Contributions

Harold Houba*, Evgenia Motchenkova and Quan Wen

The Effects of Leniency on Cartel Pricing

Abstract: We analyze how leniency affects cartel pricing in an infinitely repeated oligopoly model where the fine rates are linked to illegal gains and detection probabilities depend on the degree of collusion. A novel aspect of this study is that we focus on the worst possible outcome. We investigate the maximal cartel price, the largest price for which the conditions for sustainability hold. We analyze how the maximal cartel price supported by different cartel strategies adjusts in response to the introduction of (ex-ante and ex-post) leniency programs. We disentangle the effects of traditional antitrust enforcement, leniency, and cartel strategies on the maximal cartel price. Ex-ante leniency cannot reduce the maximal cartel price below the price under antitrust without leniency. On the other hand, for ex-post leniency, improvement is possible and granting full immunity to single-reporting firms achieves the largest reduction in the maximal cartel price. To reduce adverse effects under both leniency programs, fine reductions to multiple-reporting firms should be moderate or absent. Finally, ex-post leniency should provide less generous fine reductions to multiple-reporting firms, which is supported by the current practice in the US and the EU.

Keywords: cartel, antitrust, competition policy, leniency program, self-reporting, repeated game
JEL Classification: L41 Monopolization, Horizontal Anti-competitive Practices, K21 Antitrust Law, C72 Non-cooperative Games

DOI 10.1515/bejte-2013-0139

***Corresponding author: Harold Houba,** Department of Econometrics, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands, E-mail: harold.houba@vu.nl
Evgenia Motchenkova, Department of Economics, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands, E-mail: e.i.motchenkova@vu.nl
Quan Wen, Department of Economics, University of Washington, Box 353330, Savory 305, Seattle, WA 98195-3330, USA, E-mail: wenq2@u.washington.edu

1 Introduction

During the last 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 the cartel members who collaborate with the antitrust authority (AA) by revealing information about the cartel. The rationale for having leniency programs is based upon the economic principle that induces firms, who broke the law, to report their illegal activities if they are given proper incentives. Both the US and the EU AAs have experienced an increase in leniency applications after introducing leniency programs and claim such programs are successful, while many academics remain skeptical. The discussion generated a vast theoretical literature on the effects of leniency programs.[1] Many influential studies are conducted in the context of an infinitely repeated sequential game: in each period, the firms first choose whether to collude, and then decide whether to apply for leniency if they have colluded. In most of the existing studies, price setting is restricted to three prices that capture the profit levels of perfect competition, cartel (or monopoly) and the profit of cheating on the cartel price. In response to policy changes such as the introduction of a leniency program, the existing models focus on whether the cartel forms, i.e., can the cartel sustain the cartel (or monopoly) price. Our study focuses on the equally important issue of how cartels adjust their prices in response to such policy changes.

We investigate the impact of leniency programs on cartel prices in a general dynamic oligopoly model with price modeled as a continuous decision variable. In this framework, we analyze the impact of antitrust enforcement instruments, such as (reduced) fine schedules and monitoring and detection probabilities, that are endogenous in the collusive price. For the proper analysis of cartel price adjustments in response to policy changes, one should shift the focus away from the traditional analysis of sustainability of collusion in terms of the critical discount factor. Block, Nold, and Sidak (1981), Harrington (2004, 2005), and Chen and Harrington (2007) provide analyses of the profit-maximizing cartel price when the discount factor is sufficiently large. Alternatively, without restricting the range of discount factors, we focus

1 Optimal implementation of antitrust policy with leniency programs for cartel enforcement has been analyzed in, e.g., Motta and Polo (2003), Rey (2003), Spagnolo (2004, 2008), Harrington (2004, 2005, 2008, 2011), Hinloopen (2003, 2006), Motchenkova (2004), Buccirossi and Spagnolo (2006), Harrington and Chen (2006), Chen and Rey (2013), Chen and Harrington (2007), Choi and Gerlach (2012), and Lefouili and Roux (2012).

on the highest cartel price for which the equilibrium conditions for sustainability hold, called the maximal cartel price.

