

Houba, Harold; Motchenkova, Evgenia; Wen, Quan

Working Paper

The Effects of Leniency on Maximal Cartel Pricing

Tinbergen Institute Discussion Paper, No. 09-081/1

Provided in Cooperation with:

Tinbergen Institute, Amsterdam and Rotterdam

Suggested Citation: Houba, Harold; Motchenkova, Evgenia; Wen, Quan (2009) : The Effects of Leniency on Maximal Cartel Pricing, Tinbergen Institute Discussion Paper, No. 09-081/1, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:

<https://hdl.handle.net/10419/87019>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



TI 2009-081 / 1

Tinbergen Institute Discussion Paper

The Effects of Leniency on Maximal Cartel Pricing

Harold Houba^{1,2}

Evgenia Motchenkova¹

Quan Wen³

¹ *VU University Amsterdam;*

² *Tinbergen Institute;*

³ *Vanderbilt University.*

Tinbergen Institute

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam

Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam

Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Most TI discussion papers can be downloaded at
<http://www.tinbergen.nl>.

The Effects of Leniency on Maximal Cartel Pricing*

Harold Houba[†]
VU University Amsterdam
and Tinbergen Institute

Evgenia Motchenkova[‡]
VU University Amsterdam

Quan Wen[§]
Vanderbilt University

Abstract

We analyze maximal cartel prices in infinitely-repeated oligopoly models under leniency where fines are linked to illegal gains, as often outlined in existing antitrust regulation, and detection probabilities depend on the degree of collusion. We introduce cartel culture that describes how likely cartels persist after each conviction. Our analysis disentangles the effects of traditional antitrust regulation, leniency, and cartel strategies. Without rewards to the strictly-first reporter, leniency cannot reduce maximal cartel prices below those under traditional regulation. Moreover, in order to avoid adverse effects fine reductions should be moderate in case of multiple reporters. Our results extend the current literature and partially support existing leniency programs.

JEL Classification:

L41 Monopolization, Horizontal Anti-competitive Practices, K21 Antitrust Law, C72 Non-cooperative Games.

Keywords:

Cartel, Antitrust, Competition Policy, Leniency Program, Self-reporting, Repeated Game

*We would like to thank Rene van den Brink, Riccardo Calcagno, Arantza Estevez-Fernandez, Dave Furth, Joseph Harrington, Gerard van der Laan, Ines Lindner, Els Kleijn, Robert Porter, Patrick Rey and the participants of ACLE (2009), EEA-ESEM (2009), and EARIE (2009) conferences for stimulating discussions and valuable comments.

[†]Department of Econometrics, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, Netherlands. Email: hhoub@feweb.vu.nl.

[‡]Department of Economics, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, Netherlands. Email: emotchenkova@feweb.vu.nl.

[§]Department of Economics, Vanderbilt University, VU Station B #351819, 2301 Vanderbilt Place Nashville, TN 37235-1819, U.S.A. Email: quan.wen@vanderbilt.edu.

1 Introduction

During the last two decades, antitrust policies in the US and the EC have undergone substantial reforms and currently include leniency programs as a key ingredient, see US Department of Justice (1993) and EC (2006). Leniency programs grant total or partial immunity from fines to cartel members collaborating with the antitrust authority (AA) by revealing information about the cartel. This revelation may take place *ex-ante* before any investigation by the AA starts, or *ex-post* during an ongoing investigation. Leniency programs are based upon the economic principle that firms, who broke the law, might report their illegal activities if given proper incentives. Effective leniency programs might dissolve existing cartels or, even better, a priori deter such illegal activities. The leniency program in the US reduces the fines to cartel members, which are related to their illegal gains. Moreover, cartel members may face liabilities in the form of private law suites by their customers. The EC legislation has penalty schemes for conviction and leniency that are proportional to illegal gains.

The US Department of Justice reports some empirical evidence in favor of the major modifications of its leniency program in 1993, see D.O.J. (1998). Despite this evidence, Spagnolo (2008) asserts that the effects of leniency programs are still not fully understood theoretically. For example, it is unclear whether the observed increase in cartel detection cases is the result of unobserved increase in cartel activity or the result of improved effectiveness of leniency programs. Addressing this issue is one of the main objectives of this paper. We analyze the effects of leniency programs in an infinitely-repeated sequential game. An innovative aspect is that we analyze the maximal cartel price, i.e., the largest cartel price for which the equilibrium conditions for sustainability hold, due to the following considerations. First, the maximal cartel price represents consumers' worst-case scenario of maximal damage. Also, it is the relevant proxy for the set of sustainable cartel prices, and it puts a simple upper bound on the profit-maximizing cartel price. Second, the maximal cartel price adjusts naturally to policy changes and this is closer to reality where cartels are more likely to lower prices rather than to give up on collusion altogether. Third, since AAs often have no detailed information about demand, costs, and profits of firms, AAs' assessments of market behavior are, usually, limited to observed prices. The maximal cartel price naturally complements empirical price assessments by taking society's maximal damage as its leading principle, i.e., provides a simple worst-case scenario. Finally, it avoids some technical problems concerning monotonicity and concavity associated with profit maximization in our general framework. Since the profit-maximizing cartel price may coincide with the maximal cartel price, and the latter price characterizes the set of feasible cartel prices under profit maximization, we regard both approaches as complementary. The obtained insights are therefore also relevant for profit maximization.

We explore a convenient technique for analyzing the maximal cartel price in general oligopoly models. This technique can be illustrated graphically, appeals directly to economic intuition, and enhances a more general analysis of sustainable cartel prices than was

previously possible. We establish monotonicity properties for the maximal cartel price. This price as a function of the discount factor (or other parameters) exists, but except for some special cases, lacks a closed-form solution. Our method allows numerical implementation, which is of utmost importance in applied economics.

The characterization of the maximal cartel price disentangles the effects of traditional antitrust policies and leniency programs into three separate effects: First, we show that it is the maximum of two other sustainable prices based upon different behavioral assumptions: Such price for cartels that operate silently, and for cartels that systematically collude and report to the AA. Next, the price sustainable by silent cartels is in turn the minimum of another two sustainable prices: The price sustained under a traditional antitrust policy, and the one that induces cartel members to be silent. Each of these underlying prices can be studied separately before being combined to obtain the overall maximal cartel price. The underlying cartel prices depend upon prevailing policies and the industry characteristics, such as the industry structure, the type of competition, and the discount factor.

The aim of introducing a leniency program is to reduce the maximal cartel price below the one under traditional antitrust regulation. Our characterization of the maximal cartel price shows that, within the class of modified grim-trigger strategies, this is impossible without rewards to the single (or the strictly-first) reporting firm. In fact, if the leniency program is wrongly designed, it even may have an adverse effect by enhancing systematic collusion and reporting that increases the maximal cartel price. We derive necessary and sufficient conditions under which such adverse effects are eliminated: Individual fine reductions in case of multiple reporting firms should be moderate.

We also report positive effects of leniency programs. The above impossibility result is derived under two conditions: Reporting firms do not receive rewards, and a price-deviating firm is immune for prosecution. Relaxing the latter assumption under traditional antitrust policies describes current practice in many OECD countries. Then, introducing a leniency program in which a single reporting firm is granted full immunity and adverse effects are avoided, is sufficient to reduce the maximal cartel price to the maximal cartel price under the traditional antitrust policy as if it does not prosecute price-deviating firms. Furthermore, limiting leniency to first-time offenders reduces the adverse effects. These results offer positive support for current practices in many OECD countries.

Another issue examined by Spagnolo (2004) and Rey (2003) is whether offering rewards to single reporting firms would be a good regulation policy, or that it would only introduce adverse effects. To obtain insights, we first characterize the necessary reward to a single reporting firm that would eradicate all cartel prices above the competitive price. Substantial rewards to the strictly-first reporting firm achieve this. Next, we analyze the possibility that the requirement of moderate fine reductions in case of multiple reporting firms allows some slack for implementing such rewards. If possible, such rewards could reduce the maximal cartel price below the one under the traditional antitrust policy without introducing adverse effects. This implies that the EC system can be improved by abolishing the reduced fine

for the second reporter, similar to the US, since this would increase the slack available for introducing rewards.

Our study belongs to a growing literature on the effects of leniency programs in competition policy. Optimal implementation of antitrust policy and leniency programs for cartel enforcement have been analyzed in e.g. Motta and Polo (2003), Rey (2003), Spagnolo (2004, 2008), Harrington (2004, 2005, 2008), Hinloopen (2003, 2006), Motchenkova (2004), Buc-cirossi and Spagnolo (2006), Harrington and Chen (2006), Chen and Rey (2007) and Chen and Harrington (2007). In the next paragraphs, we relate our study to this literature.

The seminal paper on optimal revelation schemes as part of antitrust policy is Motta and Polo (2003), who study, loosely speaking, a Stackelberg game in which the AA first chooses once and for all its antitrust policy followed by the competition phase in which the firms compete with each other, which is modelled as an infinitely-repeated oligopoly game. Market competition is restricted to a discrete set of three prices that captures the three most important profit levels: The profit under perfect competition, under the cartel and the profit of optimally cheating on the cartel price. In each period, firms decide whether to reveal information about their misconduct. The cartel adopts grim-trigger strategies in which cheating on the cartel by either setting a different price or applying for leniency triggers competitive behavior forever, while the cartel continues collusion as usual each time it is caught by the AA. Under the optimal antitrust policy, introduction of ex-post leniency programs will increase the chance of the cartel being captured, but ex-ante leniency programs that grant reduced fines are ineffective. As later shown in Spagnolo (2004) and Rey (2003), effective ex-ante leniency programs require substantial rewards, i.e. pay reporting firms.

Chen and Harrington (2007) incorporate ex-ante leniency programs into a special case of the framework in Harrington (2004 and 2005) to augment traditional antitrust policy in an environment where cartels arouse suspicions. In such environment, cartels also need to manage suspicions, modeled as if the cartel keeps in mind an endogenous detection probability. This probability is modelled as a function of the cartel's price setting, where price is a continuous variable. The focus is on an exogenous antitrust policy in order to study the cartel's optimal reaction on the profit-maximizing cartel price, which implies that cartel formation and the cartel's pricing strategy have become endogenous decisions.

This model can be regarded as the competition phase in Motta and Polo (2003) in a very general setting. The cartel adopts grim-trigger strategies that are similar to those in Motta and Polo (2003), but with the difference that the cartel terminates its illegal business after being caught once by the AA. So, the cartel culture of whether to continue the cartel after being caught differs between Motta and Polo (2003) and Chen and Harrington (2007). In the last reference, the detection probability also depends upon past prices and collusive behavior induces a cumulative liability in the form of fixed fines and private law suites. These two features introduce state variables into the model and this makes the equilibrium nontractable. The analysis, therefore, has to resort to simulations of price paths. Nevertheless, the model admits a steady-state profit-maximizing cartel price that lies above the competitive price,

and this price is independent of the leniency program meaning such program is ineffective.

The foci of our paper are also detection probabilities and penalty schemes that depend upon cartel pricing, a topic that receives relative little attention in the literature. However, such modifications are relevant because these are closer to current antitrust policies. Also, in our paper cartel formation and its pricing strategy are endogenized by modelling it as the solution to an optimization program, similar to e.g. Harrington (2004 and 2005). Our model generalizes Motta and Polo (2003) by including endogenous cartel behavior, the presence of suspicions, the notion of cartel culture¹, and a general class of exogenous antitrust policies with proportional penalty schemes that include the possibility of an effective reward. Several aspects of the model in Chen and Harrington (2007) are also generalized, namely a general oligopoly model instead of Bertrand oligopoly, general penalty schemes with fine reductions that include rewards, and the notion of cartel culture. In fact, our cartel culture unifies the opposite assumptions in Motta and Polo (2003) and Chen and Harrington (2007) with respect to this parameter. Therefore, our model bridges the rudimentary market competition phase in Motta and Polo (2003) with the general approach in Chen and Harrington (2007). A major difference is that all references consider profit maximization by the cartel, whereas we consider the maximal cartel price.

Our analysis is complementary to the analyses in Harrington (2004 and 2005), where the main focus is on a sufficiently large discount factor under which the equilibrium conditions will be non-binding. For such large discount factors, they show that the profit-maximizing cartel price is decreasing in the discount factor. Combining this last result with our results, namely the profit-maximizing cartel price under binding equilibrium conditions coincides with the non-decreasing maximal cartel price, implies that the profit-maximizing cartel price is non-monotonic in the discount factor in case its entire range is considered. Finally, our results also hint at that the ineffectiveness of ex-ante leniency programs without rewards, as reported in Chen and Harrington (2007) under large discount factors, generalizes to the entire range of discount factors.

The paper is organized as follows. Section 2 outlines the model. In Section 3, we analyze the effects of traditional antitrust policies and leniency programs on the maximal cartel price. We discuss the consequences of relaxing some of the main assumptions in Section 4. Section 5 concludes the analysis and discusses several policy recommendations.

2 The Model

There are $n \geq 2$ firms that compete in every of infinitely many periods in the presence of a leniency program. In each period, the firms first choose their prices, and then decide whether

¹Cartel culture in our paper is related to the probability that the cartel resumes business as usual after each conviction by the AA. By imposing strong assumptions about the underlying stochastic process, Bryant and Eckard (1991) estimate that 14 per cent of 1300 firms are recidivist. Although our model is fundamentally different, we treat these estimates as empirical support that the cartel culture belongs to $(0, 1)$ even though the estimate is for US data prior to the major modifications of 1993.

to report in case of collusion.² Therefore, we adopt an infinitely-repeated sequential game, see e.g. Wen (2002), with a common discount factor $\delta \in (0, 1)$ per period and the following two-stage game in any period: All firms first simultaneously set their prices, and after prices are observed and profits are realized, all firms simultaneously decide whether to report. Reporting cannot be kept secret and will be known in the next period, which is called "the hard case of public reporting" by Rey (2003). If some firms report, the leniency program is executed, and otherwise, the AA investigates the market outcome at the end of every period with certain probability and, upon being caught, violators will be fined. We focus on a class of modified grim-trigger strategies to sustain cartel prices, in which deviations in either price setting or self-reporting lead to the competitive equilibrium price in every period thereafter. The underlying rationale is that any cartel is based upon trust among its members and, trust is gone after the first deviation. This setup is common in the literature.

