Effects of leniency programs on cartel stability*

Evguenia Motchenkova†

Abstract

This paper studies the effect of leniency programs on the stability of cartels under two different regimes of fines, fixed and proportional. We analyze the design of self-reporting incentives, having a group of defendants. Moreover, we consider a dynamic setup, where accumulated (not instantaneous) benefits and losses from crime are taken into account.

We obtain that cartel occurrence is less likely if the rules of the leniency programs are more strict and the procedure of application for leniency is more confidential. Moreover, we conclude that, when the procedure of application for leniency is not confidential and penalties and rate of law enforcement are low, leniency may even increase duration of cartel agreements. Another counterintuitive result is that under a fixed penalty scheme the introduction of a leniency program cannot improve the effectiveness of antitrust enforcement when the procedure of application for leniency is not confidential.

JEL-Classification: K21, L41

Keywords: Antitrust Policy, Antitrust Law, Self-reporting, Leniency Programs

1 Introduction

This paper analyzes the effects of leniency programs by employing a game between two firms, which participate in a cartel agreement and decide on the optimal time of revealing the information about the cartel to the antitrust authority. The enforcement problem we study has several ingredients. Firstly, we analyze the design of self-reporting schemes, where we have a group of defendants instead of a single defendant. Secondly, we consider a dynamic set-up, where accumulated benefits and losses from crime are taken into account. Leniency programs allow for complete or partial exemption from the fine for firms that reveal information about

*The author thanks Eric van Damme, Peter Kort, Massimo Motta and Giancarlo Spagnolo for stimulating discussions and valuable comments.

†Tilburg University, Department of Econometrics & OR and CentER, P.O.Box 90153, 5000 LE Tilburg, The Netherlands, e-mail: E.I.Motchenkova@uvt.nl.

the cartel to the antitrust authority. It is intuitively clear that a legally sanctioned opportunity for costless self-reporting changes the nature of the game played between the antitrust authority and the group of firms. To analyze the impact of this opportunity on cartel stability we apply tools of timing games. In particular, we study a dynamic game of the preemption type.

Leniency programs have been recently introduced in the European antitrust legislation and have quite a long history in the US. "Leniency programs" grant total or partial immunity from fines to firms that collaborate with the authority. To be more precise, leniency is defined as a reduction of the fine for firms, which cooperate with the antitrust authority by revealing information about the existence of the cartel before the investigation has started, or by providing additional information that can help to speed up the investigation. Leniency programs work on the principle that firms, who break the law, might report their crimes or illegal activities if given proper incentives.

In the US the first Corporate Leniency Program was introduced in 1978. Then it was refined and extended in August 1993. Later the Antitrust Division of the US Department of Justice revised its Corporate Leniency Program to make it easier for and more attractive to companies to come forward and cooperate with the Division[1]. Three major revisions were made to the program, namely, amnesty is automatic if there is no pre-existing investigation, amnesty may still be available even if cooperation begins after the investigation is underway, and all officers, directors, and employees who cooperate are protected from criminal prosecution. As a result of these changes, the Amnesty Program is the Division's most effective generator of international cartel cases. Moreover, the revised Corporate Amnesty Program has resulted in a surge in amnesty applications. Under the old amnesty policy the Division obtained roughly one amnesty application per year. Under the new policy, the application rate has been more than one per month. In the last few years, cooperation resulting from amnesty applications led to scores of convictions of over $1 billion in fines[2].

In Europe the first Leniency Programs were introduced in 1996. The modified Leniency program introduced by the EC in 2002 gives complete immunity from fines to firms, which were the first to submit evidence about the cartel to the antitrust authority. Moreover, partial reduction of fines (approximately by 50%) should take place even if firms reveal information

[1] Spratling Gray R. (Deputy assistant Attorney, Antitrust Division, US Dept. of Justice). "The corporate leniency policy: Answers to Recurring Questions", presented at the Spring 1998 ABA Meeting (Antitrust section), available for download at http://www.usdoj.gov/atr/public/speeches/1626.htm

[2] OECD report 2002, "Report on the Nature and Impact of Hard Core Cartels and Sanctions Against Cartels."

after an investigation has started. Similar programs have been introduced in 2002 in the UK and other European countries.

There is some empirical evidence that Leniency programs improve welfare by sharply increasing the probability of interrupting collusive practices and by shortening the investigation. In the US, for example, the fines collected in 1993 almost doubled the ones in 1992, which can be connected with the major modification of leniency programs. However, there are also other effects of leniency programs, which are now difficult to identify in empirical studies due to the absence of data. For example, questions of how introduction of leniency programs would influence cartel stability and duration of cartel agreement, or whether leniency facilitates collusion or reduces it, still require deeper investigation. In this paper we give some insights by analyzing these problems.

Other contributions that analyze optimal policies for the deterrence of violations of antitrust law in the presence of leniency schemes are Motta and Polo (2003), Spagnolo (2000), Malik (1993) or Hinloopen (2003). Most papers on leniency employ a discrete time framework. However, proportional penalty schemes that most closely reflect current antitrust rules were not analyzed in the discrete time repeated games models so far. Our paper studies the problem of how an additional enforcement instrument, such as a leniency program, influences the stability of cartels under two different regimes of fines, fixed and proportional. We analyze a setting with a proportional penalty scheme employing a continuous time dynamic game, in which accumulated gains from price-fixing is the state variable. We investigate intertemporal aspects of this problem using dynamic optimal stopping models and tools of dynamic continuous time preemption games. In this way we extend the existing literature.

It should be stressed that a legally sanctioned opportunity for costless whistle-blowing changes the game played between the antitrust authority and the group of firms, compared to a setting where leniency is not available. Intuitively, this opportunity should reduce cartel stability and increase the incentives for firms to reveal the cartel. In this paper we investigate the effects of leniency programs on the behavior of firms participating in price-fixing agreements. The main finding of the paper is that well designed leniency may reduce duration of cartel agreements but this result is not unambiguous. Under strict antitrust enforcement, while penalties and the rate of law enforcement are high, the possibility to self-report and be exempted from the fine increases the incentives for firms to stop cartel formation, and, hence, reduces the duration of cartels. However, when the procedure of application for leniency is not confidential and penalties and rate of law enforcement are low, introduction of leniency programs may, on the contrary, facilitate collusion. Under a fixed penalty scheme, even in the presence of leniency, the efficiency of cartel deterrence (in terms of reduction of duration of cartel agreements) depends only on the amount of the fine and the probability of law enforcement. We show that "too lenient " leniency programs may facilitate collusion, when penalties are fixed and fall below a certain threshold.

We distinguish two regimes with respect to the rules of leniency programs and application procedure. The first regime corresponds to more strict enforcement, i.e. only the firm, which is the first to self-report, is eligible for complete exemption from the fine and the application procedure is strictly confidential. The second firm bears either the full fine or, if it provides sufficient evidence, it can be exempted from up to 50% of the fine. This set up most closely reflects the rules of current guidelines for reduction of fines for firms that cooperate with antitrust authorities and reveal information about existing cartels[3]. The second regime corresponds to the case where the rules of antitrust enforcement are not too strict (more lenient). In this case also the firm, which is the second to self-report, obtains partial exemption from the fine. Moreover, the antitrust authority makes the application procedure publicly observable. Comparison of these two regimes implies that, if the rules of leniency programs and the procedure of application for leniency are more strict, cartel occurrence is less likely.

A number of earlier papers have studied the problem of self-reporting. Malik (1993) and Kaplow and Shavell (1994) were the first to identify the potential benefits of schemes which elicit self-reporting by violators. They conclude that self-reporting may reduce enforcement costs and improve risk-sharing, as risk-averse self-reporting individuals face a certain penalty rather than the stochastic penalty faced by non-reporting violators. A similar paper in this field is Innes (1999), who considers environmental self-reporting schemes.

The use of leniency programs in antitrust has been extensively studied by Motta and Polo (2003). They show that such programs might play an important role in the prosecution of cartels provided that firms can apply for leniency after an investigation has started. They conclude that, if given the possibility to apply for leniency, the firm might well decide to give up its participation in the cartel in the first place. They also find that leniency saves resources for the authority. Finally, their formal analysis shows that leniency should only be used when the antitrust authority has limited resources, so that a leniency program is not unambiguously optimal. The paper by Motta and Polo (2003) is closely related to the paper by Spagnolo (2000). He shows that only courageous leniency programs that reward self-reporting parties

[3] See OECD report (2002).

may completely and costlessly deter collusion, while moderate leniency programs that reduce or cancel sanctions for the reporting party cannot affect organized crime.

A next attempt to study the efficiency of leniency programs in antitrust enforcement was made in Feess and Walzl (2003). They compared leniency programs in the EU and the USA. For that purpose they constructed a stage-game with two self-reporting stages, heterogeneous types with respect to the amount of evidence provided, and ex post asymmetric information. Their analysis shows that self-reporting schemes are much more promising for criminal teams than for single violators, since strategic interactions between team members lead to increased expected fines, and reduce the frequency of violations. Hence, their model once again confirms the effectiveness of leniency programs in the fight against cartels.

The paper is organized as follows. Section 2 describes the basic model. In section 3 we consider the decision of a single firm, which participates in a cartel agreement, about the optimal stopping time, i.e. the moment of revelation of information about the violation to the authority in the absence of leniency programs. Further, in section 4 we introduce more players and study a timing game with two identical firms forming the cartel after leniency programs are introduced. In this part we suggest a new approach to analyze the efficiency of the leniency programs that differs from the earlier papers and that is based on the Reiganum-Fudenberg-Tirole Model. Reiganum (1981) and Fudenberg and Tirole (1985) applied timing games to a technology adoption problem. We apply a similar procedure to cartel formation game between two firms in the presence of a leniency program. Section 5 analyzes the effects of leniency while the leniency program is less strict. In Section 6 we solve the game in case penalty is fixed and compare it with the result under proportional penalty. Section 7 deals with an extension of the model of section 4 by including dynamic price competition and tacit collusion. The last section summarizes the results and suggests directions for future work.