The maximal cartel price has received little attention in the literature and this is an omission for several reasons. The maximal cartel price represents the consumers' worst-case scenario of maximal damage. Furthermore, it is the relevant proxy for the set of sustainable cartel prices, i.e., the cartel's strategic possibilities to exploit price fixing. Also, it puts a simple upper bound on the profit-maximizing cartel price avoiding technicalities involved with the possible non-monotonicity and non-concavity of the cartel's objective function. In contrast, the objective function underlying the maximal cartel price is always monotonically increasing and linear. Next, in studying the profit-maximizing cartel price, the main focus is on the subcase of nonbinding equilibrium conditions.[2] Because the maximal cartel price can be reinterpreted as the profit-maximizing cartel price under binding equilibrium conditions, our study is complementary to these studies. Finally, since AAs often do not have detailed information about demand, costs, and profits of firms, their 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 criterion.

We introduce a novel technique for analyzing the maximal cartel price. 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 will show that the maximal cartel price adjusts naturally to policy changes and that cartels are more likely to reduce prices rather than to give up on collusion altogether.

The aim of introducing a leniency program is to destabilize the cartel. In our setting, this means reducing the maximal cartel price below the one under antitrust enforcement without such a program. Under leniency programs, firms that broke the law are allowed to report their illegal activities in exchange for reduced fines. Self-reporting may take place ex-ante before any investigation by the AA starts, or ex-post during an ongoing investigation. While leniency programs introduce additional incentives for the firms to break the cartel agreement, they also broaden the range of cartel strategies, in particular the cartel

2 Profit maximization under nonbinding equilibrium conditions is investigated in Block, Nold, and Sidak (1981), Harrington (2005), and Houba, Motchenkova, and Wen (2010). Harrington (2004) shows that the equilibrium conditions must be nonbinding in case of sufficiently large discount factors. For a special case, Houba, Motchenkova, and Wen (2012) investigate both binding and nonbinding equilibrium conditions for the entire range of discount factors.

may exploit these programs if they are too generous. As extensively discussed in Chen and Rey (2013), this may manifest itself in the form of a "collude and report" strategy and generates adverse effects. This concern is also validated by experimental evidence, see Hinloopen and Soetevent (2008). We also analyze such a strategy together with a standard "collude and never report" strategy and derive the resulting endogenous maximal cartel prices supported by these two types of strategies. We observe that a "collude and report" strategy may not only become more attractive for the cartel, as reported in the literature, but also affects the maximal cartel price.

For ex-ante leniency, our characterization of the maximal cartel price shows that it is impossible to reduce this price below the maximal cartel price under antitrust enforcement without leniency. The reason is that ex-ante leniency introduces an additional incentive compatibility constraint (ICC) that will never be binding for non-reporting cartels. However, if ex-ante leniency is too generous, it has an adverse effect where the collude and report strategy is exploited and the maximal cartel price increases above the one in the absence of leniency. This result is different from that of Motta and Polo (2003), who show that collude and report strategies cannot be sustained in equilibrium under ex-ante leniency.

Our results for ex-post leniency differ from those for ex-ante leniency. Under ex-post leniency, the maximal cartel price can be reduced below the one under antitrust enforcement without leniency. In this case, the additional ICC introduced by ex-post leniency can be binding for cartels which never report and the largest reduction of the maximal cartel price is achieved if single-reporting firms are granted full immunity. Nevertheless, one should still be cautious about the possible adverse effects caused by ex-post leniency. If ex-post leniency is too generous for multiple-reporting firms, it will increase the maximal cartel price and this effect may even undo the improvement mentioned before. To reduce these adverse effects, fine reductions in case of multiple-reporting should be absent or moderate.

When comparing the ex-ante and ex-post leniency programs, we conclude that ex-post leniency should be less generous than ex-ante leniency in case of multiple-reporting firms. The exact relation depends upon the monitoring probability. This result supports the current practice in the US and the EU to differentiate between these fine reductions. It also gives a clear-cut novel policy recommendation on how the effectiveness of leniency programs can be improved.

In the next paragraphs, we relate our study to the existing theoretical literature.[3] The seminal paper on optimal revelation schemes as part of antitrust

3 In addition to the theoretical literature, there are a number of experimental and empirical articles on leniency. The empirical studies have generally investigated the effect of leniency

policy is Motta and Polo (2003), who study an infinitely repeated sequential game with three prices that capture the three most important profit levels as mentioned earlier. In each period, firms decide whether to reveal information about their misconduct. The cartel adopts a grim-trigger strategy in which cheating on the cartel agreement by either setting a different price or applying for leniency triggers competitive behavior forever after. Their agreement also specifies that the cartel continues to operate as usual each time it is caught by the AA. Under the optimal antitrust policy, introduction of ex-post leniency programs induces firms to report and destabilizes the "collude and never report" strategy. However, ex-ante leniency programs are ineffective in their setting. As is also shown in Spagnolo (2004) and Rey (2003), effective ex-ante leniency programs require substantial rewards to the reporting firms.