Price competition in each period is described by a symmetric Bertrand model with the n firms, and either homogenous or heterogeneous products.³ Let $\pi(p_1, \dots, p_n)$ be an individual firm's per-period profit for prices $p_1, \dots, p_n \in \mathbb{R}_+$. Since we mostly deal with symmetric outcomes, we denote $\pi(p, \dots, p) \equiv \pi(p)$ for simplicity. Denote the static Nash equilibrium price and the maximal collusive price by p^N and p^M , respectively. In every period, the firms decide whether to collude and if so, to what degree. In other words, all firms choose a price $p = p^N + I(p^M - p^N)$, where $I \in [0, 1]$ can be considered as a collusion index. One key element in analyzing cartel stability is a firm's profit from unilateral deviation against the cartel when all the other firms set their prices at p , denoted as $\pi^{opt}(p) = \sup_{p'} \pi(p', p, \dots, p)$. We deliberately write supremum instead of the maximum in order to have a well-defined value function for the discontinuous case of homogeneous products. As in Harrington (2004, 2005), we assume that $\pi(p)$ is continuous and strictly increasing in $p \in [p^N, p^M]$, $\pi^{opt}(p)$ is continuous, strictly increasing and $\pi^{opt}(p) > \pi(p) > 0$ for $p \in (p^N, p^M]$. Let $\lambda(p)$ be the relative size of the net cartel profit to the net gains under the best unilateral deviation:

$$\lambda(p) = \begin{cases} \frac{\pi(p) - \pi(p^N)}{\pi^{opt}(p) - \pi(p^N)}, & \text{for } p \in (p^N, p^M], \\ 1, & \text{for } p = p^N. \end{cases}$$

As will become clear later, $\lambda(\cdot)$ represents a degree of cartel stability in the sector. The higher $\lambda(\cdot)$, the less incentives each firm has to deviate, and the more stable the cartel is. The assumptions imply that $\lambda(p) \leq \lambda(p^N) \equiv 1$ is continuous in $p \in (p^N, p^M]$, but it might be discontinuous at $p = p^N$, such as the Bertrand model with homogenous products. We denote⁴ $\underline{\lambda} = \lim_{\varepsilon \rightarrow 0+} \lambda(p^N + \varepsilon) \leq 1$ and $\bar{\lambda} = \lambda(p^M)$.

We impose that $\lambda(p)$ is non-increasing for all $p \in (p^N, p^M]$, so that the higher the cartel price, the higher the incentive of each firm to deviate. This mild assumption captures the

²We adapt the economic definition of a cartel that regards pricing above the competitive (Nash) equilibrium price as cartel activity.

³Our analysis immediately applies to quantity competition by making use of the inverse demand function.

⁴The limit $\underline{\lambda}$ is well defined, because the correspondence $\hat{\lambda}(\cdot)$ that solves $\hat{\lambda}(p) [\pi^{opt}(p) - \pi(p^N)] = \pi(p) - \pi(p^N)$ for each $p \in [p^N, p^M]$ is upper semi-continuous for all $p \in [p^N, p^M]$ and it coincides with the continuous function $\lambda(\cdot)$ for all $p > p^N$.

class of heterogeneous Bertrand duopolies with linear demand and constant marginal costs. To simplify the exposition, we normalize the static model so that $\pi(p^N) = 0$ and from now on interpret $\pi(p)$ as the net profit above the profit $\pi(p^N)$ under competition.⁵

There are two different treatments in the literature of how firms behave after being convicted. Harrington (2004) assumes that being caught once is sufficient to deter cartel activity in the future. In Motta and Polo (2003) the economic sector is notorious for cartel activities despite many convictions (meaning firms pay the fines and continue illegal business). We introduce a cartel-culture parameter, denoted $\gamma \in [0, 1]$, that unifies the two different treatments. γ reflects the probability that the firms stop illegal activities after each conviction. Notorious implies $\gamma = 0$, i.e. detection does not dissolve the cartel, while $\gamma = 1$ means the sector becomes competitive after the first detection.

The market outcome is investigated at the end of every period with certain probability and the antitrust policy is implemented. Upon being caught, violators will be fined.

1. Given $p \in [p^N, p^M]$, the firms will be found guilty of collusion with probability $\beta(p) \in [0, 1]$. We assume that $\beta(p) \geq 0$ is a non-decreasing differentiable function on $p \in (p^N, p^M]$ such that $\lim_{\varepsilon \rightarrow 0^+} \beta(p^N + \varepsilon) = \underline{\beta} \geq 0$, and $\beta(p^N) = 0$.
2. If the firms are found guilty of sustaining cartel price $p \in (p^N, p^M]$, every firm will have to pay the one-time fine $k(p)\pi(p)$, where $k(p) \geq 0$ is a non-decreasing differentiable function on $p \in (p^N, p^M]$ such that $\lim_{\varepsilon \rightarrow 0^+} k(p^N + \varepsilon) = \underline{k} \geq 0$, and $k(p^N) = 0$.

The function $\beta(\cdot)$ reflects that a higher cartel price might invoke more attention about cartel abuse and make detection more likely. Any cartel will take the negative impact of its cartel price into account when deciding upon the price. Note that $\beta(p)$ depends on the price in the current period only. As in Rey (2003), only current period's misconduct is prosecuted. The benchmark case of no regulation is captured by $\beta(p)k(p) = 0$ for all $p \in [p^N, p^M]$.

Stylized facts from the OECD countries on antitrust policies, see OECD (2002), suggest detection probabilities in the range of one out of seven or six and levels of fines in the range of two to three times the illegal gains from cartel. These facts imply an expected penalty roughly between 30% to 50% of illegal gains, or $\frac{2}{7} \leq \beta(p)k(p) \leq \frac{1}{2}$, which are of the same order of magnitude as estimated in Bryant and Eckard (1991). In this paper, we assume that $0 \leq \beta(p)k(p) < 1$ for all $p \in (p^N, p^M]$. This less restrictive assumption implies that the expected fine at any price above the competitive price is lower than the (per period) cartel profit and, therefore, any cartel is tempted to set its price above the competitive price.

⁵Note that for any $p \in (p^N, p^M]$, we have

$$\lambda(p) = \frac{\pi(p) - \pi(p^N) - [\pi(p^N) - \pi(p^N)]}{\pi^{opt}(p) - \pi(p^N) - [\pi(p^N) - \pi(p^N)]}.$$

So, it is as if all profits are normalized to net gains above $\pi(p^N)$, and it is without loss of generality to normalize $\pi(p^N) = 0$.

Rey (2003) pointed to the issue whether to prosecute price-deviating firms. In Section 3, we analyze the simpler case in which price-deviating firms are never prosecuted, as in e.g. Motta and Polo (2003), and in Section 4.4 such prosecution is incorporated.

The leniency program is modeled as follows. Let $\alpha(p, s_1, \dots, s_n) \pi(p) \leq k(p) \pi(p)$ be an individual firm's reduced one-time fine in case the cartel price is $p \in [p^N, p^M]$ and the firms' reporting decisions are $s_1, \dots, s_n \in \{R, N\}$, where R (N) stands for (Not) Reporting. Since we mostly deal with symmetric outcomes, we simplify these reduced fines as follows:

1. If none of the firms reports, then the reduced fine $\alpha(p, N, \dots, N) \pi(p)$ to a firm in such a period is equal to its expected fine $\beta(p) k(p) \pi(p)$ from traditional regulation.
2. However, if a *single* firm reports, the reduced fine in this period to the only reporting firm is $\alpha(p, R, N, \dots, N) \pi(p)$, where we write $\alpha(p, N) \equiv \alpha(p, R, N, \dots, N)$.
3. Finally, if *all* firms report, then an individual firm pays the reduced fine $\alpha(p, R, \dots, R) \pi(p)$, where we use $\alpha(p, R) \equiv \alpha(p, R, \dots, R)$.

We assume that $\alpha(p, s)$, $s \in \{R, N\}$, is non-decreasing and continuous in $p \in (p^N, p^M]$ such that $0 \leq \alpha(p, N) \leq \alpha(p, R) \leq k(p)$,⁶ $\lim_{\varepsilon \rightarrow 0+} \alpha(p^N + \varepsilon, s) = \underline{\alpha}(s)$, and $\alpha(p^N, s) = 0$.⁷ Note that $\alpha(p, s) \geq 0$ excludes the possibility that a reporting firm will be rewarded, and we refer to Section 4.1 for rewards. The US leniency program is captured by $\alpha(p, N) = 0$ and the EC leniency program corresponds to $0 \leq \alpha(p, N) \leq k(p)$, where both $\alpha(p, N)$ and $k(p)$ are constant and $\alpha(p, N)$ can be reinterpreted as a constant percentage of fine reduction.

Chen and Rey (2007) address the issue of limited access to leniency for repeated offenders. In Section 3, we analyze the case in which the program always allows repeated offenders to apply for leniency, and in Section 4.3 we discuss the limited access regime. Similar to Rey (2003), we analyze the issue whether to prosecute price-deviating firms under leniency. In Section 4.4 we allow for the possibility of prosecution, but we first analyze the simpler case when such firms are never prosecuted in Section 3.

The equilibrium concept we use is subgame perfect equilibrium and equilibrium conditions are verified by applying the one-stage deviation principle, see e.g., Fudenberg and Tirole (1991). By the one-stage deviation principle, we only need to consider unilateral deviations at appropriate histories of either the price setting stage or the reporting stage. To verify equilibrium conditions, only the expected profits from unilateral one-stage deviations are required on or off the equilibrium path.

⁶This ranking captures the rules of the current leniency guidelines in most OECD countries that allow full amnesty for single self-reporting firms and only partial fine reductions in case of multiple self-reporting. Non-decreasing reduced fines conforms to the legal reasoning that more severe violations, i.e. higher cartel prices, are punished harder.

⁷Although the percentage of fine reductions specified by the current leniency programs in the US and EU do not depend on the severity of the offense, our model is richer and also includes fines and fine reductions that vary with the severity of collusion.

The variable of interest in our study is the socially worst outcome from the consumer point of view. This outcome coincides with the maximal cartel price sustained by modified grim trigger strategy profiles. A cartel forms and sets the maximal cartel price whenever the maximal cartel price is larger than p^N . Therefore, the cartel formation and cartel pricing are both endogenous in our model. Relation of our approach to profit maximization is analyzed in Section 4.2.

3 Maximal Cartel Prices

In this section, we analyze the maximal cartel price under different modes of regulation and cartel behavior. Our analysis will show that three specific cartel prices are important: the highest sustainable cartel price under a traditional antitrust policy, the highest sustainable price for silently operating cartels under leniency, and such price for cartels that systematically report under leniency. Therefore, we first derive these three cartel prices in separate subsections and, finally, analyze the maximal cartel price under leniency in Section 3.4.

3.1 Cartels under Traditional Antitrust Policy

Every cartel has its own destabilizing forces working from within, namely the incentives of individual firms to cheat on the cartel price. In this subsection, we only consider price deviations as if no leniency program is present meaning we ignore any decisions to report.

In order to express these destabilizing incentives, it is essential that we describe what happens within the cartel after some of the firms undercut the cartel price. All firms follow the modified grim-trigger strategy profile to sustain a cartel price of $p > p^N$ that is given by:

1. Firms set a price $p > p^N$ in the first period and continue to set price p as long as there was no (price) deviation. Any deviation by some of the firms leads to 3.
2. As long as there was no deviation, every time the cartel's fraud $p > p^N$ is detected, with probability $1 - \gamma$ firms continue under 1. and with probability γ go to 3.
3. All firms set the competitive equilibrium price p^N in every period.

Under this strategy profile, let $V(p)$ be the present value of a firm's expected profit if the cartel sets price $p \in [p^N, p^M]$ in every period. $V(p)$ is equal to the current illegal net gains $\pi(p)$, minus the expected fine $\beta(p) k(p) \pi(p)$, plus the expected continuation profit of maintaining the cartel after detection $\beta(p) (1 - \gamma) \delta V(p)$, and the expected continuation net gain of not being detected $(1 - \beta(p)) \delta V(p)$. Solving for $V(p)$ yields that

$$V(p) = \frac{1 - \beta(p) k(p)}{1 - \delta [1 - \gamma \beta(p)]} \pi(p). \quad (1)$$

Note that $V(p^N) = 0$ due to $\pi(p^N) = 0$. For all $p \in (p^N, p^M]$, $V(p) > 0$ because $\beta(p) k(p) < 1$. Or, for the opposite case, the cartel would be unprofitable if the AA could set the

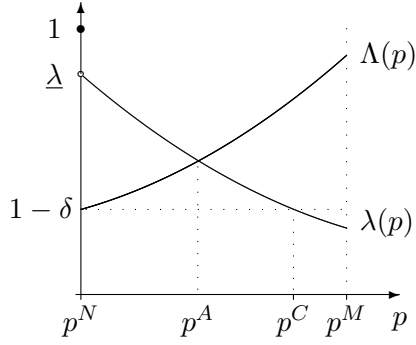


Figure 1: $p^A \leq p^C$ is the largest p such that $\lambda(p) \geq \Lambda(p)$.

expected fine above illegal gains. Since $\beta(p) \geq 0$ and $k(p) \geq 0$ for all $p \in (p^N, p^M]$, we have $V(p) \leq \pi(p)/(1 - \delta)$. The upper bound is the benchmark case $\beta(p) = k(p) = 0$ that reflects the case of no regulation.

Given the modified grim-trigger strategy profile, the profit from a unilateral deviation is equal to the short term net gain $\pi^{opt}(p)$ in the current period, minus an expected fine of zero (no prosecution), plus the normalized profit from the competitive equilibrium forever. So, the necessary and sufficient condition to support cartel price $p \in (p^N, p^M]$ is $V(p) \geq \pi^{opt}(p)$:

$$\frac{1 - \beta(p)k(p)}{1 - \delta[1 - \gamma\beta(p)]}\pi(p) \geq \pi^{opt}(p), \quad (2)$$

which implies that

$$\lambda(p) \geq \Lambda(p) \equiv \frac{1 - \delta[1 - \gamma\beta(p)]}{1 - \beta(p)k(p)}, \quad (3)$$

where $\Lambda(p)$ is a continuous function on $[p^N, p^M]$. Obviously, $p = p^N$ satisfies (3). We establish several properties of the function $\Lambda(\cdot)$. We defer all proofs to the appendix.

Lemma 1 $\Lambda(p)$ is non-decreasing in p and increasing when either $\beta'(p) > 0$ or $k'(p) > 0$. Moreover, $\Lambda(p)$ is increasing in γ , $k(p)$ and $\beta(p)$ and decreasing in δ .

Lemma 1 implies that the right-hand side of (3) increases when γ increases. Hence, a cartel that is more persistent in its collusion (a lower γ) would weaken (3) and, hence, expands the range of sustainable cartel prices. Similar, the right-hand side of (3) is decreasing in δ , meaning that an increase in δ would relax (3). As firms care more about the future, it becomes easier to sustain a cartel price. More importantly, an overall increase in detection probabilities $\beta(p)$ or fines $k(p)$ would make collusion harder to sustain.