2 Optimal stopping model. The general setup

We introduce the basic ingredients of the intertemporal optimization problem of an expected profit-maximizing firm, which participates in an illegal cartel. The key variable is the accumulated gains from prior criminal offences, $w(t)$, (in case of a cartel, these offences are price-fixing activities). Further we will call $w(t)$ the value of collusion.

Let us consider an industry with 2 symmetric firms engaged in a price-fixing agreement. Assume that they can agree and increase prices from $p^c = c$ to $p^m > c$ each, where c is the constant marginal cost in the industry. Since firms are symmetric, each of them has equal weight in the coalition and, consequently, total cartel profits will be divided equally among them. In a game theoretic model we assume that there is a possibility of strategic interaction between the firms in the coalition in the sense that they can break the cartel agreement by self-reporting. By doing this we allow for the possibility for the firms to betray the cartel and this influences the internal stability of the cartel.

The instantaneous monopoly profit in the industry under consideration is denoted by π^m. Consequently, since the firms are assumed to be symmetric, the instantaneous profit per firm will be $\frac{\pi^m}{2}$.

We consider two cases: the case where the penalty, s, is constant over time, i.e. $s(t) = F^n$, and the case where the penalty is a fraction of the accumulated gains from price-fixing activities for the firms. In the latter case the penalty is represented by the expression $s(t) = \alpha w(t)$, where α is the scale parameter of the penalty scheme. This setup will also allow us to compare the efficiency of fixed and proportional penalty schemes. Both of them are currently used in the sentencing guidelines of different countries[4].

The main feature of a leniency program is the reduction of the fine (or complete exemption from the fine) for the firm that first reveals the information about the existence of the cartel. To be more precise, in the model we assume the following set-up. If one of the firms reports the cartel, then this firm pays no fine, $s^L = 0$, while the other firm will pay the normal fine, s^n, that (according to current sentencing guidelines for violations of antitrust law) can be approximated by the amount of 10% of overall turnover of the enterprise. The current rules also imply that, if the second firm decides to cooperate before the investigation is completed, the fine for this firm will be reduced by approximately 25%, $s^F = 0.75s^n$ (or $0.75 * 10\%$ of overall turnover of the enterprise). Moreover, if both firms report the cartel simultaneously, then each of them pays the reduced fine, $s^M = 0.5s^n$. These rules are roughly consistent with partial immunity clauses that often apply if more than one cartelist reports[5].

The rate of law enforcement by the antitrust authority equals $\lambda \in (0, 1]$. This variable denotes the instantaneous probability that the firm is checked by antitrust authority and found guilty.

[4] OECD report 2002, "Report on the Nature and Impact of Hard Core Cartels and Sanctions Against Cartels."

[5] Moreover, Apesteguia, Dufwenberg and Selten (2003) use a similar mechanism to design one of the treatments in their experimental paper, which studies the effects of leniency on the stability of a cartel. Feess and Walzl (2003) also consider partial reduction of fines for both firms in case of simultaneous self-reporting.

Given this set-up, firms, participating in the cartel agreement, decide on the optimal stopping time, i.e. the moment of revelation of information about the violation to the authority. An alternative way of stopping may be described in terms of quitting the cartel without reporting to the antitrust authority. We assume that, after the cartel has been discovered due to self-reporting by one or both firms, collusion stops forever and, consequently, the stream of illegal gains also stops. However, if cartel is discovered due to the efforts of antitrust authority, collusion does not stop and firms renew their agreement in the next period. Basically, firms would not renew the agreement, if one of them betrayed the other.

The expected penalty if the firm, which was participating in the cartel, is caught at date t is given by $\lambda s(t)$. The discount rate is denoted by r. The value for the firm from revealing the cartel at time t is $V(t)$. This variable also can be considered as an option value of self-reporting. We assume that there are two identical firms that form a cartel. The infinite planning horizon is considered, on which the risk-neutral firms maximize their value at discount rate r (> 0).

3 Benchmark. Optimal stopping model without leniency

To study the effects of leniency programs on cartel stability, we, first, consider a benchmark case, where leniency is not available, i.e. the firms act without taking into account the possibility of self-reporting. Basically, in this case we consider an optimal stopping model with a single decision maker, where the representative firm maximizes its expected gains from price-fixing with respect to time. Second, in section 4 we move to the setting where the antitrust authority introduces leniency. In that case the dynamic interactions between two firms, which form a cartel but can also betray it, are modelled by employing tools of preemption games.

In the deterministic case the value of collusion changes according to the following law:

$dw = \frac{\pi^m}{2}e^{-rt}dt$ and $w(0) = 0$. This implies that $w(t) = \int\limits_0^t \frac{\pi^m}{2}e^{-rs}ds = \frac{\pi^m}{2r}(1 - e^{-rt})$

The value of stopping the cartel at time T, $St(T)$, is determined as an integral over time of instantaneous expected gains from collusion before time T. It should be positively related to the instantaneous profits from price-fixing before reporting, $\frac{\pi^m}{2}$, and negatively related to the instantaneous expected penalty, $\lambda s(t)$. We assume that cartel formation stops only in case firms decide to quit the cartel or self-report to the antitrust authority, while firms always renew collusive agreement after they are caught and punished by the antitrust authority without cooperation of cartel members. So, the value of stopping the cartel for each firm in the absence of leniency programs will be determined according to the following formula:

$$St(T) = \int_0^T [\frac{\pi^m}{2} - \lambda s(t)] e^{-rt} dt \tag{1}$$

In case the fine is proportional to the accumulated illegal gains from price-fixing, expression (1) will have the following form

$$St(T) = \max_T \int_0^T [\frac{\pi^m}{2} - \lambda \alpha w(t)] e^{-rt} dt, \tag{2}$$

where $w(t) = \int_0^t \frac{\pi^m}{2} e^{-rl} dl$.

To find the optimal time of stopping the cartel, we differentiate (2) with respect to T and obtain $\frac{\partial F(w)}{\partial T} = \frac{\pi^m}{2} e^{-rT} - \frac{\alpha\lambda\pi^m}{4r} e^{-rT} + \frac{\alpha\lambda\pi^m}{4r} e^{-2rT} = 0$. This implies that the optimal stopping time for the single firm, which takes a decision whether to quit the cartel or to continue collusion, is given by

$$T^* = \frac{\ln(\frac{\alpha\lambda}{\alpha\lambda - r})}{r} \tag{3}$$

This result coincides with the solution of a dynamic game , in which 2 symmetric firms choose whether to stop or continue cartel at each instant of time. Using the backward induction argument and assuming that in case of multiple equilibria of the matrix game played at each instant of time firms choose for the equilibrium with the highest payoff, we obtain that the game stops at T^*. At this time both firms decide to quit the cartel simultaneously.

Expression (3) shows that the optimal time of stopping the cartel decreases when either the probability or the severity of punishment increases. The higher the expected penalty, the earlier the firm decides to quit the cartel agreement, since $\frac{\partial T^*}{\partial \alpha} < 0$ and $\frac{\partial T^*}{\partial \lambda} < 0$.

4 Preemption game with leniency

Now we describe a timing game of the preemption type played between two symmetric firms. The leader in this game (i.e., the firm which is the first to self-report) has the advantage of complete exemption from the fine, i.e. $s^L = 0$. Moreover, since firms are identical it seems natural to consider symmetric strategies.

First, we consider a setting where firms cannot respond immediately to the actions of their rivals. Following the rules of application for leniency currently used by most antitrust authorities, the information about applications is kept confidential. This information normally does not become public knowledge immediately after the firm has applied for leniency. That is why in this section we analyze a setting where it is not possible to react instantaneously. The firm, which self-reports as second, can be exempted only for less than 50% of the fine, while the leader gets complete immunity from fine.

Next, in section 5 we compare the regime described above with the case where the rules of leniency programs are less strict and the procedure of application for leniency is less confidential. We model this by relaxing the assumption that instantaneous reaction is not possible. That is, we consider a setting where firms can respond immediately to their rival's decisions. This implies that actions of the firms are perfectly observable and the procedure of self-reporting is instantaneous (does not take any time). Clearly, in this case simultaneous self reporting is possible. However, this could be a too strong assumption for the model that describes leniency programs, since in most cases the procedure of application for leniency is very confidential. Nevertheless, we consider it in order to compare the results of these two regimes and show that if the rules of leniency programs and the procedure of application for leniency were more strict, cartels would be less likely.

We study a continuous time preemption game and employ the feedback equilibrium solution concept in order to solve it. First, we determine the payoffs and the objective functions for the first mover (leader), the second mover (follower), and in case of simultaneous self-reporting. Next, we determine optimal stopping times for each case. Finally, we derive the feedback equilibrium of the preemption game with leniency.

We assume that after the first firm has reported about the existence of the cartel to the authority, the cartel stops, and consequently, the stream of illegal gains also stops. In case of complete information about the actions of the rival the best response of the second firm would be to cooperate and reveal the cartel immediately after the first firm (the leader) does so. In addition, our approach represents a quite extreme form of preemption in that the follower firm loses entirely its chance to be completely exempted from the fine if it is forestalled by the leader. In the general setting the leader reports at the same time or before the follower, i.e. $0 \leq T_L \leq T_F$, where T_L and T_F are the optimal stopping times for the leader and follower, respectively.