Chen and Harrington (2007) incorporate ex-ante leniency programs into a special case of the framework in Harrington (2004, 2005) to augment antitrust policy in an environment where cartels arouse suspicions and price is a continuous variable. The price setting in Chen and Harrington (2007) can be regarded as a generalization of the competition phase in Motta and Polo (2003). The grim-trigger strategies 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. In such environment, cartels also need to manage suspicions, modeled as if the cartel keeps in mind an endogenous detection probability. 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. In Chen and Harrington (2007), 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 non-tractable. The analysis, therefore, has to resort to simulations of price paths. Nevertheless, their model admits a steady-state profit-maximizing cartel price that lies above the non-cooperative price, and this price is independent of the leniency program. Hence, the leniency program is ineffective.

The foci of our paper are also detection probabilities and penalty schemes that depend upon cartel pricing, a topic that receives relatively little attention in the literature. However, such modifications are relevant because these are

programs on cartel stability, cartel duration, and harm, see, e.g., Miller (2009) or Brenner (2009, 2011). Recent experimental studies are Apesteguia, Dufwenberg, and Selten (2007), Hinloopen and Soetevent (2008), Bigoni et al. (2012), and Hamaguchi, Kawagoe, and Shibata (2010).

closer to current antitrust policies. Also, in our paper cartel formation and its pricing strategy are endogenized, similar to those of, e.g., Harrington (2004, 2005). Our model extends the approach of Motta and Polo (2003) by including endogenous cartel behavior, the presence of suspicions, and a general class of endogenous antitrust policies with penalty schemes proportional to illegal gains and the possibility of different fine reductions to single-reporting and multiple-reporting firms. 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 endogenous fine reductions. Our model unifies some different assumptions in Motta and Polo (2003) and Chen and Harrington (2007) and bridges them together. A major difference between our focus and that of most existing studies is that we consider the maximal cartel price.[4]

Similar to the existing theoretical literature, we select particular policy rules adopted by the AA. The two essential policy rules that crucially influence the strategies of the firms are whether price-deviating firms are punished and whether repeat offenders are allowed to obtain reduced fines. Similar to Motta and Polo (2003), we adopt the setting where price-deviating firms are not punished and leniency is available to both first-time and repeat offenders. These policy rules are part of the optimal design according to available theoretical literature, see, e.g., Spagnolo (2004), Rey (2003), or Chen and Rey (2013). Spagnolo (2004) and Rey (2003) show that not punishing price-deviating firms increases the incentives of a colluding firm to deviate and, hence, destabilizes collusion the most. We select the policy rule of not punishing price-deviating firms in order to investigate the role leniency might have augmenting optimal antitrust enforcement without leniency. Chen and Rey (2013) and Wils (2008) advocate policies that grant leniency to repeat offenders as well as first-time offenders, as not allowing repeat offenders to apply for leniency will in the long run make leniency inapplicable for the majority of cartels. Moreover, this policy is currently employed in practice, see, e.g., EC Leniency Guidelines (2006).

The remainder of this paper is organized as follows. Section 2 outlines the model. In Section 3, we analyze the effects of antitrust enforcement without leniency. In Sections 4 and 5, we analyze ex-ante leniency programs and ex-post leniency programs, respectively. Section 6 concludes the analysis and discusses several policy recommendations.

4 Our analysis is complementary to the analysis of the profit-maximizing cartel price in Harrington (2004, 2005), where the main focus is on sufficiently large discount factors under which the equilibrium conditions will be nonbinding.