The highest cartel price supported by the modified grim-trigger strategy profile under a traditional antitrust policy (A) is given by

$$p^A = \max_{p \in [p^N, p^M]} p, \quad \text{s.t. (3)}. \quad (4)$$

Program (4) is a well-defined program since $p \in [p^N, p^M]$ and (3) induces a closed subinterval of $[p^N, p^M]$ that contains p^N . This insight follows from Figure 1: First, note that $\Lambda(p^N) =$

$1 - \delta$ and $\lim_{\varepsilon \rightarrow 0} \lambda(p^N + \varepsilon) = \underline{\lambda} \leq 1 = \lambda(p^N)$ implies that $p = p^N$ always satisfies (3). Second, since $\Lambda(p)$ is non-decreasing and $\lambda(p)$ is non-increasing, the intersection $\lambda(p) = \Lambda(p)$ is either a unique price or a closed subinterval of $[p^N, p^M]$. Furthermore, if $p \in (p^N, p^M]$ is any such price, then all lower prices can also be sustained by the cartel. So, the range of prices that can be sustained as cartel prices in (3) is a (possible degenerated) closed subinterval of $[p^N, p^M]$ that contains p^N . If some intersection point exists, as in Figure 1, the highest cartel price $p^A = \max\{p : \lambda(p) = \Lambda(p)\}$ implying $p^A > p^N$. Otherwise, p^A is equal to either p^N whenever $\underline{\lambda} < 1 - \delta$, or p^M if $\bar{\lambda} \geq \Lambda(p^M)$. The intuition is that higher cartel prices put an upward pressure on the right-hand side of (3). This reduces the sustainability of p and only sectors with values of $\lambda(p)$ relatively close to 1 might withstand this pressure, meaning the short-term gains of blowing up the cartel must be close enough to the cartel profit $\pi(p)$.

In what follows, we are also interested in the effectiveness of antitrust regulation (and later leniency programs) compared to the benchmark case $\beta(p) = k(p) = 0$ of no regulation. Then, the Λ function becomes the constant function $\Lambda(p) = 1 - \delta$ and Program (4) specifies the special case of the highest collusive price (C) in the absence of regulation:

$$p^C = \max_{p \in [p^N, p^M]} p, \quad \text{s.t. } \lambda(p) \geq 1 - \delta, \quad (5)$$

Figure 1 also illustrates p^C . Note that $\Lambda(p) \geq 1 - \delta$ for all $p \in [p^N, p^M]$ implies that any p that satisfies (3) also satisfies the constraint in (5). Comparing (4) and (5), we observe that regulation may reduce the highest cartel price: $p^N \leq p^A \leq p^C \leq p^M$.

Now, we explain how we characterize highest cartel prices. A direct approach would be to solve (3) for p as a function of all parameters. However, it is not clear how to solve (3). Instead, our characterization is based upon analyzing properties of the threshold level for δ as a function of $p \in [p^N, p^M]$ in the (p, δ) -space, and then, translate these properties to the highest cartel price as a function of δ in the (δ, p) -plane.

To be more specific, rewriting (3) yields the threshold on the discount factor δ for sustaining the arbitrary cartel price $p \in (p^N, p^M]$:

$$\delta \geq \Delta(p) \equiv \frac{1 - \lambda(p) [1 - \beta(p) k(p)]}{1 - \gamma \beta(p)}. \quad (6)$$

We establish the following properties of the function $\Delta(\cdot)$.

Lemma 2 *$\Delta(p)$ is continuous and non-decreasing in $p \in (p^N, p^M]$, $\Delta(p) \geq 1 - \lambda(p)$ (with strict inequality when $\beta(p) > 0$) and increasing in $k(p)$ and $\beta(p)$.*

The monotonicity property with respect to p implies that taking the inverse relationship of (6) for values of $\delta \in [\Delta(p^N), \Delta(p^M)]$ is well-defined and would in principle yield $p^A = \Delta^{-1}(\delta)$. Since the function $\Delta(p)$ is a nontrivial function of p , taking this inverse is practically impossible except for specific functional forms such as Example 5 at the end of this subsection. The following result states general properties of the curve $p^A = \Delta^{-1}(\delta)$.

Proposition 3 *Under any antitrust regulation, the highest cartel price p^A is non-decreasing in $\delta \in (0, 1)$ and non-increasing in $\gamma \in [0, 1]$. Furthermore, we have*

$$\begin{aligned} p^A &= p^N, & \text{for } \delta \in (0, 1 - \underline{\lambda}), \\ p^A &\in [p^N, p^M), & \text{for } \delta \in [1 - \underline{\lambda}, \Delta(p^M)), \\ p^A &= p^M, & \text{for } \delta \in [\Delta(p^M), 1), \end{aligned}$$

An overall increase in detection $\beta(p)$ or $k(p)$ shifts $\Delta(p^M)$ and the entire curve to the right.

As is demonstrated by Example 5, the first range must be open and the middle range can be empty. For comparison, inserting $\beta(p) = k(p) = 0$ in the function $\Delta(\cdot)$ yields

$$\begin{aligned} p^C &= p^N, & \text{for } \delta \in (0, 1 - \underline{\lambda}), \\ p^C &\in [p^N, p^M), & \text{for } \delta \in [1 - \underline{\lambda}, 1 - \bar{\lambda}), \\ p^C &= p^M, & \text{for } \delta \in [1 - \bar{\lambda}, 1). \end{aligned}$$

Since Proposition 3 also implies that $0 < 1 - \underline{\lambda} \leq 1 - \bar{\lambda} \leq \Delta(p^M)$, antitrust regulation with positive expected fines restricts the set of discount factors for which collusion can be sustained at a given $p \in (p^N, p^M]$. In other words, antitrust regulation reduces the highest sustainable cartel price. Combining all inequalities implies:

Corollary 4 *Depending on the discount factor δ , the effectiveness of regulation is given by*

$$\begin{aligned} p^N &= p^A = p^C = p^N, & \text{if } \delta \in (0, 1 - \underline{\lambda}), \\ p^N &< p^A < p^C < p^M, & \text{if } \delta \in [1 - \underline{\lambda}, 1 - \bar{\lambda}), \\ p^N &< p^A < p^C = p^M, & \text{if } \delta \in [1 - \bar{\lambda}, \Delta(p^M)), \\ p^M &= p^A = p^C = p^M, & \text{if } \delta \in [\Delta(p^M), 1). \end{aligned}$$

When $\delta \in [\Delta(p^M), 1)$, the antitrust policy is not effective to deter the cartel from setting its monopoly price, i.e. $p^A = p^M$, when interval $\delta \in [\Delta(p^M), 1)$ is non-empty. This implies that the necessary and sufficient condition for $p^A = p^M$ is $\Delta(p^M) < 1$. By (6), $\Delta(p^M) < 1$ requires that condition (3) for cartel stability is broken at $\delta = 1$, i.e.

$$\frac{\lambda(p^M)}{\gamma} > \frac{\beta(p^M)}{1 - \beta(p^M)k(p^M)}. \quad (7)$$

This inequality is written to separate the sector characteristics $(\gamma, \lambda(p))$ from the policy instruments $(\beta(p), k(p))$. In economic applications, (7) can be easily verified numerically by calibrating sector characteristics in order to obtain thresholds for policy parameters.

Given a particular antitrust policy, it is interesting to investigate whether this policy can eradicate the monopoly price for all cartel cultures. Solving condition (7) for γ yields

$$\gamma < \frac{\lambda(p^M) [1 - k(p^M)\beta(p^M)]}{\beta(p^M)}$$

To destabilize cartels for all possible cartel cultures $\gamma \in [0, 1]$, the right-hand side must be negative, i.e. $\beta(p^M)k(p^M) > 1$. Hence, under any cartel policy that satisfies condition $0 < \beta(p)k(p) < 1$ for all $p \in (p^N, p^M]$, sectors that are notorious for persistent cartel behavior

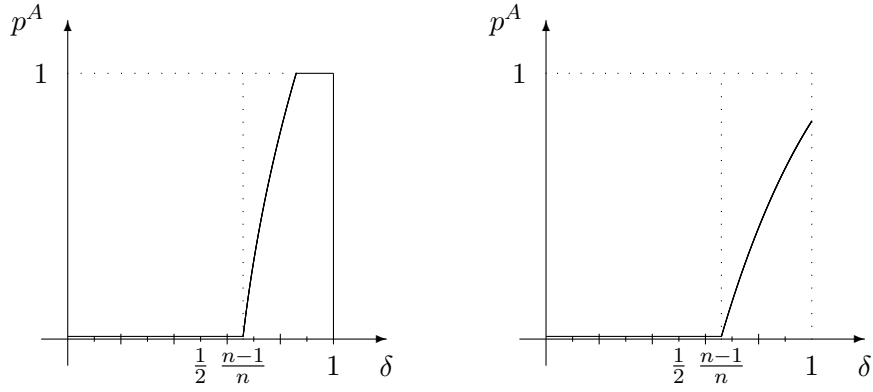


Figure 2: Two cases for the curve of p^A in Example 5.

(γ close to 0) cannot be eradicated by the antitrust policy unless one is willing to adopt a policy that fully takes away the illegal gains, i.e. $k(p)\beta(p) > 1$ for all $p \in (p^N, p^M]$.

Another important conclusion is related to the impact of the level of the function $\lambda(\cdot)$ on sustainability of the monopoly price. Recall the definition of $\lambda(\cdot)$, through (3) cartel stability is related to a particular sector characteristic. Condition (7) relaxes for sectors where the degree of cartel stability is higher, i.e. $\hat{\lambda}(\cdot) > \lambda(\cdot)$ for all $p \in (p^N, p^M]$, and this makes sustaining higher cartel prices possible. Also, the function $\Delta(p)$ shifts downward in sectors with a higher cartel stability $\hat{\lambda}(\cdot)$. So that in these sectors, we should expect that regulation will be less effective compared to sectors with lower $\lambda(\cdot)$.

The main message of this section is a mixed blessing for antitrust regulation. On the one hand, Corollary 4 identifies non-empty sets of parameter values for which antitrust regulation is effective in reducing the highest cartel price. On the other hand, as long as the legal system obeys condition (7), there always will be a non-empty set of parameter values for which $p^A = p^M$, meaning the antitrust policy is totally ineffective on this set.

In general, it is impossible to obtain a closed-form solution for $p^A = \Delta^{-1}(\delta)$ for $\delta \in [1 - \underline{\lambda}, \Delta(p^M))$. In economic applications, however, one can resort to a simple numerical implementation of our approach: Numerically calculate the monotonic values of $\Delta(p)$ in the (p, δ) -space and, then, by reversing the dependence, plot the numerical values in the (δ, p) -space for $\delta \in [1 - \underline{\lambda}, \Delta(p^M))$ to obtain $p^A = \Delta^{-1}(\delta)$. For all other values of δ , p^A is either p^N or p^M . In some cases, closed-form solutions can be derived as our next example illustrates.

Example 5 Consider a homogeneous Bertrand oligopoly model with linear demand $y = 2 - p$ and constant marginal costs of 0. The antitrust regulation is given by $\beta(p) = \beta p$ and $k(p) = k$, where $k\beta < 1$. Note that $p^N = 0$, $p^M = 1$, and $\pi^{opt}(p) = n\pi(p)$ for all $p \in (p^N, p^M]$. Consequently, $\lambda(p^N) = 1$ and $\lambda(p) = \frac{1}{n}$ for all $p \in (p^N, p^M]$ is a discontinuous function with $\underline{\lambda} = \bar{\lambda} = \frac{1}{n}$. Program (4) becomes

$$p^A = \max_{p \in [0, 1]} p, \quad \text{s.t.} \quad \frac{1}{n} \geq \frac{1 - \delta + \gamma\delta\beta p}{1 - k\beta p}.$$

Note that $p = p^N = 0$ is feasible in the constraint if and only if $\delta \geq 1 - \frac{1}{n} = 1 - \underline{\lambda}$. The

constraint can be rewritten as

$$p \leq \frac{1 - n(1 - \delta)}{(n\gamma\delta + k)\beta}$$

and the upper bound is the solution p^A whenever it is between 0 and 1. The right hand side is increasing in δ , and decreasing in γ , β and k . To summarize, we have

$$p^A = \begin{cases} 0, & \text{for } \delta \in (0, 1 - \frac{1}{n}), \\ \frac{1 - n(1 - \delta)}{(n\gamma\delta + k)\beta}, & \text{for } \delta \in [1 - \frac{1}{n}, \frac{n - (1 - \beta k)}{n(1 - \gamma\beta)}), \\ 1, & \text{for } \delta \in [\frac{n - (1 - \beta k)}{n(1 - \gamma\beta)}, 1), \end{cases}$$

where application of Proposition 3 yields the intervals for δ . Note that $\beta = k = 0$, i.e. no antitrust regulation, implies the third interval is equal to $[1 - \frac{1}{n}, 1)$ and, therefore, $p^C = 1$ for all $\delta \in [1 - \frac{1}{n}, 1)$. Note that the third interval is empty if and only if $(n\gamma + k)\beta \geq 1$. Since $\beta k < 1$, this condition can hold only when $n\gamma$ is sufficiently large. Then, p^A is less than $p^M = 1$ for all $\delta \in (0, 1)$. For sectors with a small number of firms and a persistent cartel culture (γ close to 0), the monopoly price will not be eradicated. The two mutually exclusive cases for the curve of highest prices p^A are illustrated by Figure 2, where the vertical dotted line at $\delta = 1 - \frac{1}{n}$ represents the discontinuous jump of p^C from $p^N = 0$ to $p^M = 1$.

3.2 Silent Cartels under Leniency

Since traditional antitrust policy alone may not be sufficient to eradicate all cartel prices, such policies in OECD countries are augmented with leniency programs. In this section, we analyze leniency programs under modified grim-trigger strategy profiles in which firms never report, meaning cartels operate silently.

All firms follow the modified grim-trigger strategy profile to sustain a cartel price of $p > p^N$ that is given by:

1. Firms set a price $p > p^N$ and do *not* report in the first period and continue to do so as long as there was no deviation. Any deviation by some of the firms leads to 4.
2. As long as there was no deviation, every time the cartel's fraud $p > p^N$ is detected, with probability $1 - \gamma$ firms continue under 1. and with probability γ go to 4.
3. A price-deviating firm does not report in the period of deviation.
4. All firms set the competitive equilibrium price p^N and keep silent in every period.

Note, that the cartel breaks down if a firm either undercuts the cartel price or reports.⁸ Given this strategy profile, the present value of an individual firm's expected profits is equal to $V(p)$ in (1) for the same reasons.

⁸Under the alternative assumption that reporting will not lead to a breakdown of the cartel, the interval of sustainable prices shrinks. To see this, modifying this section's arguments implies (8) below becomes $1 - \alpha(p, N) \leq \frac{1 - \delta}{\Lambda(p)}$, and any p satisfying this condition also satisfies (8), but not vice versa. Since we concentrate on the maximal cartel price, we do not further discuss this alternative assumption.

In this modified grim-trigger strategies, there are two types of unilateral one-stage deviations: undercut the cartel price and report to the AA. Furthermore, there are three relevant histories, two on the equilibrium path and one off the equilibrium path just after a unilateral price deviation. Under the assumption that a price-deviating firm is never fined, it has no incentive to seek protection from fines after any history following its own deviation and, therefore, refraining from reporting is optimal in this subgame. Next, on the equilibrium path, inequality (3) is still the necessary and sufficient condition under which no firm would undercut the cartel price.