Given the times T_L and T_F, and due to the special structure of the game, the value of the leader equals the integral over time of the instantaneous illegal gains from price-fixing, $\frac{\pi^m}{2}$, less the instantaneous expected penalty, $\lambda s(t)$. The additional term $-s^L e^{-rT_L}$ reflects the discounted value of the fine that has to be paid by the leader after the cartel is discovered due to self-reporting. By construction of the game this value equals zero. Hence, the value of the leader when $T_L < T_F$ is given by

$$V_L(T_L,T_F) = \int_0^{T_L} (\frac{\pi^m}{2} - \lambda s(t))e^{-rt}dt - s^L e^{-rT_L} = \int_0^{T_L} (\frac{\pi^m}{2} - \lambda s(t))e^{-rt}dt - 0. \qquad (4)$$

After time T_L , i.e. after the cartel was reported to the authority, the flow of illicit gains stops, so the exact value of T_F is not relevant for the determination of $V_L(T_L,T_F)$ and $V_F(T_L,T_F)$.

In the same way the value of the follower, $V_F(T_L,T_F)$, can be derived. The follower value is given by the integral over time of the instantaneous illegal gains from price-fixing, $\frac{\pi^m}{2}$, less the instantaneous expected penalty, $\lambda s(t)$. The term $-s^n(T_L)e^{-rT_L}$ reflects the discounted value of the normal (full) fine that has to be paid by the follower after the cartel is discovered[6]. Hence, the follower value, when $T_L < T_F$, is given by

$$V_F(T_L,T_F) = \int_0^{T_L} (\frac{\pi^m}{2} - \lambda s(t))e^{-rt}dt - s^F(T_L)e^{-rT_L} = V_L(T_L,T_F) - s^n(T_L)e^{-rT_L}. \qquad (5)$$

Similarly, the value of the firm in case of simultaneous self-reporting is determined by expression (6) below. Recall that in case of simultaneous self-reporting both firms pay 50% of the normal fine.

$$V_M(T_L,T_F) = \int_0^{T_c} (\frac{\pi^m}{2} - \lambda s(t))e^{-rt}dt - \frac{1}{2}s^n(T_c)e^{-rT_c}, \quad \text{iff} \quad T_L = T_F. \qquad (6)$$

Since firms are completely symmetric and the flow of illicit gains stops after one of the firms reports, it is not essential to distinguish between the follower and the leader. Hence, in order to emphasize symmetry, we call them further firm 1 and firm 2. First, we define $T_c^* = \arg\max_{T_c} V_M(T_c,T_c)$ and $T_L^* = \arg\max_{T_1 \,:\, (T_1 \le T_2)} V_L(T_1,T_2)$. Note also that $V_L(T_1,T_2) = V_M(T_1,T_2)$, when $T_1 = T_2$. From expressions (4) - (6) it is clear that $V_F(T_1,T_2) \le V_M(T_1,T_2) < V_L(T_1,T_2)$ for any $T_1 < T_2$.

4.1 Confidential Leniency Programs

In this subsection we analyze a model, where it is not possible for firms to react instantaneously to the actions of their rivals, i.e. the rules and procedure of application for leniency are very strict. This corresponds to the first regime mentioned in the introduction of the paper, namely

[6]Note that the results of the analysis below are valid for any $s^F \in (\frac{1}{2}s^n, s^n]$.

the regime with more strict enforcement, i.e. only the firm, which formally self-reports the first (even if second firm also does it voluntary), is eligible for complete exemption from the fine. The second firm bears the full fine (even it is several seconds later to self-report than the first one), and the application procedure is strictly confidential.

The objective functions of the firms can be described as follows. In a feedback equilibrium the leader (firm 1) takes into account that its stopping decision affects the decision of the follower (firm 2). However, for this particular problem it holds that the decision of the follower does not influence the value of the leader's payoff after he decides to reveal the cartel, see (4) or (7). This implies that the expressions (7) and (8) below do not depend on reaction of the follower.

Hence, we can define the following three functions

$$L(T) = \int_0^T (\frac{\pi^m}{2} - \lambda s(t))e^{-rt}dt, \tag{7}$$

$$F(T) = \int_0^T (\frac{\pi^m}{2} - \lambda s(t))e^{-rt}dt - s(T)e^{-rT}, \tag{8}$$

$$M(T) = \int_0^T (\frac{\pi^m}{2} - \lambda s(t))e^{-rt}dt - \frac{1}{2}s(T)e^{-rT}. \tag{9}$$

The function $L(T)$ $(F(T))$ is equal to the expected discounted value at time $t = 0$ of the leader (follower) when the leader reports at time T. $M(T)$ is the discounted value at time $t = 0$ of the firm when there is simultaneous self-reporting at time T.[7]

Here we assume that the firms can not react instantaneously, i.e. only a lagged reaction is possible. The implication is that the payoff of $M(t)$ is no longer available for the follower. Therefore, given the expressions (7), (8) and (9), in equilibrium the following inequalities hold

$$L(t) > M(t) > F(t) \quad for\ all \quad t \in (0, \infty) \tag{10}$$

Note that $L(0) = F(0) = M(0)$. Furthermore, we denote by π the profits of an infinitely lasing cartel.

To find feedback equilibria of this model we consider the dynamic timing game. At each instant of time t the following simultaneous move matrix game is played (see table 1 below):

[7]Note that the results of the analysis of this section are valid for any $s(T) \in (\frac{1}{2}s^n(T), s^n(T)]$. Which is in line with current leniency rules (see OECD report 2002).

	Self-report	Not self-report
Self-report	$(M(t), M(t))$	$(L(t), F(t))$
Not self-report	$(F(t), L(t))$	repeat game

Table 1. Payoffs and strategies of matrix game played at time t.

We also denote by π the value of the infinitely lasting cartel. In case of proportional penalty this value is given by the following expression: $\pi = \int_0^\infty (\frac{\pi^m}{2} - \lambda\alpha w(t))e^{-rt}dt$. This value will be extensively used in the proof of proposition 1.

The game is played at time t if no firm has reported about the existence of the cartel so far. Playing the game costs no time and if firm 1 chooses row 2 and firm 2 column 2 the game is repeated. If necessary the game will be repeated infinitely many times. Clearly, in this matrix game the outcomes, which are simultaneous self-reporting by both firms - (S, S) and the decision not to reveal the cartel by both players - (N, N), can arise as a Nash Equilibrium in pure strategies. The result depends on the magnitude of the maximal simultaneous self-reporting value and the value of the profits in case of infinitely lasting cartel.

The result of this analysis suggests that after introduction of leniency programs antitrust enforcement appears to be more efficient than in the absence of leniency. Even in combination with moderate penalties it leads to immediate self-reporting by both firms in the beginning of the game. Depending on the severity of punishment, two possible outcomes can arise. Either both firms report the cartel simultaneously in the beginning of the game, or the cartel can last forever. The results of the analysis in the setting with proportional penalty are summarized in the next proposition. Later, in section 6 of the paper, we compare these results to the solution of the model with fixed penalty.

Proposition 1 *For setting with proportional penalty, the outcome of the game with leniency, where firms cannot react instantaneously to the actions of their rivals, is*

immediate simultaneous self-reporting, i.e. $(t_1^*, t_2^*) = (0, 0)$, *if* $\alpha\lambda > r$,

or cartel forever, i.e. $(t_1^*, t_2^*) = (\infty, \infty)$, *if* $\alpha\lambda < r$.

Proof: See Appendix 1.

Let us compare this result with the conclusion of the model, where firms take decisions in the absence of leniency programs. Recall that from expression (3) we obtain that if the optimal stopping time of the model without leniency (T^*) exists, then $T^* > 0$ for any values of parameters of the model ($\alpha \in (0, \infty), \lambda \in (0, 1], r \in (0, 1]$ such that $\alpha\lambda > r$), since expression

$\frac{\alpha\lambda}{\alpha\lambda - r}$ is always greater than one, when $\alpha\lambda > r$. In the game with leniency we have immediate self-reporting by both firms in the beginning of the game, when $\alpha\lambda > r$. This result suggests that antitrust enforcement after introduction of leniency programs is more efficient than in the absence of leniency. Hence, strictly confidential leniency programs improve upon the situation without leniency.

5 Non-confidential leniency programs

In this section we discuss the preemption game with leniency under the assumption that firms can react instantaneously to the actions of their rivals. In particular, this implies that here we study the second regime, mentioned in the introduction, namely, where the rules of antitrust enforcement are not too strict and the procedure of application for leniency is less confidential.

First, we determine the objective functions of both players in case there is a first mover (leader) and a second mover (follower), and in case of simultaneous self-reporting for proportional penalty setting. Next, we find optimal stopping times for each case. Finally, we derive the feedback equilibrium of the preemption game with leniency under the assumption that instantaneous reaction is possible.

Now, we describe in more detail the derivation of the optimal stopping times for the leader and in case of simultaneous self-reporting, T_L and T_c , in a setting where the penalty is proportional to the amount of illicit gains, i.e. $s(t) = \alpha w(t)$ for all $t \in [0, \infty)$.