2 The Model

There are $n \geq 2$ firms that compete in each of infinitely many periods in the presence of an antitrust policy with a leniency program. In each period, the firms first choose their prices, and then decide whether to apply for leniency if they have colluded in setting prices. Since the firms choose these two types of actions sequentially, the model we adopt in this paper is an infinitely repeated sequential game of Wen (2002). The antitrust policy consists of two stages: (1) after observing the market outcome, the AA will investigate possible collusion among the firms probabilistically; (2) without additional evidence (particularly from the colluding firms), the AA will be able to prosecute and fine the firms successfully with some probability. Leniency programs allow amnesty or reduced fines if some firms report evidences to the AA so that the AA will be able to prosecute the case with certainty. There are two versions of the leniency program, ex-ante and ex-post. The difference depends on when leniency may apply, before (ex-ante) the AA starts its investigation or after (ex-post) its investigation but before the AA is able to conclude the case on its own. We will analyze three policy regimes, antitrust without leniency as well as antitrust with an ex-ante or ex-post leniency program.

The firms will take the antitrust policy with or without a leniency program as exogenous. It is well known that repeated games such as the one used in this paper admit multiple subgame perfect equilibria, commonly known as the folk theorem. In a subgame perfect equilibrium, a firm always takes the optimal action at any stage of the game no matter how the game has been played, given how the other firms react in the future as specified by the equilibrium strategies. We will focus on a special class of stationary subgame perfect equilibria that have been the main focus in this literature,[5] in which the firms take the same actions (cartel agreement in pricing and reporting strategies) as long as no firm has deviated; otherwise, the firms will compete non-cooperatively in the future. The underlying rationale is that cartels are based upon trust and this trust will be gone once some firms deviate from their cartel agreement. We call the strategies underlying these equilibria modified grim-trigger strategies, because our model is a modified repeated game with a two-stage sequential game of pricing and reporting in every period.

Price competition in each period is modeled as a symmetric Bertrand model among the n firms with either homogeneous or heterogeneous products. Let $\pi(p_1, \ldots, p_n)$ be a firm's per-period profit for prices $p_1, \ldots, p_n \in \mathbb{R}_+$. Since

5 See, e.g., Motta and Polo (2003), Rey (2003), Spagnolo (2004, 2008), and Chen and Rey (2013).

we mostly deal with symmetric outcomes, denote $\pi(p,\ldots,p) \equiv \pi(p)$ for simplicity. We denote the static Nash/non-cooperative equilibrium price and the maximal collusive/monopoly price by p^N and p^M, respectively. In each period, the firms first decide a symmetric collusive price, say $p \in (p^N, p^M]$.[6] To analyze cartel stability, denote a firm's profit from unilateral deviation when all the other firms choose price p by $\pi^{\text{opt}}(p) = \sup_{p'} \pi(p', p, \ldots, p)$.[7] As in Harrington (2004, 2005), we assume that $\pi(p)$ is continuous and strictly increasing in $p \in [p^N, p^M]$, $\pi^{\text{opt}}(p)$ is continuous, strictly increasing, and $\pi^{\text{opt}}(p) > \pi(p) > 0$ for $p \in (p^N, p^M]$. Without loss of generality, we normalize this static Bertrand model so that $\pi(p^N) = 0$, so one may interpret $\pi(p)$ as the net profit above the non-cooperative profit $\pi(p^N) = 0$.

The antitrust policy has the following structure in each period: Given $p \in [p^N, p^M]$,

1. the AA will launch an investigation on the firms for possible colluding activities with probability $\mu(p)$;
2. upon being investigated, the firms will be proved to be guilty of collusion with probability $\theta(p)$, and upon being convicted, every firm will be fined by $k(p)\pi(p)$.

Assume that both $\mu(\cdot)$ and $\theta(\cdot)$ are non-decreasing in p, as a higher cartel markup will invoke more attention from the AA, and more likely the AA can prove and detect the cartel (see Harrington (2004, 2005) and Connor and Bolotova (2006)). We also assume that the AA does not make a type-I error in prosecuting with $\theta(p^N) = 0$, but may make a type-II error with $\theta(p) < 1$ for all $p \in (p^N, p^M]$. As in Rey (2003), we assume that the AA will prosecute the firms only based on the firms' misconduct during the current period, but not from early periods.[8] Moreover, we adopt the treatment of Motta and Polo (2003) that a price-deviating firm is not prosecuted. Based on the current practice in many countries, $k(\cdot)$ is assumed to be non-decreasing as the fine is often positively

6 Bageri, Katsoulacos, and Spagnolo (2013) and Katsoulacos and Ulph (2013) show that antitrust fines based upon turnover could induce Cournot oligopolies to collude at a price even above the monopoly price.

7 We deliberately write "sup" instead of "max" so that $\pi^{\text{opt}}(\cdot)$ is well defined even for the Bertrand model with homogeneous products.