We now consider refraining from reporting on the equilibrium path. The expected continuation profit of the one-stage deviation by reporting consists of paying the reduced fine $\alpha(p, N) \pi(p)$ in the current period followed by the continuation profits from the competitive equilibrium price p^N forever after. Therefore, a firm will not report if and only if

$$-\alpha(p, N) \pi(p) \leq \beta(p) [-k(p)\pi(p) + \delta(1 - \gamma)V(p)] + [1 - \beta(p)] \delta V(p) = V(p) - \pi(p),$$

which simplifies to

$$1 - \alpha(p, N) \leq \frac{V(p)}{\pi(p)} = \frac{1}{\Lambda(p)}. \quad (8)$$

Condition (8) implies that, in order to break the silence of no reporting, the fine reduction should be sufficiently low, i.e. $\alpha(p, N) \leq k(p)$. Since $\Lambda(p) \geq 0$, any leniency program that offers a reduced fine of $\alpha(p, N) \geq 1$, i.e. the reduced fine is higher than a single period's profit, trivially satisfies (8) and fails to be effective. Therefore, it is without loss of generality to consider $0 \leq \alpha(p, N) < 1$ for all $p \in [p^N, p^M]$. Then, (8) can be rewritten as

$$\frac{1}{1 - \alpha(p, N)} \geq \Lambda(p). \quad (9)$$

Since $\alpha(p, N)$ is non-decreasing in p , we have the following result.

Lemma 6 $\frac{1}{1 - \alpha(p, N)}$ is non-decreasing in p and increasing when $\alpha(p, N)$ is increasing in p . Moreover, $\lim_{\varepsilon \rightarrow 0} \frac{1}{1 - \alpha(p^N + \varepsilon, N)} = \frac{1}{1 - \alpha(N)} \geq 1 = \frac{1}{1 - \alpha(p^N, N)}$.

Lemma 1 and 6 imply that the functions on both sides of (9) are non-decreasing in p . Therefore, these functions may cross more than once.

Under a leniency program, the highest cartel price for a silent cartel (S) is given by

$$p^S = \max_{p \in [p^N, p^M]} p, \quad \text{s.t. (3) and (9)}. \quad (10)$$

Program (10) is a well-defined program. This result follows from Figure 3: First, $p = p^N$ always satisfies (9) due to $\Lambda(p^N) = 1 - \delta$ and $\lim_{\varepsilon \rightarrow 0} \frac{1}{1 - \alpha(p^N + \varepsilon, N)} \geq 1$. Then by the results in Section 3.1, $p = p^N$ satisfies both constraints in (10). Next, although $\Lambda(p)$ and $\frac{1}{1 - \alpha(p, N)}$ might have multiple crossings, continuity of both functions implies that the set of prices such that (9) holds is the union of possibly degenerated non-empty and compact subintervals of

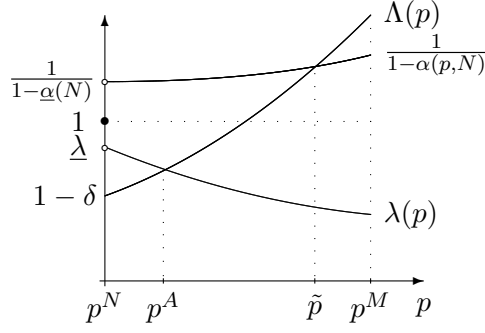


Figure 3: The price \tilde{p} is the smallest p such that $\frac{1}{1-\alpha(p,N)} = \Lambda(p)$. And $\tilde{p} \geq p^A$.

$[p^N, p^M]$. Taking the intersection of the union of subintervals with the subinterval implied by (3) preserves the last property on a subset of the subinterval $[p^N, p^A]$. Combining these facts, implies that Program (10) induces a well-defined highest price on $[p^N, p^M]$.

For our analysis, let $\tilde{p} \in (p^N, p^M]$ be the smallest intersection point p such that $\frac{1}{1-\alpha(p,N)} = \Lambda(p)$, then all lower prices satisfy (9). So, the cartel can keep secret the range of prices $[p^N, \tilde{p}]$. It is without loss of generality to treat \tilde{p} as the highest cartel price that can be sustained by (9). To see this, both \tilde{p} and p^A are intersections of some curve with the curve $\Lambda(\cdot)$. By assumption, $\frac{1}{1-\alpha(p,N)} \geq 1$ for all $p \in [p^N, p^M]$, and by definition, $\lambda(p) \leq 1$ for all $p \in [p^N, p^M]$ implies $\lambda(p) \leq 1 \leq \frac{1}{1-\alpha(p,N)}$ for all $p \in [p^N, p^M]$. Since $\Lambda(p^N) < 1$ and $\Lambda(\cdot)$ is non-decreasing, the curve $\Lambda(\cdot)$ must first intersect $\lambda(\cdot)$ before it intersects the curve $\frac{1}{1-\alpha(\cdot,N)}$, as illustrated by Figure 3. Hence, $p^A \leq \tilde{p}$. Or, (3) is always binding in Program (10), whereas (9) is never binding and can therefore be disregarded. Hence, we have:

Proposition 7 *Under (ex-ante) leniency programs and modified grim-trigger strategy profiles in which firms operate silently, we have $p^S = p^A$.*

The implication of Proposition 7 is that if the cartel can sustain the cartel price $p \in [p^N, p^A]$ under traditional regulation, then introducing an leniency program without rewards allows the cartel to maintain its illegal activity with the same cartel price by operating silently. So, the introduction of such a program to existing antitrust regulation is not effective in reducing cartel prices. Proposition 7 confirms the findings in Motta and Polo (2003), Spagnolo (2004) and Rey (2003) and extends these to our generalized setting. Treating the cartel price as a continuous variable allows to employ a more powerful technique and derive richer results. Given the low expected fines implied by the current policies of most OECD countries, i.e. $\beta(p)k(p) < 1$, we obtain that any fine reduction for the first-reporting firm under a leniency program (when rewards are not allowed) would not be sufficient neither to reduce the existing cartel price $p \in (p^N, p^A]$, nor to block cartel formation.

3.3 Systematically Reporting Cartels under Leniency

Spagnolo (2004) argues that cartels may collude and apply for leniency every period as part of the collusive agreement. This behavior may reduce fines below expected fines from operating

silently. Such behavior is observed in experiments by Hinloopen and Soetevent (2008) with pricing below the monopoly price. In this section, we analyze modified grim-trigger strategy profiles in which firms systematically collude and report in every period.

All firms follow the modified grim-trigger strategy profile to sustain a cartel price of $p > p^N$ that is given by:

1. Firms set a price $p > p^N$ and *report* in the first period and continue to do so as long as there was no *price* deviation. Any price deviation by some of the firms leads to 3.
2. A price-deviating firm does not report in the period of deviation.
3. All firms set the competitive equilibrium price p^N and keep silent in every period.

Note that the cartel breaks down after a firm undercuts the cartel price, but not if a firm refrains from reporting. For $p \in [p^N, p^M]$, this strategy profile induces a present value equal to $[1 - \alpha(p, R)] \pi(p) / (1 - \delta)$.

As in Section 3.2, there are two types of deviations and three relevant histories. A firm that refrains from reporting after all firms set $p \in (p^N, p^M]$ will have to pay a higher fine, which hurts in the current period while it does not cause the cartel to break down. So, reporting after setting cartel price $p \in (p^N, p^M]$ is trivially optimal. Similar as in Section 3.2, since a price-deviating firm is never fined, such firm has no incentive to report after his own price deviation. So, both conditions can be ignored and only unilateral price deviations on the equilibrium path matter. A firm does not set a deviating price if and only if

$$\frac{1 - \alpha(p, R)}{1 - \delta} \pi(p) \geq \pi^{opt}(p).$$

Since $\alpha(p, R) < 1$ is necessary for this condition to hold, it can be rewritten as

$$\lambda(p) \geq \frac{1 - \delta}{1 - \alpha(p, R)} > 0. \quad (11)$$

Since $\alpha(p, R)$ is non-decreasing, the following properties hold.

Lemma 8 $\frac{1-\delta}{1-\alpha(p,R)}$ is non-decreasing in p , increasing in p when $\alpha(p, R)$ is increasing in p , and decreasing in δ . Finally, $\lim_{\varepsilon \rightarrow 0} \frac{1-\delta}{1-\alpha(p^N+\varepsilon, R)} = \frac{1-\delta}{1-\alpha(R)} \geq 1 - \delta = \frac{1-\delta}{1-\alpha(p^N, R)}$.

Lemma 8 implies that both functions in (11) have opposite weak monotonicity properties, which is similar to (3). Furthermore, the right-hand side of (11) decreases in δ , meaning that an increase in δ would relax (11), and it becomes easier to sustain a collusive price.

Under leniency, the highest cartel price of always colluding and reporting (R) is given by

$$p^R = \max_{p \in [p^N, p^M]} p, \text{ s.t. (11)}. \quad (12)$$

For reasons similar to Section 3.1, (12) is a well-defined program since $p \in [p^N, p^M]$ and (11) induces a closed subinterval of $[p^N, p^M]$ that contains p^N , where the latter follows from

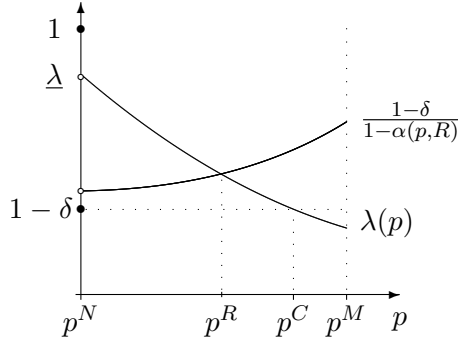


Figure 4: The maximal cartel price p^R is the largest p such that $\frac{1-\delta}{1-\alpha(p,R)} \geq \lambda(p)$.

$\frac{1-\delta}{1-\alpha(p^N,R)} = 1-\delta < 1 = \lambda(p^N)$. If some intersection point of (11) exists, as Figure 4 illustrates, the highest cartel price $p^R = \left\{ \max p : \lambda(p) = \frac{1-\delta}{1-\alpha(p,R)} \right\}$ implying $p^R > p^N$. Otherwise, p^R is equal to either p^N whenever $\underline{\lambda} < 1 - \delta$, or p^M if $\bar{\lambda} \geq \frac{1-\delta}{1-\alpha(p^M,R)}$. The intuition is that higher cartel prices put an upward pressure on the right-hand side of (11). This reduces the sustainability of p and only sectors with values of $\lambda(p)$ relatively close to 1 might withstand this pressure. Since the right-hand side of (11) is at least $1 - \delta$, comparing (5) and (12) implies: $p^N \leq p^R \leq p^C \leq p^M$.

Similar to Section 3.1, we characterize the highest cartel price p^R . Rewriting (11) yields the threshold on the discount factor for sustaining the arbitrary cartel price $p \in [p^N, p^M]$;

$$\delta \geq \Phi(p) \equiv 1 - \lambda(p) [1 - \alpha(p, R)]. \quad (13)$$

Lemma 9 $\Phi(p)$ is continuous and non-decreasing in $p \in (p^N, p^M]$, and $\Phi(p) \geq 1 - \lambda(p)$ is increasing in $\alpha(p, R)$.

Lemma 9 implies a higher threshold for δ when compared to the benchmark $1 - \lambda(p)$.

The monotonicity property with respect to p implies that taking the inverse relationship of (13) for values $\delta \in [\Phi(p^N), \Phi(p^M)]$ is well-defined and yields $p^R = \Phi^{-1}(\delta)$. For reasons similar to Section 3.1, a closed-form solution for this inverse only exists for specific functional forms. Therefore, our next result states general properties of the curve $p^R = \Phi^{-1}(\delta)$.

Proposition 10 The highest cartel price p^R is non-decreasing in $\delta \in (0, 1)$. Furthermore,

$$\begin{aligned} p^R &= p^N, & \text{for } \delta \in (0, 1 - \underline{\lambda}), \\ p^R &\in [p^N, p^M), & \text{for } \delta \in [1 - \underline{\lambda}, \Phi(p^M)), \\ p^R &= p^M, & \text{for } \delta \in [\Phi(p^M), 1), \end{aligned}$$

where $\Phi(p^M) < 1$.

To conclude this subsection, we characterize pairs in the (p, α) -space consisting of a cartel price p and a reduced fine $\alpha(p, R)$ such that the strategy profile with cartel price p and reporting form an equilibrium. To do so, we rewrite (11) as

$$\alpha(p, R) \leq 1 - \frac{1 - \delta}{\lambda(p)}. \quad (14)$$

We establish the following result.

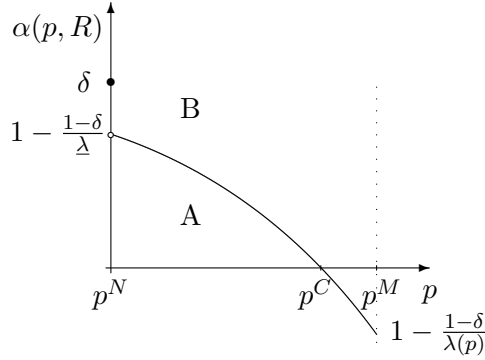


Figure 5: Systematic reporting is sustainable if p and $\alpha(p, R)$ belong to region A .

Lemma 11 $1 - \frac{1-\delta}{\lambda(p)}$ is non-increasing in p , decreasing in p when $\lambda(p)$ is decreasing, increasing in δ , and has image $[0, 1 - \frac{1-\delta}{\lambda}] \subset [0, \delta]$ if and only if $p \in [p^N, p^C]$.

Condition (14) implies two regions, called A and B , as illustrated in Figure 5 under $\alpha(p, N) = 0$ for explanatory reasons.⁹ For any pair p and $\alpha(p, R)$ in region A , condition (14) holds and the strategy profile with systematic collusion on cartel price p and reporting forms an equilibrium. So, if the curve $\alpha(p, R)$ crosses region A , then $p^R \in (p^N, p^C]$ and systematic collusion and reporting occurs. In region B , condition (14) does not hold and such profile cannot form an equilibrium.

Hence, systematic collusion and reporting cartel sustains price p if and only if p and $\alpha(p, R)$ belong to region A . Since upper bound (14) on the fine reduction is non-negative, there always exists a range of substantial fine reductions (without rewards) marked by region A where leniency programs provide sufficient incentives for reporting. But this is also the region where we observe, so-called, adverse effects of leniency in the sense that reporting becomes a part of the collusive agreement without dissolving the cartel.

Cartels do dissolve whenever the curve $\alpha(\cdot, R)$ runs entirely through region B . Then, any $p \in (p^N, p^M]$ fails (14) and $p^R = p^N$. In other words, whenever $\alpha(p, R) > 1 - \frac{1-\delta}{\lambda(p)}$ for all $p \in (p^N, p^M]$, then equilibria in strategy profiles with systematic collusion and reporting are ruled out. Moreover, the cartel dissolves due to price-deviations. This implies that the regulator can eliminate the adverse effects of leniency programs and prevent cartel formation. For that the design of leniency guidelines should avoid substantial fine reductions $\alpha(p, R)$ in case of multiple reporting, i.e. such fine reductions should at most be moderate.

The next example illustrates that in some cases closed-form solutions can be derived.