In this case the value of the leader is obtained by substituting $s(t) = \alpha w(t)$ into expression (4):

$$L(T) = \int_0^T (\frac{\pi^m}{2} - \lambda\alpha w(t))e^{-rt}dt. \tag{11}$$

Similarly, after substitution of $s(t) = \alpha w(t)$ into expression (5) the value of the follower equals

$$F(T) = \int_0^T (\frac{\pi^m}{2} - \lambda\alpha w(t))e^{-rt}dt - \alpha w(T)e^{-rT}. \tag{12}$$

Finally, the value of simultaneous self-reporting is determined by

$$M(T) = \int_0^T (\frac{\pi^m}{2} - \lambda\alpha w(t))e^{-rt}dt - \frac{1}{2}\alpha w(T)e^{-rT} \tag{13}$$

where $w(t) = \int_0^t \frac{\pi^m}{2}e^{-rs}ds$, $w(0) = 0$. For further analysis, similarly to the section 4, we define

T_c to be the optimal time of simultaneous self-reporting and $T_L(T_F)$ the optimal time of self-reporting by the leader (follower).

Following the benchmark model, taking the derivative of (11) with respect to T and equalizing it to zero, we obtain the optimal stopping time of the leader, i.e. the time T_L, which maximizes $L(T)$.

$$T_L = \frac{\ln(\frac{\alpha\lambda}{\alpha\lambda - r})}{r} = T^* = \arg\max_T \; St(T) \tag{14}$$

The necessary condition for a maximum is satisfied since $\frac{\partial L^2(T_L)}{\partial^2 T} < 0$.

From expression (14) we obtain, that the earliest time of revelation (i.e. breaking the cartel agreement) by one of the firms will decrease when either α or λ increases. This result is quite intuitive, because it means that the cartel stability should be reduced when either severity or probability of punishment increases. At the same time, the effect of an increase in the discount rate on the optimal time of self reporting gives $\frac{\partial T_L}{\partial r} < 0$. Hence, the firms will find it more attractive to stop earlier if the discount rate is higher, since future illicit gains become less valuable.

Similarly to the above analysis we take the derivative of (13) with respect to T and equalize it to zero. In this way we obtain the optimal stopping time in case both firms report the cartel simultaneously, i.e. the time T_c which maximizes $M(T)$.

$$T_c = \frac{\ln(\frac{2\alpha(\lambda - r)}{2\alpha\lambda - 2r - \alpha r})}{r} \tag{15}$$

The necessary condition for the existence of maximum is satisfied since $\frac{\partial M^2(T_c)}{\partial^2 T} < 0$.

From this expression we obtain, that the earliest time of revelation (i.e. breaking the cartel agreement) by both firms simultaneously will decrease (move closer to the origin) when either α or λ increase. So, the cartel stability is lower when either severity or probability of punishment increase.[8]

Moreover, the solution of this problem exists only when $\lambda > r$ (i.e. the rate of law enforcement is higher than the discount rate) and $\alpha\lambda > \frac{r(2+\alpha)}{2}$ (i.e. the coefficient of expected penalty is greater than the sum of the discount rate and half of the product of the scale parameter and discount rate)[9]. In other words, the expected penalty is high enough to outweigh the current

[8]It also should be mentioned that $\ln(\frac{2\alpha(\lambda-r)}{2\alpha\lambda-2r-\alpha r}) > 0$ only if $\alpha < 2$. For any $\alpha \geq 2$, we obtain $\ln(\frac{2\alpha(\lambda-r)}{2\alpha\lambda-2r-\alpha r}) \leq 0$, consequently, $t_c^* = 0$, since feasible domain for $t_c \in [0, \infty)$.

[9]Note, that $\alpha\lambda > \frac{r(2+\alpha)}{2}$ implies $\alpha\lambda > r$. Hence, existence of non-negative value for optimal stopping time of simultaneous self-reporting in the presence of leniency implies existence of non-negative optimal stopping time in case when leniency is not available.

benefits from crime compared to the future penalties. Comparison of expressions (14) and (15) implies the following lemma.

Lemma 2 *Given* $\lambda > r$ *and* $\alpha\lambda > \frac{r(2+\alpha)}{2}$, *there exist* $T_L = \arg\max_T L(T) = T^* = \arg\max_T St(T)$ *and* $T_c = \arg\max_T M(T)$ *such that*

$$T^* < T_c, \text{ when } r < \alpha\lambda < 2r \quad \text{and} \quad T^* > T_c, \text{ when } 2r < \alpha\lambda.$$

Proof: See Appendix 2.

This result shows that when the multiplier of the expected penalty is lower than twice the discount rate, in the absence of leniency programs the firm stops cartel formation sooner than in case of simultaneous self-reporting after introduction of leniency. And vice versa, when the instantaneous expected penalty is high enough, the firm that decides about the optimal time of quitting the cartel on its own, in the absence of leniency programs, will choose to report later than in case the firms coordinate their actions after introduction of the leniency program. The result of this lemma will also be used later when we consider the implications of the feedback equilibrium of the preemption game with leniency.

5.1 Derivation of the Feedback Equilibrium

The above described preemption game has a special feature in that the leader payoff is not influenced by the decision of the follower. However, still in the feedback equilibrium the reaction of the follower should influence the decision of the leader about optimal time of self-reporting. The leader should take into account that the second firm can react instantaneously to the actions of the leader. This implies that the second firm will choose the same action as the leader at each instant of time due to the fact that its fine will be halved in this way. Hence, $T_F = T_L$ for any $T_L \in [0, \infty)$. This implies that the firm that moves first maximizes the value of simultaneous self reporting, $M(T)$, at each instant of time. Hence, $T_c = \arg\max_{T\geq 0} M(T)$, and $T_L = \arg\max_{T\geq 0} L(T)$ and $L(T_L) > M(T_c)$.

Now we turn to the special case, where the penalty schedule is proportional. Due to the assumptions of symmetry and the possibility of instantaneous reaction for the second firm, from the expressions (11)-(13) it is clear that in equilibrium the following condition is satisfied $L(t^*) = F(t^*) = M(t^*)$.

To find feedback equilibria of this model we recall the matrix game played at each instant of time:

	SR	Not SR
SR	$(M(t^*), M(t^*))$	$(L(t^*), F(t^*))$
Not SR	$(F(t^*), L(t^*))$	repeat

$\rightarrow$

	SR	Not SR
SR	$(M(t^*), M(t^*))$	$(M(t^*), M(t^*))$
Not SR	$(M(t^*), M(t^*))$	repeat

Table 2. Payoffs and strategies of matrix game played at time t under the assumption of possibility of instantaneous reaction.

We also denote by π the value of the infinitely lasting cartel. In case of proportional penalty this value is given by the following expression:

$$\pi = \int_0^\infty (\frac{\pi^m}{2} - \lambda\alpha w(t))e^{-rt}dt$$

So the equilibrium where both firms self-report, (S, S), arises as a pure strategy Nash Equilibrium of the matrix game described above in case $M(t^*) > \pi$. On the other hand, the decision not to reveal cartel by both players, (N, N), is a pure strategy Nash Equilibrium of this matrix game in case $\pi > M(t^*)$ or $\pi > M(t)$ for all $t \in [0, \infty)$. Recall that the maximal payoff in case of simultaneous self-reporting equals

$$M(t^*) = \int_0^{t^*} (\frac{\pi^m}{2} - \lambda\alpha w(t))e^{-rt}dt - \frac{1}{2}\alpha w(t^*))e^{-rt^*},$$

where t^* is the equilibrium time of simultaneous self-reporting.

After we simplify these expressions, we obtain

$$\pi = \frac{\pi^m}{2r}(1 - \frac{\alpha\lambda}{2r}) \quad (16)$$

$$M(t^*) = \pi - \frac{\pi^m}{2r}e^{-rt^*}(1 - \frac{\alpha\lambda}{r} + \frac{\alpha}{2} + \frac{\alpha\lambda}{2r}e^{-rt^*} - \frac{\alpha}{2}e^{-rt^*}) \quad (17)$$

Based on expressions (16) and (17) we conclude that $\pi > M(t)$ for all $t \in [0, \infty)$ only in case $\alpha\lambda < r + \frac{\alpha r}{2}$. Hence, when $\alpha\lambda < r + \frac{\alpha r}{2}$ the unique SPNE of the game is $(N, N)_t$ for all $t \in [0, \infty)$. This means that self-reporting is never optimal when $\alpha\lambda < r + \frac{\alpha r}{2}$, because firms prefer to keep the cartel forever. In this case, introduction of leniency programs does not have any effect on cartel stability.

To complete the analysis, we consider the setting where $\alpha\lambda > r + \frac{\alpha r}{2}$. In this case, self-reporting occurs at the moment, T_c , when $M(t)$ reaches its maximum. Hence, the unique SPNE of the game is to play $(N, N)_t$ for all $t \in [0, T_c)$ and to play $(S, S)_t$ when $t = T_c$. Hence, the game stops after period T_c. In this case, introduction of leniency programs can influence the cartel stability. We will study these effects in more detail in the next proposition.

In case $\alpha\lambda > r + \frac{\alpha r}{2}$, according to lemma 2 two possible outcomes can arise: $T^* < T_c$ or $T^* > T_c$. The first inequality implies that the result, obtained in case we consider the game with leniency, leads to a later time of self-reporting compared to the solution of the problem of the individual decision maker when leniency is not available. And the latter case implies an earlier stopping time after introduction of leniency. In both cases the result described in the following proposition holds.

Proposition 3 *In the feedback equilibrium of the game both firms report simultaneously at time* $T_c = \frac{\ln(\frac{2\alpha(\lambda - r)}{2\alpha\lambda - 2r - \alpha r})}{r}$.

Proof: See Appendix 3.

In short, the intuition behind the proof of this proposition is as follows. There exists a continuum of simultaneous self-reporting equilibria, from which simultaneous self-reporting at time $t = T_c$ Pareto dominates all other equilibria. In this Pareto dominant equilibrium, firms "tacitly cooperate" by keeping the cartel until time T_c and then reveal it simultaneously and pay half of the fine, which is most beneficial for both of them.