8 Having the penalty depend on the current price markup relaxes a restrictive assumption of fixed penalties imposed in most of the theoretical leniency analyses. It also allows us to obtain tractable results. However, our setup still does not fully reflect the current sentencing guidelines, where the penalty also depends on the duration of the cartel. Unfortunately, incorporating duration will be at the expense of stationarity, see, e.g., Harrington (2014). Therefore, we leave this extension to future research.

related to the social harm caused by the cartel and the firms' illegal gains. We assume $k(p^N) = 0$ by default as p^N is the non-cooperative market equilibrium price. The fine schedule $k(\cdot)$ could also include private damages, as in Harrington (2004, 2005). For technical tractability, we assume that all these three policy functions are continuous in $p \in (p^N, p^M]$ and have well-defined right limits at p^N as they may not be continuous at p^N due to $\theta(p^N) = 0$ and $k(p^N) = 0$. Finally, the analyses in Bageri, Katsoulacos, and Spagnolo (2013) and Katsoulacos and Ulph (2013) suggest detection probabilities around 0.15 and levels of fines in the range of two to three times the illegal gains of the cartel. This implies an expected penalty roughly between 30% and 50% of illegal gains, which are of the same order of magnitude as estimated by Bryant and Eckard (1991). Accordingly, we assume that $0 \leq \mu(p)\theta(p)k(p) < 1$ for all $p \in (p^N, p^M]$. Hence, the expected fine is always lower than the abnormal profit and the firms are always tempted to collude in setting their prices.

As described early in this section, leniency programs may grant total amnesty or partially reduced fines to the firms that cooperate with the AA, in which case the AA will be able to prosecute the other colluding firms with certainty. We will be more specific on how to incorporate ex-ante and ex-post leniency programs into the model in Sections 4 and 5, respectively.

There are two different treatments in the literature of how the firms behave after being detected and convicted by the AA. While Harrington (2004) assumes that the cartel dissolves forever once this occurs, Motta and Polo (2003) assume that cartel re-establishes every time it is detected and convicted by the AA, which is consistent with profit-maximizing behavior. To include these two polarized cases, we assume that every time the firms are detected and convicted by the AA, they will compete non-cooperatively in the future with probability γ and re-establish the cartel with probability $1-\gamma$ in the following period. How the firms behave after being convicted does not affect the main results in the mentioned papers, but simply enhances or reduces the impact of detection and conviction, which is also the case in our analysis.

Given the antitrust policy as described so far, the firms will first choose prices, and then decide whether to report to the AA if there is a leniency program in place in every period. The firms evaluate their strategies based on their total discounted profit over infinitely many periods with a common discount factor $\delta \in (0,1)$ per period. We follow the mainstream literature by focusing on two classes of stationary subgame perfect equilibria:[9] (1) the firms always collude and never report to the AA, called silent cartel, and (2) the firms always

9 See, e.g., Motta and Polo (2003), Spagnolo (2004), and Chen and Rey (2013).

collude and always apply for leniency, called systematically reporting cartel. Even with these two classes of equilibria, there are still many sustainable cartel prices. Moreover, the union of all these cartel prices covers the range of subgame perfect equilibrium prices that are sustainable in the class of grim-trigger strategies.

For each of the three possible policy regimes, we are interested in the maximal cartel price. The maximal cartel price is the highest cartel price among the cartel prices that are sustainable within the class of stationary grim-trigger strategies we consider. In what follows, we express the maximal cartel price as the solution of an optimization program in which we maximize the cartel price over the subset of all cartel prices that are sustainable. Such cartel price represents the worst possible outcome in terms of deadweight loss in social welfare. On the technical side, the equilibrium conditions are the standard ones based upon the ICCs of cartel participants not to deviate either in choosing prices or deciding whether to report to the AA. In contrast to the literature, these constraints are solved explicitly in the variable of interest that is the maximal cartel price. In other words, the maximal cartel price is determined by these constraints only.