Example 12 Consider once more the homogeneous Bertrand oligopoly model of Example 5 and the linear reduced-fine function $\alpha(p, R) = \alpha p$, where $\alpha \in [0, 1)$ satisfies $\alpha(p, R) < 1$ for all $p \in [p^N, p^M]$. Program (12) becomes

$$p^R = \max_{p \in [0, 1]} p, \quad s.t. \quad \frac{1}{n} \geq \frac{1 - \delta}{1 - \alpha p}$$

⁹Since $\alpha(p, R) \geq \alpha(p, N)$, the non-decreasing curve $\alpha(p, N) \geq 0$ also puts a lower bound on region A , which is not drawn.

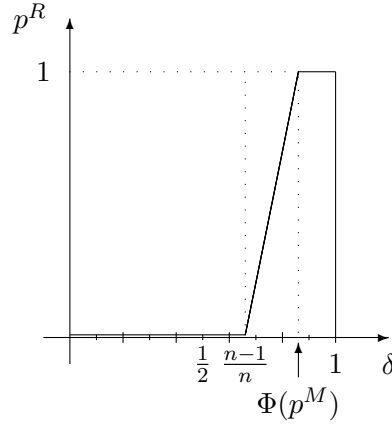


Figure 6: The curve of the maximal cartel price p^R in Example 12.

Note that $p = p^N = 0$ is feasible in the constraint if and only if $\delta \geq 1 - \frac{1}{n} = 1 - \underline{\lambda}$. Proceeding similar as in Example 5, we rewrite the constraint as $p \leq [1 - n(1 - \delta)] / \alpha$, and obtain

$$p^R = \begin{cases} 0, & \text{for } \delta \in (0, 1 - \frac{1}{n}), \\ \frac{1-n(1-\delta)}{\alpha}, & \text{for } \delta \in [1 - \frac{1}{n}, 1 - \frac{1}{n}[1 - \alpha]), \\ 1, & \text{for } \delta \in [1 - \frac{1}{n}[1 - \alpha], 1), \end{cases}$$

where application of Proposition 10 yields the intervals for δ . Since $1 - \frac{1}{n}(1 - \alpha) < 1$, the monopoly price will not be eradicated if firms are sufficiently patient. For $\delta \in [1 - \frac{1}{n}, 1)$, the curve in Figure 5 is the constant function $1 - n(1 - \delta) \geq 0$ implying $p^C = p^M = 1$ and region A is a rectangle. Figure 6 illustrates the highest cartel price as a function of δ .

3.4 The Maximal Cartel Price under Leniency

In this section, we characterize the maximal cartel price for the entire class of grim-trigger strategy profiles, which consists of the union of those studied in Section 3.2 and 3.3.¹⁰

Firms follow a modified grim-trigger strategy profile that serves the cartel's purpose of supporting cartel prices. In some cases, it is just one type of the strategy profiles considered in Section 3.2 and 3.3 that serves this goal, in other cases, both types do. In the latter case, we investigate whether the leniency program is exploitable. Since prices are implicit in the definition proposed in Spagnolo (2004), we extend his definition and call a leniency program exploitable at cartel price p whenever both strategy profiles of Section 3.2 and 3.3 support p as an equilibrium price and $\frac{1-\alpha(p,R)}{1-\delta}\pi(p) > V(p)$.

The maximal cartel price is defined as the maximal cartel price subject to the appropriate equilibrium conditions. These are either the equilibrium conditions (3) and (9) in Section 3.2, or (11) in Section 3.3. Formally, the maximum cartel price under leniency (L) for the entire class of grim-trigger strategy profiles is given by

$$p^L = \max_{p \in [p^N, p^M]} p, \text{ s.t. either (3) and (9), or (11).} \quad (15)$$

¹⁰Obviously, the union also includes the modified grim-trigger strategy profiles discussed in Footnote 8.

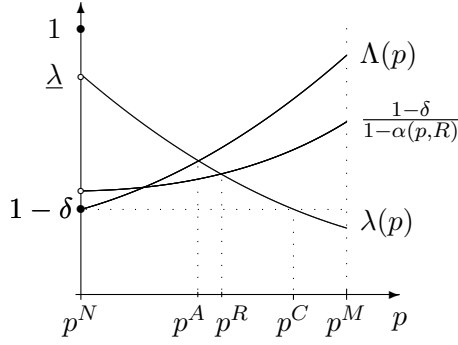


Figure 7: The maximal cartel price p^L is the maximum of p^A and p^R .

By our previous results, the conditions can be translated into two intervals of sustainable cartel prices, i.e., either $p \in [p^N, p^S]$ or $p \in [p^N, p^R]$, and therefore the constraints in (15) are equivalent to $[p^N, p^S] \cup [p^N, p^R] = [p^N, \max \{p^S, p^R\}]$. Since also $p^S = p^A$ by Proposition 7, the following result is immediate.

Proposition 13 *Under (ex-ante) leniency programs and modified grim-trigger strategy profiles, we have $p^L = \max \{p^S, p^R\} \geq p^A$.*

The implication of Proposition 13 is negative: Given that we consider arbitrary policy functions $\beta(p)$ and $k(p)$ such that $\beta(p)k(p) < 1$, it is impossible to find a policy function $\alpha(p, R) \geq 0$ that will reduce the range of sustainable cartel prices to a strict subinterval of $[p^N, p^A]$. To the contrary, introducing a leniency program to existing antitrust regulation may even enhance cartels by allowing them to sustain higher prices. At best, the leniency program achieves $p^L = p^A$ and, at worst, it achieves $p^L > p^A$. If $p^L = p^A$, then the maximal cartel price can be supported by a strategy profile in which the cartel operates silently and the leniency program fails to break this price and the silence. Observing silence is evidence of $p^L = p^A$. Otherwise, i.e. $p^L > p^A$, the maximal cartel price cannot be sustained under traditional regulation, but it can be sustained by a strategy profile in which the cartel systematically colludes and reports to the AA. In this case, observing reporting is evidence of an increased range of sustainable cartel prices!

The necessary and sufficient conditions for $p^L = p^R > p^A$ are somewhat involved. For explanatory reasons, consider Figure 7. It illustrates the case in which $p^R > p^A$ is associated with a unique crossing of the curves $\Lambda(\cdot)$ and $\lambda(\cdot)$ at a lower cartel price than the unique crossing at some higher price of the curves $\frac{1-\delta}{1-\alpha(\cdot, R)}$ and $\lambda(\cdot)$. So, the curve $\Lambda(\cdot)$ at $p = p^A$ lies above the curve $\frac{1-\delta}{1-\alpha(\cdot, R)}$ at $p = p^A$ and this suffices for $p^L = p^R > p^A$. To put it differently, p^A and $\alpha(p^A, R)$ belong to Figure 5's region A, but not to this region's upper boundary. However, since intersections of these curves may involve horizontal line pieces it may also be the case that all three curves have line pieces that partly overlap and if the curve $\Lambda(p)$ starts increasing at $p = p^A$ while the other two curves remain horizontal for p in some non-empty interval $(p^A, p^A + \varepsilon]$, for some $\varepsilon > 0$, then $p^L = p^R \geq p^A + \varepsilon > p^A$. The next lemma states the necessary and sufficient conditions for $p^L > p^A$.

Proposition 14 $p^L > p^A$ if and only if either i) $\alpha(p^A, R) < 1 - \frac{1-\delta}{\lambda(p^A)}$, or ii) $\alpha(p^A, R) = 1 - \frac{1-\delta}{\lambda(p^A)} = 1 - \frac{1-\delta}{\Lambda(p^A)}$, and there exists some $\bar{\varepsilon} > 0$ such that both $\lambda(\cdot)$ and $\frac{1-\delta}{1-\alpha(\cdot, R)}$ are constant on $[p^A, p^A + \bar{\varepsilon}]$ and $\Lambda(\cdot)$ is strictly increasing on $(p^A, p^A + \bar{\varepsilon}]$.

Proposition 14 states conditions under which the maximal cartel price increases due to leniency. In short, if the fine reductions for multiple reporting firms are too generous, introducing a leniency program increases the maximal cartel price instead of reducing it. In order to avoid this type of adverse effects, fine reductions should be limited such that $\alpha(p^A, R) > 1 - (1 - \delta) / \lambda(p^A)$. So, moderate fine reductions in case of multiple reporting firms not only allow to avoid the systematic collusion and reporting strategy profile of Section 3.3, but also prevents that introduction of leniency will increase the maximal cartel price.

Next, similar to previous subsections, we characterize the maximal cartel price p^L as a function of $\delta \in (0, 1)$. Consider the cartel price $p \in (p^N, p^M]$, then for $\delta \geq \Delta(p) \in (0, 1)$ this cartel price can be sustained by a strategy profile in which the cartel operates silently, and for $\delta \geq \Phi(p) \in (0, 1)$ the same cartel price can be sustained by a strategy profile in which the cartel systematically colludes and reports. Hence, the lowest threshold on δ suffices to sustain p . Formally, for any cartel price $p \in (p^N, p^M]$ and any $\delta \geq \min\{\Delta(p), \Phi(p)\} \in (0, 1)$ there exists at least one strategy profile supporting p . Inverting this minimum function $\min\{\Delta(p), \Phi(p)\}$ yields $p^L = \max\{\Delta^{-1}(\delta), \Phi^{-1}(\delta)\}$. We have the following result.

Proposition 15 Under antitrust regulation with a leniency program, the maximal cartel price p^L is non-decreasing in $\delta \in (0, 1)$, and

$$\begin{aligned} p^L &= p^N, & \text{if } \delta \in (0, \min\{\Delta(p^N), \Phi(p^N)\}), \\ p^L &\in [p^N, p^M), & \text{if } \delta \in [\min\{\Delta(p^N), \Phi(p^N)\}, \min\{\Delta(p^M), \Phi(p^M)\}), \\ p^L &= p^M, & \text{if } \delta \in [\min\{\Delta(p^M), \Phi(p^M)\}, 1), \end{aligned}$$

where $\min\{\Delta(p^M), \Phi(p^M)\} < 1$.

To conclude this section, we classify all possible leniency programs applicable for the case of multiple reporting $\alpha(p, R)$ in the (p, α) -space. In this classification, two issues stand out: By what strategy profile can we sustain the cartel price p (if sustainable) and whether the program is exploitable. The first question is answered for a silent strategy profile if $p \leq p^A$ and a systematic collusion and reporting strategy profile if $\alpha(p, R) \leq 1 - \frac{1-\delta}{\lambda(p)}$, i.e. region A in Figure 5. The second question requires the following two results.

Lemma 16 The leniency program is exploitable at $p \in [p^N, p^M]$ if and only if $\alpha(p, R) < 1 - \frac{1-\delta}{\Lambda(p)}$. Furthermore, $1 - \frac{1-\delta}{\Lambda(p)} \leq 1 - \frac{1-\delta}{\lambda(p)}$ if and only if $p \in [p^N, p^A]$.

Lemma 17 $1 - \frac{1-\delta}{\Lambda(p)} \geq 0$ is non-decreasing in p and increasing in p if either $\beta'(p) > 0$ or $k'(p) > 0$. Furthermore, $1 - \frac{1-\delta}{\Lambda(p)}$ is increasing in γ , $k(p)$ and $\beta(p)$ and decreasing in δ .

Combining all results implies that the (p, α) -space is divided by the vertical line $p = p^A$ and the curves $1 - \frac{1-\delta}{\Lambda(\cdot)}$ and $1 - \frac{1-\delta}{\lambda(\cdot)}$ that intersect at $p = p^A$ and, in case of two partially

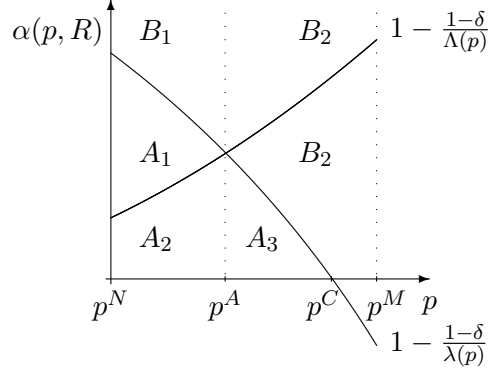


Figure 8: Classification of leniency programs $\alpha(\cdot, R)$.

overlapping horizontal line pieces, intersect at some range of prices $p \leq p^A$. Figure 8 illustrates these curves, and also subdivides the regions A and B in Figure 5 into subregions. We discuss these cases when $p \in (p^N, p^M]$ and $\alpha(p, R)$ belong to each of these regions separately:

A_1 : Since $p \leq p^A$ and $\alpha(p, R) \leq 1 - \frac{1-\delta}{\lambda(p)}$, both strategy profiles sustain p as an equilibrium price. Since also $\alpha(p, R) \geq 1 - \frac{1}{\Lambda(p)}$, the leniency program is not exploitable at p , and the cartel prefers to operate silently.

A_2 : As in A_1 , both strategy profiles sustain p as an equilibrium price. Since also $\alpha(p, R) < 1 - \frac{1}{\Lambda(p)}$, the leniency program is exploitable at p , and the cartel prefers to systematically report.

A_3 : Since $p > p^A$ and $\alpha(p, R) \leq 1 - \frac{1-\delta}{\lambda(p)}$, only the systematic collusion and reporting strategy profile forms an equilibrium. Even though the traditional antitrust policy would eradicate this cartel price, the leniency program annihilates this positive effect by enhancing collusion on this cartel price.

B_1 : Since $p \leq p^A$ and $\alpha(p, R) > 1 - \frac{1-\delta}{\lambda(p)}$, only the strategy profile in which the cartel operates silently forms an equilibrium.

B_2 : Since $p > p^A$ and $\alpha(p, R) > 1 - \frac{1-\delta}{\lambda(p)}$, none of the strategy profiles forms an equilibrium. Hence, the cartel price p cannot be sustained, firms prefer to deviate, and no cartel formation takes place.

These results have the following policy implications. Introducing leniency program $\alpha(p, R)$ in either of the regions A_1 , B_1 or B_2 does not change cartel practices, and the cartel either operates silently to sustain p or does not form at all. However, adverse effects do arise for $\alpha(p, R)$ in regions A_2 and A_3 due to the leniency program. In region A_2 , the leniency program does not eradicate a previously sustainable cartel price p , but it is exploitable and might induce systematic reporting. This can be seen as a first type of adverse effects of (ex-ante) leniency programs. In region A_3 , introduction of the leniency program makes it possible to sustain a previously unsustainable cartel price p implying that leniency programs in this region are harmful. This can be identified as a second type of adverse

effects. For the class of non-decreasing policy functions $\alpha(\cdot, R)$, regions A_2 and A_3 can be avoided by allowing only moderate (or no) fine reductions in case of multiple reporters, i.e. $\alpha(p^A, R) > 1 - (1 - \delta) / \lambda(p^A)$.

The above discussion implies that introduction of wrongly designed leniency programs (e.g. leniency programs that give too generous fine reductions $\alpha(p, R)$ to all reporting firms in case of multiple reporting) can lead to an improvement in cartel stability with higher prices sustained by the cartel. At the same time, under silently operating cartels, in case only one firm reports (or first reporter is clearly identified and is treated differently from the rest), our results of Section 3.2 imply that leniency programs without rewards to a single reporting firm, i.e. $\alpha(p, N)$, do not have any impact on the range of sustainable cartel prices. We will extend our analysis to the discussion of rewards $\alpha(p, N) < 0$ in the next section.

