent firms confess even if the marker holder remains silent only when \( k_N \geq \max \{ k_{N-1}^0, k_1^0 \} \Leftrightarrow z_c < z < z_b \). If \( z < z_c \) and marker holder denies to confess, at most \( N-2 \) reporting firms exist. Consider that \( \rho_0 > \frac{1}{N(N-2)^2} \). For \( z_c < z < z_b \) the marker holder knows that remaining silent implies that \( N-1 \) others are going to report and its incentive to confess is:

\[
1 - \gamma_1(k_f)\mu \geq (1 - \rho_{N-1}) + \rho_{N-1}(1 - \mu)
\]

which yields

\[
k_f \geq k_N^{1'} = \frac{(1-z)(1-\rho_0)}{1 + z(N-1)}
\]

For \( z_c < z < z_b \) the value of the cartel becomes

\[
\frac{1 - a\mu\hat{\gamma}(k_N^{1'}, k_N)}{1 - \delta(1 - a)} = \hat{\nu}_N^0
\]

where \( \hat{\gamma}(k_N^{1'}, k_N) = \frac{\gamma_1(k_N^{1'}) + \gamma(k_N) + (N-2)}{N} \), \( \gamma(k_N) = \frac{1 + (z + \rho_0 - \rho_0\rho_0)(N-1)}{N} < 1 \) and \( \gamma_1(k_N^{1'}) = \frac{N-1 + (z + \rho_0 - \rho_0\rho_0)}{N} > \gamma(k_N) \). Therefore, under the non-scrolling marker the collusive value is equal to the value of the cartel under the no marker regime for \( z_c < z < z_b \). For \( z_a < z < z_b \) the collusive value under the no marker system is \( \hat{\nu}_N^0 > \hat{\nu}_N^0 \).

If \( z < z_c \) the marker holder knows that remaining silent implies that less than \( N-1 \) subsequent are going to come forward and therefore the holder’s confession requires \( k_f > k_N^{1'} \). Consequently the collusive value in this case is always greater than \( \hat{\nu}_N^0 \).

If \( z > z_b \) the unique marker holder recognizes that remaining silent implies that firms will coordinate on the most profitable equilibrium where no investigated firm confesses and that the conviction probability will be \( \rho_0 \) (see lemma 3). The marker holder confesses only if \( 1 - \gamma_1(k_f)\mu \geq 1 - \rho_0 + \rho_0(1 - \mu) \) or if \( \rho_0 \geq \gamma_1(k_f) \) which yields

\[
k_f \geq k_N^{1''} = \frac{N(1 - \rho_0)}{1 + z(N - 1)}
\]

The value of the cartel is
$$\tilde{V}_N^1 = \frac{1 - a\mu \tilde{\gamma}(k_N^{1''}, k_N)}{1 - \delta(1 - a)}$$

where $\tilde{\gamma}(k_N^{1''}, k_N) = \frac{\gamma_1(k_N^{1''}) + \gamma(k_N) + (N-2)}{N} = \frac{N(N-(1-\rho_0)(2-z)+(1-z)(1-\rho_0))}{N^2}, \quad \gamma_1(k_N^{1''}) = \rho_0 < \gamma(k_N).$ Notice that the expected fine is lower under the marker regime, that is

$$\tilde{\gamma}(k_0^{1''}) - \gamma(k_N^{1''}, k_N) = \frac{(N-1)(1-\rho_0)(1-\rho_0)(1 + (N-1)(1-\rho_0)z + \rho_0(2N-1))}{N^2 [1 + (N-1)(1-\rho_0)z + \rho_0(N-1)]} > 0$$

which implies $\tilde{V}_N^1 > \tilde{V}_N^0.$