Clearly, in contrast with the benchmark case, in the preemption game, which takes into account the possibility of leniency, the antitrust authority can influence the outcome of the game, i.e. the decision about the time of breaking the cartel agreement by both firms, not only by changing the fine and the probability of law enforcement. The introduction of leniency programs also appears to be an important factor that may either reduce cartel stability or facilitate collusion.

The above result also states that in the model without possibility of instantaneous reaction, leniency programs appear to be more efficient, since they enforce immediate self-reporting for lower expected fines compared to the model where instantaneous reaction is possible. Recall the model without possibility of instantaneous reaction. There we get that self-reporting becomes dominant strategy already when $\alpha\lambda > r$. On the other hand, in the model, where instantaneous reaction is possible, self-reporting becomes a dominant strategy only when $\alpha\lambda > r + \frac{\alpha r}{2}$. This comparison clearly gives the result of earlier self-reporting in case the rules are more strict, i.e. there is no possibility of instantaneous reaction. This implies that the incentives for the firms to break the cartel are stronger under the assumption that they cannot react instantaneously to the actions of their rivals.

We conclude that if the rules of leniency programs were more strict and the procedure of application for leniency was more confidential, cartel occurrence would be less likely. This can happen due to the fact that the absence of the possibility to react to the actions of a rival instantaneously increases the expected future losses if the cartel is revealed, since the payoff of $M(t)$ is no longer available for the follower.

5.2 Effects of leniency programs in case instantaneous reaction is possible

The equilibrium of the game with leniency may lead to either earlier or later deterrence than in case the firms take the decision about stopping cartel agreement in the absence of leniency programs.

Earlier deterrence happens if $2r < \alpha\lambda$, while the result of later deterrence arises if $r < \alpha\lambda < 2r$. A special case occurs when $r > \alpha\lambda$. In this case maxima of $M(t)$ and $St(t)$ in the positive orthant do not exist and $M'(t) > 0$ and $St'(t) > 0$ for all $t \in [0, \infty)$. Hence, the best strategy is cartel forever, since self-reporting is never profitable. This situation is depicted in Figure 1.

Moreover, for any $\alpha \geq 2$ we obtain from expression (15) that $T_c \leq 0$. This means that cartel formation stops immediately, $T_c = 0$. I.e. in the equilibrium of preemption game with leniency when instantaneous reaction is possible it is optimal for both firms to reveal the cartel immediately after the introduction of the leniency program. So, we can conclude that,- for proportional penalty similarly to fixed penalty,- in combination with a strict enforcement policy (when α is high, $\alpha \geq 2$) leniency programs appear to be quite efficient. They allow to achieve immediate deterrence.

If we compare the impact of the penalty of the form $s(t) = \alpha w(t)$ (with $\alpha \geq 2$) in the absence of the leniency programs, we do not observe the outcome with complete deterrence in the beginning of the planning horizon for any parameter values, whereas with the introduction of leniency programs this result becomes unambiguous[10].

Moreover, for any $\alpha < 2$, thus when penalties are low, introduction of leniency programs does not lead to the outcome with immediate complete deterrence, since $T_c > 0$.

To illustrate the above analysis, in Figures 2 and 3 the two functions , $M(t)$ and $St(t)$, are plotted for cases $r < \alpha\lambda < 2r$ and $2r < \alpha\lambda$, respectively. The solid lines correspond to the value of stopping in situation without leniency, and the dotted line represents the value of simultaneous self-reporting that is relevant value in the game with leniency where instantaneous reaction is possible.

[10]For complete derivation of this result see paper by Motchenkova and Kort (2003).

In case $r < \alpha\lambda < 2r$[11], we get that $T_c > T^*$, where $T_c = \arg\max_{t\geq 0} M(t)$ and $T^* = \arg\max_{t\geq 0} St(t)$. This implies that the result, obtained in case we consider equilibria of the preemption game with leniency, leads to a later optimal stopping time. Hence, compared to the benchmark case where no leniency is available, greater harm is done to the consumers. Recall that $T^* = \arg\max_{t\geq 0} St(t)$ reflects the optimal time of stopping the cartel formation in the benchmark model, where firms do not take strategic considerations into account, see expression (3) in Section 3. So, the fact that the firms take into account the reaction of the other firm clearly increases the stability of cartel for intermediate values of α and λ, i.e. $r < \alpha\lambda < 2r$, compared to the optimum of a single decision maker in the situation without leniency. However, in case $2r < \alpha\lambda$, equilibrium of the game with leniency (T_c, T_c) leads to an earlier stopping time than in the benchmark model. In this case, see Figure 3, in the solution of the game with leniency duration of cartel agreement is reduced, since $\arg\max_{t\geq 0} M(t) < \arg\max_{t\geq 0} St(t)$.

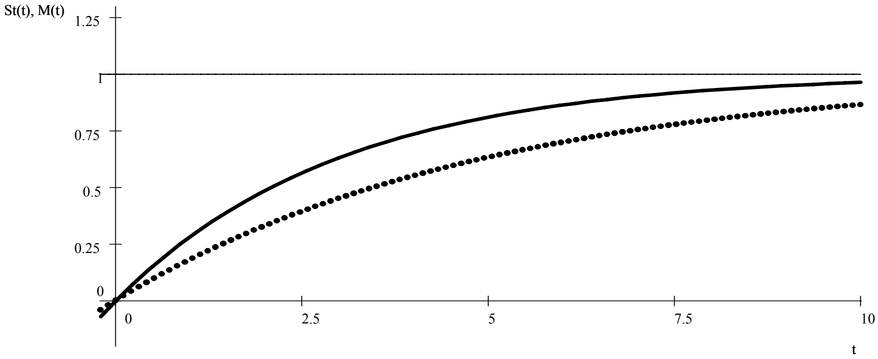

Figure 1: Graphs of π, $St(t)$ and $M(t)$ for case $\alpha\lambda < r$. Parameter values are $\alpha = 1.5, \lambda = 0.2, r = 0.1, \pi^m = 1$.

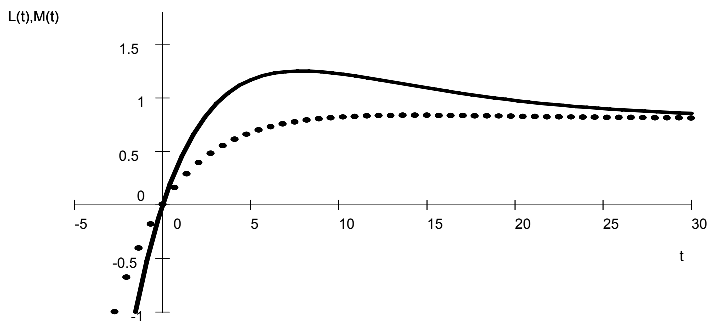

Figure 2: Graphs of $St(t)$ and $M(t)$ for case $r < \alpha\lambda < 2r$. Parameter values are $\alpha = 1, \lambda = 0.2, r = 1/8, \pi^m = 1$.

[11] See Figure 2.

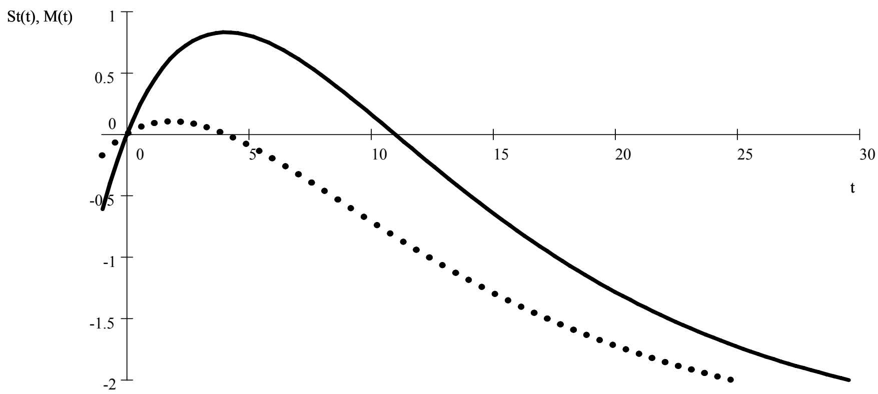

Figure 3: Graphs of $St(t)$ and $M(t)$ for case $\alpha\lambda > 2r$. Parameter values are $\alpha = 1.5, \lambda = 0.2, r = 0.1, \pi^m = 1$

The main conclusion of the above analysis is that, when the procedure of application for leniency is not confidential, leniency may still reduce duration of cartel agreements, but not in all cases. When penalties and rate of law enforcement are low, introduction of leniency programs may, on the contrary, facilitate collusion.

6 Analysis of the model with fixed penalty

6.1 Benchmark model without leniency

In case the penalty is fixed the value of stopping the cartel formation is determined as in expression (1) with $s(t) = F^n$. So that it has the following form:

$$St(T) = \max_T \int_0^T [\frac{\pi^m}{2} - \lambda F^n] e^{-rt} dt \tag{18}$$

In order to find the optimal time of stopping the cartel agreement, we maximize (18) with respect to time. This implies that the optimal time of quitting the cartel agreement for the single firm is given by

$$\begin{aligned} T^* &\to \infty \text{ if } \frac{\pi^m}{2} > \lambda F^n \\ T^* &= 0 \text{ if } \frac{\pi^m}{2} \leq \lambda F^n \end{aligned} \tag{19}$$

So, we can conclude that, while taking the decision about the optimal time of quitting the cartel agreement , the firm just compares expected instantaneous benefits from price-fixing and expected punishment. Moreover, from expression (19) it follows that when the expected penalty is high enough, i.e. $\lambda F^n > \frac{\pi^m}{2}$, cartel formation stops immediately at time zero.