Finally, we discuss a combination of a tougher antitrust policy and substantial fine reductions in the leniency program. According to Figure 8, this reduces $p^S = p^A$ since such a policy change shifts the curve $1 - \frac{1-\delta}{\Lambda(\cdot)}$ upwards, and simultaneously, expands the combined region of A_2 and A_3 enhancing adverse effects that might increase p^R . These opposite effects have an ambiguous effect on the maximal cartel price. The largest reduction of the maximal cartel price is achieved if the new fine reductions are moderate for multiple self-reporting firms, i.e. the new curve $\alpha(\cdot, R)$ is confined to the new regions A_1 , B_1 and B_2 . Otherwise, the reduction in this price is less (and could even be an increase), and the model predicts $p^L = p^R > p^A > p^N$ supported by systematic collusion and reporting in the new situation. Reporting should therefore not be taken as evidence that the cartel dissolves.

4 Discussion

In this section, we discuss the consequences of relaxing some of the assumptions imposed in Section 2. These concern rewards for a single reporting firm, profit maximization by the cartel, limited access to leniency for repeated offenders and punishing a price-deviating firm. Each topic is discussed in a separate subsection.

4.1 Rewards for Single Reporting Firms

Spagnolo (2004) and Rey (2003) proposed an alternative point of view that in order to fully eradicate cartels, leniency programs should offer sufficiently large rewards to reporting firms. Our analysis can be modified to characterize the effective reward needed to upset cartel prices sustainable under traditional antitrust regulation. We confirm the insights in Spagnolo (2004) and Rey (2003) and extend it to our general setting.

Consider a silent cartel, the leniency program is effective to upset the cartel price $p \in (p^N, p^A]$, i.e. $\Lambda(p) \leq \lambda(p)$, if and only if (9) fails at $p \in (p^N, p^A]$. Accordingly,

$$\frac{1}{1 - \alpha(p, N)} < \Lambda(p) \leq \lambda(p) \iff \alpha(p, N) < 1 - \frac{1}{\Lambda(p)} \leq 1 - \frac{1}{\lambda(p)} \leq 0. \quad (16)$$

The smallest upper bound on $\alpha(p, N)$ in (16) has several important implications. First, independent of the policy functions $\beta(p)$ and $k(p)$ and the sector characteristics $\lambda(p)$, δ and γ , the leniency program can only be effective if the first reporting firm is substantially rewarded as $\alpha(p, N) < 0$. This result is quite intuitive. Since a firm's future expected cartel profit is positive (i.e. higher than its competitive profits), this firm can only be induced to report if it will be compensated for foregoing its positive future expected profit. Second, the upper bound $1 - \frac{1}{\lambda(p)}$ is the minimal effective reward to break the silence at cartel price p , which is non-decreasing in p . After substituting (3) into (16), we obtain

$$\alpha(p, N) < \frac{\beta(p) k(p) - \delta [1 - \gamma \beta(p)]}{1 - \delta [1 - \gamma \beta(p)]}. \quad (17)$$

In Rey (2003), it is shown that such reward is unbounded if δ goes to 1. Taking this limit evaluated at arbitrary $p \in (p^N, p^M]$ shows that this limit is bounded for all $\gamma > 0$ and that it is only unbounded for the boundary case $\gamma = 0$, which is the case analyzed in Rey (2003).

The fact that cartels that set higher prices arouse more suspicions makes them more vulnerable to be uncovered. This implies a positive effect on the expected penalty and a negative effect on the expected time of enjoying the benefits from the cartel before the first conviction. Lower expected net benefits makes coming forward less costly and, require lower minimal effective rewards. So, the presence of suspicions has a dampening effect on the minimal effective reward. Cartels that set low prices, and by doing so arouse less suspicions, are the most costly to eradicate. We now summarize our main findings:

Proposition 18 *There exists a non-decreasing (ex-ante) leniency program with rewards for single reporting firms that satisfies (17) for all $p \in (p^N, p^A]$. This program eradicates all cartel prices in $(p^N, p^A]$, which could otherwise be sustained under traditional antitrust enforcement.*

The leniency program in Proposition 18 prevents cartel formation even when the expected fines are too low and not sufficient to prevent cartel formation, i.e. $\beta(p)k(p) < 1$. This reward scheme should only be allowed for single reporting firms, or for the strictly-first reporter.

Based on this discussion and the discussion in Section 3.4, we stress that the issue of single versus multiple reporting firms (or issue of different treatments for the first reporter and the other reporting firms) is particularly important when discussing the effects of leniency programs. Moreover, our results provide partial support for the leniency programs present in OECD countries, where leniency programs allow for large differences between the amount of fine reduction available for the first reporting firm (up to full amnesty) and much smaller or no reductions for further reporters.

Proposition 18, however, should not be understood as advocating leniency programs with rewards, since that would require a welfare analysis and also neglects the issue of exploitable leniency programs. The latter is important, because individual reduced fines for multiple reporting firms, i.e. $\alpha(p, R)$, may depend upon the reward for the strictly first reporting firm, i.e. $\alpha(p, N)$. In case the reward is too generous, i.e. too negative, it may push $\alpha(p, R)$ below

the condition characterizing region B in Figure 5. For example, Spagnolo (2004) considers $\alpha(p, R) = \frac{1}{n}\alpha(p, N) + \frac{n-1}{n}k(p)$ and combined with region B implies $\alpha(p, N) > n(1 - \frac{1-\delta}{\lambda(p)}) - (n-1)k(p)$, which is non-increasing in p . The minimal effective reward reduces the maximal cartel price and avoids adverse effects if and only if $p > p^{ER}$, where $p^{ER} \in [p^N, p^M]$ is the largest intersection point such that $n(1 - \frac{1-\delta}{\lambda(p)}) - (n-1)k(p) = 1 - \frac{1}{\lambda(p)}$.¹¹ Without adding more structure, we cannot answer whether $p^{ER} < p^A$. In general, the moderate individual fine reductions in case of multiple reporting firms allow for some slack in rewarding single reporting firms that might be used to implement such rewards without adverse effects. Currently, the ex-ante leniency programs in some EC countries have a reduced fine for the second reporter and the US system does not. Therefore, the EC system has less slack than the US system in implementing minimal effective rewards.

4.2 Profit-Maximizing Cartels

The maximal cartel price for the entire class of grim-trigger strategy profiles may differ from the standard profit-maximizing cartel price for the same class of strategy profiles. We argue, however, that our approach is complementary to profit maximization, and is also convenient in circumventing technicalities involved with profit maximization.

A profit maximizing cartel always chooses the strategy profile supporting cartel price $p \in (p^N, p^M]$ that yields the highest profit in case multiple profiles support such price. So, the cartel's profit function is equal to

$$\max \left\{ V(p), \frac{1 - \alpha(p, R)}{1 - \delta} \pi(p) \right\}, \quad (18)$$

which is continuous on $p \in (p^N, p^M]$. Substitution of (18) for the objective function p in Program (15) yields profit maximization, and this implies a different objective under the same equilibrium conditions. Therefore, the profit-maximizing cartel price is at most p^L . In other words, the socially worst outcome p^L puts an upper bound on welfare losses to consumers due to, for instance, profit-maximizing behavior.

In case (18) is monotonically non-decreasing, the profit-maximizing cartel price is the boundary solution p^L of Section 3. Since $\pi(p)$ is non-decreasing in $p \in (p^N, p^M]$, this applies to the benchmark case, and to most of the literature where all policy functions are assumed to be constant. The profit function (18), however, may fail monotonicity of either $V(p)$ and $\frac{1-\alpha(p,R)}{1-\delta}\pi(p)$. The product form of $V(p)$ might also destroy concavity, as Example 19 below illustrates. So, maximization of $V(p)$ would also require investigation of second-order conditions. Similar considerations about concavity apply for $\frac{1-\alpha(p,R)}{1-\delta}\pi(p)$.

Since the maximal cartel price does not involve the above technical difficulties it is therefore a more convenient approach. Moreover, whenever the profit-maximizing price is a boundary solution, it coincides with p^L . So, the maximal cartel price and the profit-maximizing cartel price are complementary to each other. We provide with an illustrating example.

¹¹Similar as before, p^{ER} is the largest intersection point of a non-increasing and a non-decreasing curve with the same caveats of Section 3.1 about existence.

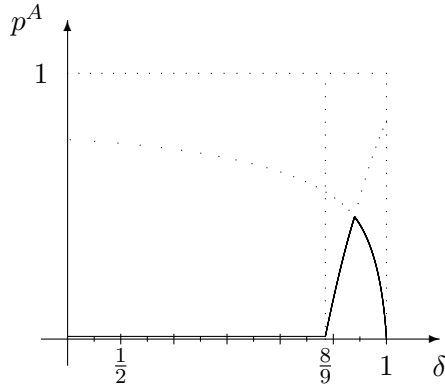


Figure 9: The profit-maximizing cartel price under traditional antitrust regulation.

Example 19 *In this example, we consider profit maximization under traditional antitrust policy in the homogeneous Bertrand oligopoly of Example 5. By (1), $V(p) = \frac{1-\beta kp}{1-\delta(1-\gamma\beta p)} \cdot \frac{(2-p)p}{n}$ fails both monotonicity and concavity on $[p^N, p^M]$.¹² The function is single peaked. For parameter values $k = 3$, $\beta = \frac{1}{6}$ and $\gamma = \frac{2}{3}$, neglecting the equilibrium conditions as in Harrington (2005), MAPLE obtains from $\partial V(p)/\partial p = 0$ the profit-maximizing cartel price*

$$\frac{-27(1-\delta) + 3\sqrt{65\delta^2 - 146\delta + 81}}{4\delta} \in [0, \frac{3}{4})$$

for all $\delta \in [0, 1]$. The profit-maximizing cartel price is decreasing in $\delta \in [0, 1]$. Introducing the equilibrium conditions implies that the profit-maximizing cartel price is the minimum of the above expression and p^A , i.e. the solid curve in Figure 9. These two price curves intersect at $\delta \approx .955$. So, on the interval $[0, 0.955]$, the profit-maximizing cartel price coincides with p^A , while for the interval $(.955, 1]$ the profit-maximizing cartel price is lower than p^A because of non-binding equilibrium conditions.

This example confirms the assertion in Harrington (2004 and 2005) that the equilibrium conditions are always non-binding for sufficiently large $\delta < 1$. Our results, however, also show that the profit-maximizing cartel price is non-monotone in δ on $[0, 1]$ if the equilibrium conditions are taken into account. As δ goes to 1, the profit-maximizing cartel price goes to the competitive price p^N , and the limit profit-maximizing cartel price seriously underestimates the potential maximal damage to consumers under sustainable cartel behavior.

4.3 Limited Access to Leniency

Many OECD countries do not offer unlimited access to their leniency programs. Restricting availability of leniency to repeated offenders prevents cartels from systematically colluding and reporting, see Chen and Rey (2007). In this subsection, we analyze the impact of limited access to leniency when the cartel would otherwise exploit leniency by systematically

¹²Without going into details, for δ close to 1 it holds that $\partial V(p)/\partial p > 0$ for p close to 0, while $\partial V(p)/\partial p < 0$ for p close to 1. For parameter values $k = 3$, $\beta = \frac{1}{6}$, $\gamma = \frac{2}{3}$ and $\delta = .99$, software package MAPLE returns the inflexion point $p = 0.64433 \in [p^N, p^M]$ when solving $\partial^2 V(p)/\partial p^2 = 0$.

colluding and reporting, i.e. when $p^A < p^R < p^M$. In particular, we examine when “collude and report as long as leniency applies and then collude silently” is effective to support the maximal cartel price. We show that the leniency program that blocks repeated offenders and only allows first-time offenders to apply, which is the program considered in Chen and Rey (2007), is optimal within the class of leniency programs with limited access.

Consider the situation where cartel members may only apply for leniency a finite number of times, where this number represents the maximum number of convictions and each conviction reduces the number left for all firms by one. After the leniency program is exhausted, the continuation game is equivalent to that of Section 3.1 where no leniency is offered. Based on our analysis in Section 3.1 and 3.2, the maximal cartel price is p^A and it is sustained by a “collude and never report” strategy profile. From Section 3.1 it follows that $p^A < p^R < p^M$ implies

$$V(p^A) = \frac{1 - \beta(p) k(p^A)}{1 - \delta [1 - \gamma \beta(p^A)]} \pi(p^A) = \pi^{opt}(p^A).$$

In a period where the firms may still apply for leniency one more time, the cartel has two options, either always collude and never report or collude and report in the current period followed by colluding silently. The first option enables them to sustain any cartel price $p \leq p^A$ in every period as in Section 3.2. In the second option, each firm’s profit is $[1 - \alpha(p, R)] \pi(p) + \delta V(p^A)$, where $p \in [p^N, p^M]$ is the cartel price in the current period. Price $p \in [p^N, p^M]$ can be supported if and only if $\pi^{opt}(p) \leq [1 - \alpha(p, R)] \pi(p) + \delta V(p^A)$, which can be rewritten as

$$\frac{1 - \alpha(p, R)}{1 - \delta} \pi(p) - \pi^{opt}(p) \geq \frac{\delta}{1 - \delta} [\pi^{opt}(p) - V(p^A)]. \quad (19)$$

When $p^A < p^R < p^M$, (11) evaluated at $p = p^A$ implies that the left-hand side of (19) is non-negative. Then also, $\pi^{opt}(p^A) = V(p^A)$ implies that the right-hand side of (19) is equal to 0 at $p = p^A$. Therefore, (19) holds at $p = p^A$, and hence we must have $p^1 \geq p^A$, where p^1 is the maximal cartel price where (19) holds. When $p^A < p^R < p^M$, however, $p^1 = p^A$ is impossible. To see this, note that the leniency program is exploitable at price $p = p^A$, and, by its definition, we have

$$\pi^{opt}(p^A) = V(p^A) < (1 - \delta) \left[\frac{1 - \alpha(p^A, R)}{1 - \delta} \pi(p^A) \right] + \delta V(p^A).$$

This implies that (19) has a slack at $p = p^A$, and by continuity in p , this condition still holds for all p slightly above p^A . So, we conclude that $p^1 > p^A$. Since $\pi^{opt}(p)$ is increasing, the right-hand side of (19) is positive whenever $p > p^A$, and therefore, (19) is at least as restrictive as (11) for all $p \in (p^A, p^M]$. When $p^A < p^R < p^M$, then p^R satisfies (11) with an equal sign, and necessarily, also p^1 satisfies (19) with an equality, i.e.

$$\frac{1 - \alpha(p^1, R)}{1 - \delta} \pi(p^1) - \pi^{opt}(p^1) = \frac{\delta}{1 - \delta} [\pi^{opt}(p^1) - V(p^A)] > 0.$$

From this we conclude that $p^1 < p^R$. To summarize, whenever the unlimited leniency program predicts adverse effects, i.e. $p^A < p^R < p^M$, limiting access to leniency to just one time reduces the maximal cartel price to $p^1 \in (p^A, p^R)$.