Expression (19) shows that the optimal decision is either to stop collusion immediately or never. The higher the expected penalty the more likely that cartel formation stops immediately. On the other hand, the higher the instantaneous illegal gains the more likely that the cartel will last forever.

6.2 Analysis of the game with leniency, where the penalty is fixed

6.2.1 Regime with confidential procedure of application for leniency

Depending on the severity of punishment, two possible outcomes can arise in a feedback equilibrium of a preemption game with leniency where instantaneous reaction is not possible. Either both firms report the cartel simultaneously in the beginning of the game or the cartel will last forever.

Proposition 4 *The feedback equilibria of the game, when penalty is fixed and equals F^n, are immediate simultaneous self-reporting, i.e. $(t_1^*, t_2^*) = (0, 0)$, if $\lambda F^n > \frac{\pi^m}{2}$ or cartel forever, i.e. $(t_1^*, t_2^*) = (\infty, \infty)$, if $\lambda F^n \leq \frac{\pi^m}{2}$.*

We refrain from presenting the proof of proposition 4, since it is similar to the proof of proposition 1 with a number of simplifications.[12] Clearly, in case, firms cannot react instantaneously to the actions of their rivals, under the fixed penalty scheme the solution of the game with leniency coincides with the outcome of the benchmark model, where leniency is not available.

6.2.2 Non-confidential procedure of application for leniency

In this subsection we consider a situation with less strict leniency programs, where firms can react instantaneously to the actions of their rivals. If we compare the optimal stopping time in a setting without leniency and the equilibrium of the preemption game with leniency we can conclude that for any positive discount rate the optimal time of simultaneous self-reporting in case of leniency is more likely to be greater than the optimal stopping time, which maximizes the individual payoff when leniency is not available. To be more precise, due to the discontinuity result of the model with fixed penalty, in case of leniency the outcome of infinitely lasting

[12] The proof of proposition 4 is available from the author upon request.

cartel is more likely than the outcome of immediate simultaneous self-reporting compared to the benchmark case. These results are summarized in the following proposition.

Proposition 5 *Consider the situation where the penalty is fixed and* $(\lambda - \frac{r}{2})F^n < \frac{\pi^m}{2} < \lambda F^n$. *In the setting without leniency both firms report at* $t_1^* = t_2^* = 0$. *However, if we consider the equilibrium of the game with leniency, immediate self-reporting does not occur: both firms report at* $T_c = t_c^* \to \infty$.

Proof: See Appendix 4.

So, with a fixed penalty scheme, even in the presence of leniency, the efficiency of deterrence depends only on the amount of the fine and the probability of law enforcement. Moreover, we show that, when penalties are fixed and fall below a certain threshold, leniency programs may well facilitate collusion.

Hence, if we consider the setting where self-reporting is not possible, we can conclude that cartel formation stops immediately, at the beginning of the planning horizon, only when the penalty is fixed and high enough to outweigh the expected benefits from collusion. However, in case the government introduces leniency, even an expected penalty being greater than instantaneous gains from price-fixing, cannot ensure immediate success of the leniency program. Only the condition $\frac{\pi^m}{2} \leq (\lambda - \frac{r}{2})F^n$ which implies $\lambda F^n > \frac{\pi^m}{2} + \frac{r}{2}F^n$, can ensure immediate self-reporting in case firms take into account the possibility of leniency and are able to react instantaneously to the actions of their rivals. So, when penalties are fixed and procedure of application for leniency is not confidential, the introduction of leniency programs reduces the effectiveness of antitrust enforcement. This implies that the authority will have to increase either the amount of the penalty or the rate of law enforcement in order to achieve whistle-blowing by both firms immediately in the beginning of the planning horizon. Otherwise, when the penalty is low, i.e. $F^n < \frac{\pi^m}{2\lambda - r}$, introduction of leniency makes the cartel more stable. This is a very surprising result, since intuitively leniency should increase the incentives for firms to betray the cartel and, hence, reduce cartel stability. However, when the penalty is low, and does not depend on the amount of illegal gains, it may be the case that the reduced (as a consequence of leniency) net expected fine is, actually, less than the instantaneous gains from price-fixing, and this drives the result.

7 Effects of leniency in the model of dynamic price competition and "tacit collusion"

In this section we study the effects of leniency programs on the behavior of the firms in the model of dynamic price competition where "tacit collusion" may arise[13]. In the previous sections the situation of a formalized cartel was considered. In particular, we analyzed a model where there is a formal cartel agreement. This can be discovered by the antitrust authority and punished on the basis of official documents, which provide evidence of illegal price-fixing agreement. However, it is often the case that firms do not form an explicit cartel, but sustain high prices by means of "tacit collusion", which harms consumers. This is also an illegal activity and can be punished according to the Article 81 of the EC Treaty. Recall, for example, the Soda-Ash case. In that case the Commission decided that tacit collusion between ICI, a British company, and Solvay, a Belgian company, was an infringement of Article 81 (ex-85). The Commission motivated the decision by the fact that the term "concerted practices" mentioned in Article 81 among the prohibited practices also covered the case of tacit collusion between these two companies.

The situation of "tacit collusion" assumes that when there is no formal agreement between 2 firms, but they still keep prices above competitive level, both of them have incentives to undercut and obtain monopoly profits. Hence, this situation involves the possibility of unercutting. This special feature makes this case different from the assumptions of the preemption game described above[14].

In this section we incorporate the possibility of undercutting into the model of leniency without instantaneous reaction[15]. We consider a game between two symmetric firms that may cooperate by charging the monopoly price, and obtain half of the monopoly profits in the industry, $\frac{\pi^m}{2}$, each period. However, there is a threat that this violation will be discovered by the antitrust authority. There are two other options for the firms: self-reporting or undercutting. The second option is to self-report to the authority and obtain leniency (reduction of the fine).

[13] See, e.g., Tirole (1988) "The Theory of Industrial Organization" chapter 6.

[14] In the case of an explicit cartel undercutting is not so easy but also possible. In addition, in case of an explicit cartel there should be no possibility of renegotiation in order to sustain an agreement. So, if we assume that renegotiation is either impossible or very costly, then we can include an additional strategy in the form of "possibility of undercutting" in the model of explicit collusion as well.

[15] Clearly, the model of dynamic price competition and "tacit collusion", as it is described in Tirole (1988) rests on the assumption that instantaneous reaction is not possible.

The third option is to undercut by reducing the monopoly price by a minimal amount. Then it obtains monopoly profits, π^m, for one or more periods. We also assume that after one of the firms betrayed and another firm discovers it, collusion stops forever. We define here an information lag, which delays the punishment phase and allows the firm to enjoy extra profits for several periods, by ε. This setup gives us a number of interesting results that differ from the model where undercutting is not possible. We can summarize these results in the following proposition.

Proposition 6 *The feedback equilibria of the game with proportional penalty are immediate stopping by undercutting, i.e.* $(t_1^*, t_2^*) = (0, 0)$, *if* $\alpha\lambda > 2r - re^{r\varepsilon}$ *or cartel forever, i.e.* $(t_1^*, t_2^*) = (\infty, \infty)$, *if* $\alpha\lambda \leq 2r - re^{r\varepsilon}$.

We refrain here from presenting the proof of this proposition, since it uses the same methods as the proof of proposition 1. We sketch the main arguments of the proof here. First, we describe the objective function of the firm that chooses the undercutting option. If the firm 1 decides to undercut at instant T, it obtains half of cartel profits, $\frac{\pi^m}{2}$, from the initial period until T and full monopoly rents, π^m, from T until $T+\varepsilon$. At instant $T+\varepsilon$, the second firm discovers that firm 1 betrayed the cartel and, hence, collusion stops forever. However, there is a threat of expected punishment throughout periods 0 to $T+\varepsilon$. Hence, the value of undercutting at instant T is given by the following expression

$$U(T) = \int_0^T (\frac{\pi^m}{2} - \lambda\alpha w(t))e^{-rt}dt + \int_T^{T+\varepsilon} (\pi^m - \lambda\alpha w(t)e^{-rt}dt.$$

At the same time the leader value or the value of self-reporting at instant T has the form $L(T) = \int_0^T (\frac{\pi^m}{2} - \lambda\alpha w(t))e^{-rt}dt$. Clearly, the value of undercutting intersects the value of self-reporting from above in the point where $\int_T^{T+\varepsilon} (\pi^m - \lambda\alpha w(t)e^{-rt}dt = 0$, hence when $T^{**} = \frac{\ln(\frac{\lambda\alpha(1+e^{-r\varepsilon})}{2(\lambda\alpha-2r)})}{r}$. This implies that the best option and, hence, the option that is chosen by both firms (due to symmetry) up to the time $T^{**} = \frac{\ln(\frac{\lambda\alpha(1+e^{-r\varepsilon})}{2(\lambda\alpha-2r)})}{r}$ is the undercutting option, if an additional condition, $\lambda\alpha > 2r$, for existence of T^{**} is satisfied. Hence, up to $T^{**} = \frac{\ln(\frac{\lambda\alpha(1+e^{-r\varepsilon})}{2(\lambda\alpha-2r)})}{r}$ the undercutting value, $U(T)$, will be compared to $\pi = \int_0^\infty (\frac{\pi^m}{2} - \lambda\alpha w(t))e^{-rt}dt$, being the value of the infinitely lasting cartel. After time $T^{**} = \frac{\ln(\frac{\lambda\alpha(1+e^{-r\varepsilon})}{2(\lambda\alpha-2r)})}{r}$ the leader value, $L(T) = \int_0^T (\frac{\pi^m}{2} - \lambda\alpha w(t))e^{-rt}dt$, will be compared to the value of the infinitely lasting cartel[16].