In a period where the firms may still apply for leniency for two more times, the firms have three options, the two options discussed before or collude and exhaust the leniency program within two periods followed by always collude and never report. The first two options do not exhaust the leniency program, but nevertheless enable the firms to sustain either silent collusion on p^A as in Section 3.2, or collude and report on $p^1 > p^A$ in the current period followed by silent collusion on p^A as in Section 3.2. Then in the third option, setting cartel price $p \in [p^N, p^M]$ in the current period followed by collude and report on p^1 in the next period and then p^A as in Section 3.1, each firm's profit is $[1 - \alpha(p, R)] \pi(p) + \delta \bar{V}(p^1, p^A)$, where $\bar{V}(p^1, p^A) > V(p^A)$ denotes $[1 - \alpha(p^1)] \pi(p^1) + \delta V(p^A)$.¹³ Price $p \in [p^N, p^M]$ can be supported in the current period if and only if $\pi^{opt}(p) \leq [1 - \alpha(p, R)] \pi(p) + \delta \bar{V}(p^1, p^A)$, and similar as before, implies

$$\frac{1 - \alpha(p, R)}{1 - \delta} \pi(p) - \pi^{opt}(p) \geq \frac{\delta}{1 - \delta} [\pi^{opt}(p) - \bar{V}(p^1, p^A)]. \quad (20)$$

Condition (19) evaluated at $p = p^1$ implies that the left-hand side of (19), which is also the left-hand side of (20), is at least $\frac{\delta}{1 - \delta} [\pi^{opt}(p^1) - V(p^A)]$ and this expression is larger than the right-hand side of (20) evaluated at $p = p^1$. Therefore, $p = p^1$ can be sustained, and hence we must have $p^2 \geq p^1$, where p^2 is the maximal cartel price where (20) holds. Moreover, (20) has a slack at $p = p^1$, and by continuity in p , this condition holds for all p slightly above p^1 . So, we can conclude that $p^2 > p^1 > p^A$. Similar as for p^1 , when $p^A < p^R < p^M$ the maximal cartel price p^2 is found at

$$\frac{1 - \alpha(p^1, R)}{1 - \delta} \pi(p^2) - \pi^{opt}(p^2) = \frac{\delta}{1 - \delta} [\pi^{opt}(p^2) - \bar{V}(p^1, p^A)] > 0.$$

This implies $p^2 \in (p^1, p^R)$, and

$$[1 - \alpha(p^2)] \pi(p^2) + \delta \bar{V}(p^1, p^A) = \pi^{opt}(p^2) > \pi^{opt}(p^1) = \bar{V}(p^1, p^A).$$

To summarize, whenever the unlimited leniency program predicts adverse effects, i.e. $p^A < p^R < p^M$, limiting access to leniency to two times will yield the sequence of decreasing maximal cartel prices $p^2 > p^1 > p^A$ where $p^2 \in (p^A, p^R)$.

The arguments for p^1 and p^2 can be repeated up to any finite number representing the maximum of times a cartel may apply for leniency. In principle, this only requires to redefine the value function on the right-hand side of (20). Doing so yields a finite sequence of decreasing maximal cartel prices above p^A followed by silent collusion. In case the law dictates

¹³By $\pi^{opt}(p)$ is increasing and the equilibrium conditions for p^1 and p^A , we have

$$\bar{V}(p^1, p^A) = [1 - \alpha(p^1)] \pi(p^1) + \delta V(p^A) \geq \pi^{opt}(p^1) > \pi^{opt}(p^A) = V(p^A).$$

the presence of a leniency program, the consumers' best of these sequences of maximal cartel prices is the shortest: meaning p^1 is sustained once, followed by silent collusion.

These results offer positive support for leniency programs in those OECD countries that allow for fine reductions to first-time offenders only. Limiting access to leniency for repeated offenders has two positive effects: First, it minimizes the length of time firms can exploit adverse effects in the form of "collude and report" strategies, and second, it minimizes the maximal cartel price sustainable under adverse effects. These results show that the exogenously given leniency program offered to first-time offenders only, as in e.g. Chen and Rey (2007), is optimal. In that respect, our interpretation of the effects of limited access to leniency as positive differs from the negative conclusion in Chen and Rey (2007), who measure the effectiveness of restricted access solely by its ability to eradicate cartels.

4.4 Prosecution of Price-Deviating Firms

Spagnolo (2004) and Rey (2003) argue that antitrust enforcement would be most effective if the incentives for price deviations are maximal and this requires that price-deviating firms are never prosecuted, as analyzed in Section 3. In this section, we assume price-deviating firms can be prosecuted and extend our analysis of Section 3.

Consider traditional antitrust regulation first. Let (p_i, p) be a history in which firm i deviated in prices in the current period, $\beta(p_i, p) \geq 0$ be the cartel's detection probability after firm i 's unilateral deviation, and $k(p_i, p) \geq 0$ be firm i 's fine. For $p_i = p$ we impose $\beta(p, p) = \beta(p)$ and $k(p, p) = k(p)$. Firm i 's expected profit from price deviation p_i is equal to $\pi(p_i, p) [1 - \beta(p_i, p) k(p_i, p)]$, and firm i 's optimal price deviation solves

$$\max_{p_i \in [p^N, p^M]} \pi(p_i, p) [1 - \beta(p_i, p) k(p_i, p)] \leq \max_{p_i \in [p^N, p^M]} \pi(p_i, p) = \pi^{opt}(p). \quad (21)$$

Define $\lambda_\beta(p)$ as the relative size of the net cartel profit to the net gains under the best unilateral deviation, i.e.

$$\lambda_\beta(p) = \frac{\pi(p)}{\max_{p_i \in [p^N, p^M]} \pi(p_i, p) [1 - \beta(p_i, p) k(p_i, p)]} \geq \lambda(p), \quad (22)$$

which is continuous in $p \in (p^N, p^M]$. For explanatory reasons, we assume $\lambda_\beta(\cdot)$ is non-increasing. More importantly, when $\beta(p_i, p) k(p_i, p) > 0$ for all $p \in (p^N, p^M]$, the main difference with Section 3.1 is that $\lambda_\beta(\cdot)$ can be regarded as an upward shift of $\lambda(\cdot)$ in Figure 1. Then, replicating the analysis of Section 3.1 implies condition (3) becomes $\lambda_\beta(p) \geq \Lambda(p)$, and it sustains a non-empty subinterval of cartel prices $[p^N, p_\beta^A]$, where $p_\beta^A \geq p^A$ is the largest intersection point of $\lambda_\beta(p) \geq \Lambda(p)$. Obviously, $\beta(p_i, p) k(p_i, p) = 0$ for all $p_i, p \in (p^N, p^M]$ and $p_i \neq p$, suffices to obtain $p_\beta^A = p^A$. In words, the lowest maximal cartel price sustainable under traditional antitrust regulation is p^A and it can be attained by granting single price-deviating firms full immunity, i.e. $\beta(p_i, p) k(p_i, p) = 0$. This confirms the insights of Spagnolo (2004) and Rey (2003) and extends these to our general setting.

Next, consider a leniency program for a single reporting firm under the class of silent strategy profiles of Section 3.2 that also specify that firms remain silent even after a price deviation (p_i, p) .^{14,15} For any history (p_i, p) , $\alpha(p_i, p, N) \geq 0$ denotes price-deviating firm i 's reduced fine after reporting with $\alpha(p, p, N) = \alpha(p, N)$. Firm i 's expected profit from the price deviation p_i is equal to $\pi(p_i, p) [1 - \alpha(p_i, p, N)]$, and firm i 's optimal price deviation solves $\max_{p_i \in [p^N, p^M]} \pi(p_i, p) [1 - \alpha(p_i, p, N)] \leq \pi^{opt}(p)$. Define $\lambda_\alpha(p)$ as the relative size of the net cartel profit to the net gains under the best unilateral deviation, i.e.

$$\lambda_\alpha(p) = \frac{\pi(p)}{\max_{p_i \in [p^N, p^M]} \pi(p_i, p) [1 - \alpha(p_i, p, N)]} \geq \lambda(p). \quad (23)$$

Duplication of the arguments for $\lambda_\beta(\cdot)$ yields that $\lambda_\alpha(\cdot)$ is an upward shift of $\lambda(\cdot)$ in Figure 1 with intersection point $p_\alpha^S \geq p^A$ and $[p^N, p_\alpha^S]$ as the range of sustainable cartel prices. Moreover, setting $\alpha(p_i, p) = 0$ for all $p_i, p \in (p^N, p^M]$ and $p_i \neq p$ suffices to obtain $p_\alpha^S = p^A$.

The implication for silently operating cartels is that Program (10) needs to be modified such that $\lambda(p) \geq \Lambda(p)$ is replaced by both $\lambda_\beta(p) \geq \Lambda(p)$ and $\lambda_\alpha(p) \geq \Lambda(p)$. Hence, $p^S = \min\{p_\beta^A, p_\alpha^S, \tilde{p}\} \geq p^A$. Whenever $0 \leq \alpha(p_i, p, N) < \beta(p_i, p) k(p_i, p)$ for all $p_i, p \in (p^N, p^M]$ and $p_i \neq p$, the curve $\lambda_\alpha(\cdot)$ lies between $\lambda(\cdot)$ and $\lambda_\beta(\cdot)$ and we obtain $p^A \leq p_\alpha^S < p_\beta^A$. This implies that if the leniency program offers sufficiently generous fine reductions to a price-deviating firm in case it is the single (or strictly the first) reporting firm, then such firm is induced to seek protection from fines. So, introducing leniency helps in reducing the maximal cartel price from p_β^A to p_α^S . However, no (ex-ante) leniency program without rewards can ever obtain a maximal cartel price below p^A .

Most importantly, in case the traditional antitrust policy does not minimize the maximal cartel price, i.e. $\beta(p_i, p) k(p_i, p) > 0$ for all $p_i, p \in (p^N, p^M]$, the introduction of the ex-ante leniency program that grants a single price-deviating firm that reports full amnesty, i.e. $\alpha(p_i, p, N) = 0$ for all $p_i, p \in (p^N, p^M]$, is an improvement, since it allows to attain the lowest maximal cartel price under *silent* strategy profiles, i.e. $p^S = p_\alpha^S = p^A$. This again confirms the insights of Spagnolo (2004) and Rey (2003).

In summary, we confirm and extend two perspectives brought forward by Spagnolo (2004): First, the incentives for price deviations are maximal if the leniency program does not prosecute price-deviating firms, and therefore such firms should be exempted from fines. Second, in case the traditional regulation does prosecute price-deviating firms, then the leniency program might induce such firms to report in order to protect themselves from expected fines, if the latter are high while the reduced fines for the single reporting firm are low.

We forego a detailed analysis of the highest cartel price under systematic collusion and reporting, denoted as p_α^R . In essence, $\alpha(p, R)$ should also be extended to $\alpha(p_i, p, R)$, $\lambda(p)$ has to be modified similar to (23), and then straightforward modification of (11) follows. Since this is also an upward shift of the curve $\lambda(\cdot)$, this yields $p_\alpha^R \geq p^R$. Consequently,

¹⁴We only consider the perspective of a price-deviating firm in order to minimize the discussion.

¹⁵Alternatively, the silent cartel agreement may specify that all firms report after a price deviation, i.e. a race to report. We forego analyzing such strategy profiles, because the analysis and conclusions would be similar after substituting $\alpha(p_i, p, N)$ by i 's reduced fine for this case, which is $\alpha(p_i, p, R)$ as defined below.

modifying the analysis in Section 3.4 would yield $p^L = \max \{p^S, p_\alpha^R\}$ and, therefore, $p^L \geq \min \{p_\beta^A, p_\alpha^S\} \geq p^A$, which demonstrates that the negative result of Section 3.4 persists.

These results provide important support for the leniency programs present in OECD countries, where price-deviating firms are prosecuted and the single reporting firm is granted full immunity: Introducing such leniency program reduces the maximal cartel price set by a silent cartel from p_β^A to p^R . If the design of leniency programs avoids the adverse effects discussed in Section 3.3 and 3.4, then (ex-ante) leniency programs have a non-increasing effect on the number of cartels and reduce the maximal damage to consumers compared to the situation prior to the leniency program. Remaining cartels operate silently. Furthermore, a properly-designed leniency program should allow fine reductions only for the first reporting firm, like in the US. In case leniency programs are wrongly designed (which seems the case if multiple reporting firms can obtain substantial fine reductions, like in some EU countries) cartels will switch from operating silently to systematically colluding and reporting. Then an increase in reporting after introducing a leniency program is a negative phenomenon.

5 Concluding Remarks

In this paper, the maximal cartel price in an infinitely-repeated sequential game with two-stages is studied as a proxy for both the set of sustainable cartel prices, and consumers' worst-case scenario of maximal damage. This allows for the simultaneous analysis of two important decisions firms face in the presence of leniency programs: The decision to deviate from the cartel price, and the decision to report to the AA. We show that the maximal cartel price is the maximum of two other prices: The maximal cartel price sustained by cartels that operate silently and the price sustained by cartels that systematically collude and report. This characterization disentangles the effects of traditional antitrust policies, leniency programs, and behavioral assumptions into three separate effects, which are studied in Sections 3.1, 3.2 and 3.3 before being combined in Section 3.4. Each of these maximal cartel prices, and as a result the effectiveness of antitrust regulation and leniency programs, depends upon the prevailing policies and the sector characteristics, such as the sector structure, the type of competition, the sector's cartel culture, and the discount factor. Furthermore, the effects of leniency in Section 3.2 are positive in nature, because these concern incentives to break the cartel, whereas the effects of leniency in Section 3.3 are adverse since the maximal cartel price increases and systematic reporting becomes a part of the collusive cartel strategy.

We provide policy recommendations on how to improve the design of antitrust policy and ex-ante leniency, how to eliminate adverse effects, and what is necessary to prevent cartel formation in the first place: 1) Price-deviating firms should not be prosecuted in order to give firms maximal incentives for price deviations. 2) The design of leniency programs should avoid adverse effects since these raise the maximal cartel price, and therefore increase the maximal damage to consumers. 3) Individual fine reductions in case of multiple reporting firms should be moderate in order to obey the necessary and sufficient conditions under

which adverse effects are eliminated. Based on Section 3.4, it seems that the current ex-ante leniency programs in most EC countries could be improved by abolishing the reduced fine for the second reporting firm, similar to the current US system. The analysis in Section 4.3 implies that leniency should be offered to first-time offenders only. Our results offer positive support for the policies currently employed in OECD countries that prosecute price-deviating firms while simultaneously granting full immunity to single reporting firms, and policies that restrict leniency for repeated offenders to first-time offenders only. Whether these OECD leniency programs also avoid adverse effects is left as an empirical question.