[16] See proof of Propsition 1.

Proposition 6 states that in the preemption game with leniency, where firms illegally fix the prices above competitive level and can undercut each other, the no collusion (or immediate stopping) outcome arises when the coefficient of the expected penalty, $\alpha\lambda$, is greater than $2r - re^{r\varepsilon}$. Now, if we compare the result of proposition 6 to the result of proposition 1, we can conclude that, since $r > 2r - re^{r\varepsilon}$ for any $r \in (0, 1)$, the possibility of undercutting improves the result: a smaller expected penalty, $\alpha\lambda > 2r - re^{r\varepsilon}$, insures immediate stopping compared to the model where only self-reporting option is available to the firms. Hence, the antitrust authority has to put less efforts into control in order to achieve the outcome of complete deterrence. This result is also quite intuitive, since the possibility of undercutting increases the incentives for the firms to betray the cartel and, hence, reduces the stability of price-fixing agreements.

Another interesting observation is connected with the influence of the size of the informational lag in case of undercutting on the stability of cartel agreement. From Proposition 6 it follows that the bigger the ε (information lag) the easier for the antitrust authority to block the violation, since a smaller expected penalty insures immediate stopping. Moreover, for $\varepsilon > \frac{\ln 2}{r}$ collusion will never arise in equilibrium. This can be explained by the fact that, when the information lag is bigger the cartel is less stable due to the fact that undercutting brings benefits for a longer period and, hence, it is a more attractive option.

8 Conclusions

The main problem addressed in this paper is how leniency programs influence the stability of cartels under two different regimes of fines. First, we study the effects of leniency in case the penalty is an increasing function of the accumulated illegal gains from price-fixing to the firm. Next, we look at the case where the penalty is fixed. We denote the former system by proportional penalty scheme. The enforcement problem we study has several ingredients. We analyze the design of self-reporting incentives, having a group of (and not a single) defendants. Moreover, we consider a dynamic setup, where accumulated benefits and losses from crime are taken into account.

For this purpose we use the tools of optimal stopping and timing games. In particular, the preemption game is studied in order to identify the advantages of being the leader in the race to the court game between the members of the existing cartel after the introduction of leniency programs. The approach, we use, is based on the Reiganum-Fudenberg-Tirole approach, who

The complete proof of proposition 6 is available from author upon request.

applied the concept of timing games to a technology adoption problem. We apply a similar procedure to a cartel formation game between two firms in the presence of leniency programs, which allows taking into account the possibility to influence the internal stability of the cartel.

Comparison of results in the situations with and without leniency suggests that antitrust enforcement after introduction of leniency programs is more efficient than in the absence of leniency. Hence, leniency improves upon the situation without leniency.

We also obtain that in the settings with a strictly confidential procedure of application for leniency, the outcome is immediate self-reporting by both firms in case the expected penalty is sufficiently high (but still below the threshold of the model where instantaneous reaction is possible). This implies that strict leniency programs unambiguously increase the efficiency of antitrust enforcement and reduce cartel stability. The reason is that the impossibility to react instantaneously to the actions of the rival increases expected future losses in case the cartel is revealed. This happens due to the fact that it is no longer possible for the follower to obtain reduction of the fine from simultaneous self-reporting. Hence, we conclude that if the rules of leniency programs are more strict and the procedure of application for leniency is more confidential, cartel occurrence is less likely.

We find that in most cases leniency reduces duration of cartel agreements but this result is not unambiguous. In case leniency programs are not too strict and fines are proportional to the accumulated illegal gains from price-fixing the result is as follows. Under strict antitrust enforcement[17], the possibility to self-report and be exempted from the fine increases the incentives for the firms to stop cartel formation, and, hence, reduces the duration of cartels. However, when penalties and rate of law enforcement are low, introduction of leniency programs may, on the contrary, facilitate collusion.

Under a fixed penalty scheme, even in the presence of leniency, the efficiency of deterrence depends only on the amount of the fine and the probability of law enforcement. Moreover, we have shown that in some cases, when penalties are fixed and fall below a certain threshold, less strict leniency programs facilitate collusion.

Another interesting conclusion comes from a numerical comparison of the efficiency of antitrust enforcement under proportional penalty in the absence of leniency and after introduction of leniency. In the earlier case we do not observe the outcome of complete deterrence in the beginning of the planning horizon for any relevant (from the legislation point of view) parameter values, whereas after the introduction of leniency programs this result becomes unambiguous

[17] Here by strict antitrust enforcement we mean high fins and rate of law enforcement.

for sufficiently high (but still in the range of legally acceptable) values of the scale parameter of the penalty scheme.

We also study the effects of leniency programs on the behavior of firms in the model of dynamic price competition and "tacit collusion", where firms can undercut each other in prices. The result of this model implies that in environments where undercutting is possible it is easier for competition authority to prevent price-fixing, since a smaller expected penalty insures immediate stopping. This implies that the antitrust authority has to put less efforts into control in order to achieve the result of complete deterrence.

Another interesting extension would be to study the behavior of asymmetric firms. They may differ either in costs or size. This extension would make the model much closer to real world situation but solution of dynamic games with asymmetric information it not a trivial task.

9 Appendixes

9.1 Appendix 1: Proof of proposition 1

Let the penalty be proportional to the accumulated illegal gains from cartel formation $s(t) = \alpha w(t)$. In this case two possible outcomes can arise depending on the parameters of the model. Either both firms report the cartel simultaneously in the beginning of the game or the cartel will last forever.

1. Consider the case $\alpha\lambda \leq r$.

In this case we compare the value of infinitely lasting cartel $\pi = \int_0^\infty (\frac{\pi^m}{2} - \lambda s(t))e^{-rt}dt$, which can be rewritten as $\pi = \frac{\pi^m}{2r}(1 - \frac{\alpha\lambda}{2r}) > 0$, with the leader's value that can be rewritten as $L(t) = \pi - \frac{\pi^m}{2r}e^{-rt}(1 - \frac{\alpha\lambda}{r} + \frac{\alpha\lambda}{2r}e^{-rt})$. Hence, given $\alpha\lambda \leq r$, we obtain $\pi > L(t)$ for all $t \in [0, \infty)$. This implies that the matrix game played at each instant t, which has been described in section 4.1, has two pure strategy Nash Equilibria, and the equilibrium $(N, N)_t$ Pareto dominates $(S, S)_t$ for all $t \in [0, \infty)$. This implies that, when $\alpha\lambda \leq r$ the unique SPNE of the dynamic game is $(N, N)_t$ for all $t \in [0, \infty)$. Hence, $(t_1^*, t_2^*) = (\infty, \infty)$ is the unique feedback equilibrium of the preemption game if $\alpha\lambda \leq r$.

The above considerations imply that in case $\alpha\lambda \leq r$ the cartel will last forever, and self-reporting is never a dominant strategy for any of the firms. See also Figure 1 above in section 5.2.

End of the proof of part 1.

2. Consider the setting with $\alpha\lambda > r$. Two possible sub-cases can arise here.

a) $\pi < 0$, so that $\pi < L(t)$ for all $t \in [0, \infty)$ This can hold only if $\alpha\lambda > 2r$. See Figure 3.

b) $\pi < L(t)$ for some $t \in [0, \infty)$ holds when $r < \alpha\lambda < 2r$. This situation is depicted in Figure 2.

In both cases the dominant strategy for each firm is to play S_t at each instant of time. This implies that $(t_1^*, t_2^*) = (0, 0)$ is the unique feedback equilibrium of the preemption game if $\alpha\lambda > r$.

We prove this statement by backward induction.

We can show that in both cases, $\alpha\lambda > 2r$ and $r < \alpha\lambda < 2r$, the function $L(t)$ approaches π from above when t tends to infinity, i.e. there exists a finite number $\widehat{t}$ such that $L(t) > \pi$ for all $t > \widehat{t}$,[18] where $\widehat{t}$ satisfies $L(t) - \pi = 0$. This implies that $\widehat{t} = \frac{\ln(\frac{\alpha\lambda}{2(\alpha\lambda - r)})}{r}$. It is clearly finite for any finite values of α, λ and r when $\alpha\lambda > r$. Moreover, it is easily verified that $L(t) - \pi = -\frac{\pi^m}{2r}e^{-rt}(1 - \frac{\alpha\lambda}{r} + \frac{\alpha\lambda}{2r}e^{-rt}) > 0$ for all $t > \widehat{t}$.

Since $M(t) > F(t)$ for all $t \in (0, \infty)$, given $s^F \in (\frac{1}{2}s^n, s^n]$ and $s^m = \frac{1}{2}s^n$, and $L(t) > \pi$ for $t > \widehat{t}$, we can conclude that for both firms the strategy S_t (self-report at t) strictly dominates strategy N_t for all $t > \widehat{t}$. Hence, for any $t \in [\widehat{t}, \infty)$, there is a unique Nash Equilibrium of simultaneous move matrix game played at instant t of a dynamic game, which is described by $(S, S)_t$.

Now we apply the backward induction argument. We look at the matrix game played at instant t and assume that, if the game continues for one more period, the equilibrium of game at t^+ will be $(S, S)_{t^+}$, since simultaneous self-reporting should be part of the subgame perfect strategy given the result above. Then the payoff matrix at t will have following form

	Self-report	Not self-report
Self-report	$(M(t), M(t))$	$(L(t), F(t))$
Not self-report	$(F(t), L(t))$	$(M(t^+), M(t^+))$

By assumption of the model, function $L(t)$ is always above the function $M(t)$ and, hence, $L(t) > M(t^+)$ for any t and $t^+ \in (0, \infty)$. This inequality implies that the strategy S_t (to self-report at t) is dominant for both firms. This implies that the matrix game at t has a unique pure strategy Nash Equilibrium : (S, S). Repeating this argument backwards to the initial period of the game, and taking into account the fact that $L(t) > M(t^+)$ for any t, $t^+ \in (0, \infty)$, we obtain that self-reporting is a dominant strategy for both players at each instant of time. Consequently, immediate simultaneous self-reporting at $t = 0$ is a SPNE of the

[18] Recall Figure 2 and Figure 3.

infinitely repeated game. And, hence, $(t_1^*, t_2^*) = (0,0)$ is the unique feedback equilibrium of the preemption game when $\alpha\lambda > r$.

In other words, when $\alpha\lambda > r$, both firms want to become leader and report at $T_L = \arg\max L(t)$. As a result a firm will try to preempt the other firm by reporting at time $T_L - \varepsilon$. But then the other will try to preempt by reporting at $T_L - 2\varepsilon$ and so forth and so on. This process stops at time $t = 0$, where both values $L(t)$ and $F(t)$ are equal.

End of the proof of part 2.

End of proof of proposition 1.

9.2 Appendix 2: Proof of Lemma 2

Proof of lemma 2:

First , we reformulate expressions (14) and (15) in order to make them comparable. The (14) gives $T_L = \frac{\ln(\frac{1}{1-\frac{r}{\alpha\lambda}})}{r}$. Similarly, the expression (15) gives $T_c = \frac{\ln(\frac{1-\frac{r}{\lambda}}{1-\frac{r}{\alpha\lambda}-\frac{r}{2\lambda}})}{r}$.

Second, we consider the difference

$$\frac{1}{1-\frac{r}{\alpha\lambda}} - \frac{1-\frac{r}{\lambda}}{1-\frac{r}{\alpha\lambda}-\frac{r}{2\lambda}} = \alpha r \frac{\alpha\lambda - 2r}{(\alpha\lambda - r)(2\alpha\lambda - 2r - r\alpha)} \tag{20}$$

It is clear that expression (20) is negative when $r < \alpha\lambda < 2r$. This implies that $T_c > T_L$ when $r < \alpha\lambda < 2r$.

In case $2r < \alpha\lambda$ expression (20) becomes strictly positive and, consequently, we obtain that $T_L > T_c$ when $2r < \alpha\lambda$.

End of the proof.

9.3 Appendix 3: Proof of Proposition 3

If the leader reports at time T_L, then the best response of the follower, given the possibility to react instantaneously, is to report at time T_L as well. But this means simultaneous self-reporting and, consequently, the payoffs for both firms will be $M(T_L)$, which is less than $M(T_c)$ by definition. So, the rational leader will anticipate this and take into account this best response of the follower. Consequently, his optimal strategy would be to wait until T_c and then both firms report simultaneously at time T_c.

Here $(t_1^*, t_2^*) = (T_c, T_c)$ is a feedback equilibrium of the preemption game with leniency, since no one of the firms has incentives to deviate either by waiting with self-reporting till $t > T_c$ or by preempting the other firm by playing $t < T_c$. In both cases, given the assumption that firms can react instantaneously and, hence, the second firm will also self-report immediately after the first firm does so, both firms obtain lower payoffs: $M(t) < M(T_c)$ for any $t \neq T_c$, since by definition $T_c = \arg\max_{t\geq 0} M(t)$.

End of the proof.

Note, that this result holds only under assumptions that firms are completely symmetric and can react instantaneously to the actions of their opponents. In case when we relax the assumption that firms can react instantaneously, the feedback equilibrium of the game is $(t_1, t_2) = (0, 0)$ if $\alpha\lambda > r$ or cartel forever, i.e. $(t_1, t_2) = (\infty, \infty)$ if $\alpha\lambda < r$.

9.4 Appendix 4: Proof of proposition 5

To prove the result of proposition 5, we first derive the optimal stopping time for the firm when there is no leniency and the optimal stopping time in case of simultaneous self-reporting with leniency in the setting where the penalty is fixed.

Following the result of the benchmark model, we obtain that the optimal stopping time in the model without leniency is given by

$$\begin{aligned} T_L &\rightarrow \infty \text{ if } \frac{\pi^m}{2} > \lambda F^n, \\ T_L &= 0 \text{ if } \frac{\pi^m}{2} \leq \lambda F^n. \end{aligned} \tag{21}$$

In the game with leniency, following the reasoning similar to Proposition 3, we conclude that $(t_1, t_2) = (T_c, T_c)$ is a feedback equilibrium of the preemption game with leniency, where $T_c = \arg\max_{t\geq 0} M(t)$. In this case the firms have no incentives to deviate either by waiting with self-reporting till $t > T_c$ or by preempting the other firm by playing $t < T_c$. In both cases, given the assumption that firms can react instantaneously and, hence, the second firm will also self-report immediately after the first firm does so, both firms obtain lower payoffs: $M(t) < M(T_c)$ for any $t \neq T_c$.

The value of simultaneous self-reporting in the setting with fixed penalty is given by

$$M(T_c) = \int_0^{T_c} (\frac{\pi^m}{2} - \lambda F^n) e^{-rt} dt - \frac{1}{2} F^n e^{-rT_c} \tag{22}$$

Next, we derive exact formulas for the feedback equilibrium of the game with leniency where the penalty is fixed. Recall the game described in Table 1. The outcome, i.e. whether (S, S) or (N, N) will occur, depends on the magnitude of gains from cartel formation and the expected fine. Maximizing expression (22) with respect to time we conclude that $T_c \rightarrow \infty$ if $\frac{\pi^m}{2} > (\lambda - \frac{r}{2})F^n$ and $T_c = 0$ if $\frac{\pi^m}{2} \leq (\lambda - \frac{r}{2})F^n$. Hence, $(S, S)_t$ with $t = 0$ is a SPNE when

$\frac{\pi^m}{2} \leq (\lambda - \frac{r}{2})F^n$, while $(N, N)_t$ is a SPNE for all $t \in [0, \infty)$ when $\frac{\pi^m}{2} > (\lambda - \frac{r}{2})F^n$. We conclude that, two outcomes can arise as an equilibrium in feedback strategies: one is immediate self-reporting at $T_c = 0$, i.e. $(t_1^*, t_2^*) = (0, 0)$ and the other equilibrium is never self-report, i.e. $T_c \to \infty$, so that $(t_1^*, t_2^*) = (\infty, \infty)$.

End of the proof.

References:

Apesteguia, J., M. Dufwenberg and R. Selten, 2003, Blowing the whistle, mimeo, *Bonn Economic Discussion papers* 9.

Becker, G.S., 1968, Crime and punishment: an economic approach, *Journal of Political economy* 76(2),169-217

Emons, W., 2001, A note on optimal punishment for repeat offenders, mimeo, University of Bern and CEPR.

European Commission, 2002, "Commission adopts new leniency policy for companies which give information on cartels", press release, Brussels, Feb. 13, 2002.

[http://europa.eu.int/rapid/start/cgi/guesten.ksh?p_action.gettxt=gt&doc=IP/02/247 |0|RAPID&lg=EN]

Feess, E. and M. Walzl, 2003, Corporate leniency programs in the EU and the USA, mimeo, *German Working Papers in Law and Economics* 24.

Fudenberg, D. and J. Tirole ,1985, Preemption and rent equalization in the adoption of new technology, *The Review of Economic Studies*, 52, 393-401.

Garoupa, N., 1997, The theory of optimal law enforcement, *Journal of Economic Surveys* 11, 267-295.

Garoupa, N., 2001, Optimal magnitude an probability of fines, *European Economic Review* 45, 1765-1771.

Hinloopen, J., 2003, An economic analysis of leniency programs in antitrust law, *De Economist* 151 (4), 415-432.

Innes, R., 1999, Remediation and self-reporting in optimal law enforcement, *Journal of Public Economics* 72, 379-393.

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

Motchenkova, E. and P. Kort, 2003, Analysis of the properties of current penalty schemes for violations of antitrust law, mimeo.

Motta, M. and M.Polo, 2003, Leniency Programs and Cartel Prosecution, *International journal of industrial organization* 21(3), 347-379 .

OECD report 2002, "Report on the Nature and Impact of Hard Core Cartels and Sanctions Against Cartels."

Polinsky, A. and S. Shavell , 1979, The optimal trade-off between the probability and magnitude of fines, *The American Economic Reveiw* 69(5), 880-891.

Reinganum, J.F., 1981, On the diffusion of new technology: A game theoretic approach, *The Review of Economic Studies,* 48, 395-405.

Spagnolo, G., 2000, "Optimal Leniency Programs" *FEEM Working Paper* No. 42.2000: http://ssrn.com/abstract=235092

Spratling Gray R. (Deputy assistant Attorney, Antitrust Division, US Dept. of Justice). "The corporate leniency policy: Answers to Recurring Questions", presented at the Spring 1998 ABA Meeting (Antitrust section), available for download at http://www.usdoj.gov/atr/public/speeches/1626.ht

Tirole, J., 1988, "The Theory of Industrial Organization".

Wils, W.P.J., 2002, The optimal enforcement of EC antitrust law". *Essays in law and economics.*

European system. Guidelines on the method of setting fines imposed (PbEG 1998.)

US system . Guidelines manual (chapter 8-Sentencing of organizations)

Dutch system . Guidelines for the setting of fines. (Section 57(1) of Competition Act.)

US Department of Justice, Official Guidelines

The twelfth annual report. DOJ 1998. [http://www.usdoj.gov/atr/public/speeches/1583.html]