Effective ex-ante leniency programs that eradicate all cartel prices always exist in theory. Such leniency programs, however, necessarily involve rewards to the single reporting firm (or strictly the first reporting firm). Such rewards might introduce adverse effects. The conditions that avoid adverse effects seem to allow for some slack in rewarding single reporting firms, but we leave it to future research in how far the minimum effective reward can be implemented without adverse effects to eradicate the most damaging prices.

We stress that our results are robust. First of all, our results hold for general oligopoly models and general policy functions. Furthermore, since there is a substantial class of oligopoly models in which the profit-maximizing cartel price coincides with the maximal cartel price, our results complement the cartel profit-maximization approach for this class. Most important is the robustness with respect to relaxing the assumption never to prosecute price-deviating firms, which is discussed above.

Our focus is on general policy functions for a methodological reason. In future research, the optimal design of traditional antitrust and leniency programs remains to be an important research issue. Studying the optimal design requires a well-defined framework for analyzing the effects of changes in antitrust policies and leniency programs on consumers' welfare. Such changes can be thought of as shaping the policy functions and, ideally, one would like a flexible and large class of such policy functions that are a priori neither constant or linear. Our framework allows for such a rich class of potential policy functions and a characterization of the maximal cartel price related to the equilibrium conditions. Future research will verify whether the currently popular fixed detection probabilities and fixed fines are indeed the optimal policies within the class of policy functions we consider.

As mentioned in the introduction, Spagnolo (2008) questions whether the observed increases in cartel detection are the result of unobserved increases in cartel activity or are the result of improved effectiveness of leniency programs. Based on our results we are able to address this issue. For wrongly designed leniency programs introducing such program implies that cartels switch from operating silently to systematically colluding and reporting, and also the maximal cartel price increases. Then, the observed increase in reporting in many OECD countries should be viewed as a negative phenomenon. If the leniency program is properly designed, i.e. it avoids adverse effects and grants full immunity to single reporting firms, then our model cannot explain the observed increase in reporting, but it excludes an unobserved increase in cartel activity as a possible explanation. Moreover, if such properly

designed leniency program is accompanied by tougher traditional antitrust regulation, than the maximal cartel price will be reduced due to such policy change.

6 Appendix with Proofs

Proof of Lemma 1. Taking derivative of $\Lambda(p)$ with respect to p , we have

$$\frac{d\Lambda(p)}{dp} = \frac{\delta\gamma\beta'(p)[1 - \beta(p)k(p)] + [\beta'(p)k(p) + \beta(p)k'(p)][1 - \delta(1 - \gamma\beta(p))]}{[1 - \beta(p)k(p)]^2} \geq 0,$$

due to the assumptions on $\beta(p)$ and $k(p)$. In addition, when either $\beta'(p) > 0$ or $k'(p) > 0$, the above inequality is strict. Next, $\hat{\beta}(p) > \beta(p)$ implies

$$[(1 - \delta)k(p) + \gamma\delta][\hat{\beta}(p) - \beta(p)] > 0 \iff \frac{1 - \delta[1 - \gamma\hat{\beta}(p)]}{1 - \hat{\beta}(p)k(p)} > \frac{1 - \delta[1 - \gamma\beta(p)]}{1 - \beta(p)k(p)}.$$

The effects of γ , δ and $k(p)$ are obvious. \square

Proof of Lemma 2. Since $\lambda(p)$, $\beta(p)$, and $k(p)$ are all continuous in $p \in (p^N, p^M]$, so is $\Delta(p)$. Taking the derivative of $\Delta(p)$ with respect to p , we have

$$\frac{d\Delta(p)}{dp} = \frac{-\lambda'(p)[1 - \beta(p)k(p)] + \lambda(p)[\beta'(p)k(p) + \beta(p)k'(p)]}{1 - \gamma\beta(p)} + \Delta(p)\frac{\gamma\beta'(p)}{1 - \gamma\beta(p)} \geq 0$$

due to $\lambda'(p) \leq 0$, $\beta'(p) \geq 0$ and $k'(p) \geq 0$. Note also that

$$\Delta(p) \equiv \frac{1 - \lambda(p)[1 - \beta(p)k(p)]}{1 - \gamma\beta(p)} \geq 1 - \lambda(p)[1 - \beta(p)k(p)] \geq 1 - \lambda(p).$$

The last inequality is strict when $\beta(p) > 0$. Next, $\hat{\beta}(p) > \beta(p)$ implies

$$[\gamma(1 - \lambda(p)) + \lambda(p)][\hat{\beta}(p) - \beta(p)] > 0 \iff \frac{1 - \lambda(p)[1 - \hat{\beta}(p)k(p)]}{1 - \gamma\hat{\beta}(p)} > \frac{1 - \lambda(p)[1 - \beta(p)k(p)]}{1 - \gamma\beta(p)}$$

The effect of $k(p)$ is obvious. \square

Proof of Proposition 3. Lemma 1 and monotonicity of $\lambda(p)$ imply that p^A is non-decreasing in δ and non-increasing in γ . Note that $\Delta(p^N) = 0$ and $\lim_{\varepsilon \rightarrow 0+} \Delta(p^N + \varepsilon) = 1 - \underline{\lambda}$. When $\delta \in (0, 1 - \underline{\lambda})$, (3) fails for all $p \in (p^N, p^M]$ and, hence, $p^A = p^N$. For $\Delta(p^M) \leq \delta < 1$, (3) holds at $p = p^M$ and, hence, $p^A = p^M$. Otherwise, we have $p^A \in [p^N, p^M)$. \square

Proof of Lemma 9. Since $\lambda(p)$, and $\alpha(p, R)$ are continuous in $p \in (p^N, p^M]$, so is $\Phi(p)$. Next, $\alpha(p, R) \geq 0$ implies $\Phi(p) \equiv 1 - \lambda(p)[1 - \alpha(p, R)] \geq 1 - \lambda(p)$.

The last inequality is strict when $\alpha(p, R) > 0$. Taking the derivative of $\Phi(p)$ with respect to p , we have $\frac{d\Phi(p)}{dp} = -\lambda'(p)[1 - \alpha(p, R)] + \lambda(p)\alpha'(p, 1) \geq 0$, due to $\lambda'(p) \leq 0$, $\alpha(p, R) \geq 0$, and $\alpha'(p, 1) \geq 0$. Finally, the effect of $\alpha(p, R)$ is obvious. \square

Proof of Proposition 10. Observe that $\frac{1 - \delta}{1 - \alpha(\cdot, 1)}$ is decreasing in $\delta \in (0, 1)$ and increasing δ will relax (11). Hence, p^R is non-decreasing in δ . Substitution of $\Phi(p)$ for $\Delta(p)$ in the proof of

Proposition 3 yields the results for each interval of δ . Finally, $\lambda(p) < 1$ and $0 \leq \alpha(p, R) < 1$ for all $p \in (p^N, p^M]$ imply $\Phi(p) = 1 - \lambda(p)[1 - \alpha(p, R)] < 1$. So $\Phi(p^M) < 1$. \square

Proof of Lemma 11. Note that $1 - \frac{1-\delta}{\lambda(p)}$ responds similar to p as $\lambda(p)$, and the latter is non-increasing or increasing in p . Next, since $\underline{\lambda} \leq 1$ we have that $\lim_{\varepsilon \rightarrow 0} 1 - \frac{1-\delta}{\lambda(p^N + \varepsilon)} = 1 - \frac{1-\delta}{\underline{\lambda}} \leq \delta$. Finally, $1 - \frac{1-\delta}{\lambda(p)} \geq 0 \iff \delta \geq 1 - \lambda(p)$, which is the constraint of Program (5) $\iff p \in [p^N, p^C]$. \square

Proof of Proposition 14. Note that $p^L > p^A$ if and only if $p^R > p^A$ if and only if either one of the following two distinct cases holds:

1. The curve $\Lambda(\cdot)$ at $p = p^A$ lies above the curve $\frac{1-\delta}{1-\alpha(\cdot, R)}$ at $p = p^A$ (the case illustrated by Figure 7). This case is equivalent to $\alpha(p^A, R) < 1 - \frac{1-\delta}{\lambda(p^A)}$.
2. $\alpha(p^A, R) \geq 1 - \frac{1-\delta}{\lambda(p^A)}$. Then, we must have $\alpha(p^A, R) = 1 - \frac{1-\delta}{\lambda(p^A)} = 1 - \frac{1-\delta}{\Lambda(p^A)}$, where the latter equality holds by definition of p^A , and $>$ instead of $=$ would imply $p^R < p^A$, which contradicts that we derive conditions for $p^R > p^A$. Next, the continuous and weakly monotonic functions $\lambda(\cdot)$ and $\Lambda(\cdot)$ cannot be both constant on $[p^A, p^A + \varepsilon_1]$ for some $\varepsilon_1 > 0$, because then we obtain the contradiction that $p^A = \max_p \{p : \lambda(p) \geq \Lambda(p)\} \geq p^A + \varepsilon_1 > p^A$. So, $\lambda(p) < \Lambda(p)$ for all $p > p^A$. Furthermore, both the continuous functions $\lambda(\cdot)$ and $\frac{1-\delta}{1-\alpha(\cdot, R)}$ cannot be monotonic (i.e. not constant) on $[p^A, p^A + \varepsilon_2]$ for all $\varepsilon_2 > 0$, because then $\frac{1-\delta}{1-\alpha(p, R)} > \lambda(p)$ for all $p > p^A$ and, hence, $p^R \leq p^A$, which again contradicts that we derive conditions for $p^R > p^A$. So, for some $\varepsilon_2 > 0$ we must have that both $\lambda(\cdot)$ and $\frac{1-\delta}{1-\alpha(\cdot, R)}$ are constant on $[p^A, p^A + \varepsilon_2]$. Since $\lambda(\cdot)$ is constant while we already concluded that $\lambda(\cdot)$ and $\Lambda(\cdot)$ cannot be both constant for $p > p^A$ and, in particular for $p \in [p^A, p^A + \varepsilon_2]$, it follows that for some $\bar{\varepsilon} \in (0, \varepsilon_2)$ the continuous function $\Lambda(\cdot)$ is increasing on $[p^A, p^A + \bar{\varepsilon}]$. \square

Proof of Proposition 15. By Proposition 3 and 10, both $\Delta(p)$ and $\Phi(p)$ are non-decreasing in $\delta \in (0, 1)$ and, therefore, $\min\{\Delta(p), \Phi(p)\}$ is also non-decreasing in δ . Therefore, p^L is non-decreasing in $\delta \in (0, 1)$. Substitution of $\min\{\Delta(p), \Phi(p)\}$ for $\Delta(p)$ in the proof of Proposition 3 yields the results for each interval of δ . Finally, by Proposition 10, $\min\{\Delta(p^M), \Phi(p^M)\} \leq \Phi(p^M) < 1$. \square

Proof of Lemma 16 + 17. All results follow immediately from Lemma 1 after observing that $1 - \frac{1-\delta}{\Lambda}$ responds similar to changes in variables and parameters as Λ . Next, $\frac{1-\alpha(p, R)}{1-\delta} \pi(p) > V(p) = \frac{\pi(p)}{\Lambda(p)}$ if and only if $\alpha(p, R) < 1 - \frac{1-\delta}{\Lambda(p)}$. By definition of p^A , $p \in [p^N, p^A]$ if and only if $\lambda(p) \geq \Lambda(p)$. The latter is equivalent to $1 - \frac{1-\delta}{\Lambda(p)} \leq 1 - \frac{1-\delta}{\lambda(p)}$. Finally, from Section 3.1, $\Lambda(p^N) = 1 - \delta$. Then, $1 - \frac{1-\delta}{\Lambda(p^N)} = 0$. Combined with $1 - \frac{1-\delta}{\Lambda(p)}$ non-decreasing ensures non-negativity. \square

References

Bryant, P. and E. Eckard (1991). The probability of getting caught. *Review of Economics and Statistics* 73, 531–536.

- Buccirossi, P. and G. Spagnolo (2006). Leniency programs and illegal transactions. *Journal of Public Economics* 90(6-7), 1281–1297.
- Chen, J. and J. Harrington (2007). The impact of the corporate leniency program on cartel formation and the cartel price path. In V. Ghosal and J. Stennek (Eds.), *The Political Economy of Antitrust*. Elsevier.
- Chen, Z. and P. Rey (2007). On the design of leniency programs. IDEI Working Papers 452, Institut d'conomie Industrielle (IDEI), Toulouse.
- D.O.J. (1993). Us corporate leniency policy. <http://www.usdoj.gov/atr/public/guidelines/0091.htm>.
- D.O.J. (1998). The twelfth annual report of the department of justice. <http://www.usdoj.gov/atr/public/speeches/1583.html>.
- EC (2006). Commission notice on immunity from fines and reduction of fines in cartel cases. *Official Journal of the European Union* (2006/C 298/11). Brussels, <http://europa.eu.int/comm/competition/antitrust/leniency/>.
- Fudenberg, D. and J. Tirole (1991). *Game Theory*. MIT Press, Cambridge.
- Harrington, J. (2004). Cartel pricing dynamics in the presence of an antitrust authority. *The Rand Journal of Economics* 35, 651–673.
- Harrington, J. (2005). Optimal cartel pricing in the presence of an antitrust authority. *International Economic Review* 46, 145–170.
- Harrington, J. (2008). Optimal corporate leniency programs. *The Journal of Industrial Economics* LVI(2), 215–246.
- Harrington, J. and J. Chen (2006). Cartel pricing dynamics with cost variability and endogenous buyer detection. *The International Journal of Industrial Organization* 24(6), 1185–1212.
- Hinloopen, J. (2003). An economic analysis of leniency programs in antitrust law. *De Economist* 151, 415–432.
- Hinloopen, J. (2006). Internal cartel stability with time-dependent detection probabilities. *The International Journal of Industrial Organization* 24(6), 1213–1229.
- Hinloopen, J. and A. Soetevent (2008). Laboratory evidence on the effectiveness of corporate leniency programs. *RAND Journal of Economics* 39, 607–616.
- Motchenkova, E. (2004). Effects of leniency programs on cartel stability. CentER Discussion Papers Series 2004-98, Tilburg University, Tilburg.
- Motta, M. and M. Polo (2003). Leniency programs and cartel prosecution. *International Journal of Industrial Organization* 21, 347–379.
- OECD (2002). Fighting hard-core cartels: Harm, effective sanctions and leniency programs. OECD Report 2002, OECD, Paris, France, <http://www.SourceOECD.org>.

- Rey, P. (2003). Towards a theory of competition policy. In M. Dewatripont, L. Hansen, and S. Turnovsky (Eds.), *Advances in Economics and Econometrics: Theory and Applications*. Cambridge University Press.
- Spagnolo, G. (2004). Optimal leniency programs. Centre for Economic Policy Research, Discussion paper series, 4840, <http://www.cepr.org/pubs/new-dps/showdp.asp?dpno=4840>, (revised 2008).
- Spagnolo, G. (2008). Leniency and whistleblowers in antitrust. In P. Buccirossi (Ed.), *Handbook of Antitrust Economics*. MIT Press.
- Wen, Q. (2002). A folk theorem for repeated sequential games. *Review of Economic Studies* 69, 493–512.