3 Antitrust Policy without Leniency

In this section, we analyze the maximal cartel price under antitrust policy without a leniency program as a benchmark. Since colluding firms are not eligible for reduced fines, the firms only need to decide prices in each period. Accordingly, there is only one ICC for cartel stability under which no individual firm has any incentive to deviate from the cartel price. Keeping in line with the class of stationary subgame perfect equilibria, we will focus on the following modified grim-trigger strategy profile to sustain a cartel price $p > p^N$:

1. Firms set price $p > p^N$ in the first and in all subsequent periods as long as there was no price deviation. Any price deviation by some firms leads to 3.
2. Every time the cartel is detected and convicted by the AA, the firms continue to collude under 1 with probability $1-\gamma$, or compete non-cooperatively under 3 with probability γ.
3. The firms compete at static Nash equilibrium price p^N in every period.

Let $V(p)$ be the present value of a firm's expected profit from the modified grim-trigger strategy profile described above and it is determined by the following recursive dynamics:

$$V(p) = \pi(p) + \beta(p)\left[-k(p)\pi(p) + \delta\left(\gamma\frac{\pi(p^N)}{1-\delta} + (1-\gamma)V(p)\right)\right] + (1-\beta(p))\delta V(p), \quad [1]$$

where we denote $\beta(p) = \mu(p)\theta(p)$ for convenience wherever possible. In words, $\beta(p)$ is the probability that AA can successfully investigate and prosecute the cartel when the firms collude at price $p > p^N$. Given our assumptions on $\mu(\cdot)$ and $\theta(\cdot)$, $\beta(\cdot)$ is non-decreasing and continuous for $p \in (p^N, p^M]$, and has a well-defined right limit at p^N because $\beta(p^N) = 0$ and $\beta(\cdot)$ may be discontinuous at p^N. With normalization $\pi(p^N) = 0$, solving for $V(p)$ from eq. [1] yields

$$V(p) = \frac{1-\beta(p)k(p)}{1-\delta+\delta\gamma\beta(p)}\pi(p). \quad [2]$$

Note that for all $p \in (p^N, p^M]$, we have $V(p) > 0$ due to $\beta(p)k(p) < 1$.

To sustain this modified grim-trigger strategy profile as an equilibrium, the one-stage deviation principle (see Fudenberg and Tirole 1991) leads to the following ICC:

$$V(p) \geq \pi^{\text{opt}}(p) + \frac{\delta}{1-\delta}\pi(p^N) \Rightarrow \frac{\pi(p)}{\pi^{\text{opt}}(p)} \geq \frac{1-\delta+\delta\gamma\beta(p)}{1-\beta(p)k(p)}. \quad [3]$$

The left-hand side of eq. [3] $\pi(p)/\pi^{\text{opt}}(p)$ measures the relative size of cartel gains to the net gains under a firm's best unilateral deviation, which also plays an important role in our later analysis. Accordingly, denote $\lambda(p) = \pi(p)/\pi^{\text{opt}}(p) < 1$ for $p \in (p^N, p^M]$ and $\lambda(p^N) = 1$ by default. The higher the $\lambda(\cdot)$, the less incentive each firm would have to deviate from a cartel, the more stable the cartel could be. For technical convenience, we assume $\lambda(\cdot)$ is non-increasing and has a well-defined right limit at p^N, say $\bar{\lambda} \leq 1$. Note that $\lambda(\cdot)$ may not be continuous at p^N, such as in the Bertrand model with homogeneous products.

The right-hand side of eq. [3], denoted as $\Lambda(p)$ for exposition simplicity, is mostly determined by the antitrust policy. It is straightforward to verify that $\Lambda(\cdot)$ is non-decreasing and continuous for $p \in (p^N, p^M]$, and lowering $\beta(\cdot)$ and/or $k(\cdot)$ will reduce $\Lambda(\cdot)$. Similarly, $\Lambda(\cdot)$ decreases as δ increases.

The highest cartel price supported by the grim-trigger strategy profile under an antitrust policy without leniency (A) is given by

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

Program [4] is well defined because $p \in [p^N, p^M]$ and eq. [3] induce a closed subinterval of $[p^N, p^M]$ that contains p^N. Given the monotonicity of $\lambda(\cdot)$ and $\Lambda(\cdot)$,

condition [3] alone determines the maximal cartel price under antitrust without leniency.[10] These functions are illustrated in Figure 1, where an interior solution $p^A \in (p^N, p^M)$ is depicted. The boundary solutions are $p^A = p^N$ if the right limit of $\Lambda(\cdot)$ at p^N is at least $\bar{\lambda}$ and $p^A = p^M$ if $\Lambda(p^M) \leq \lambda(p^M)$. The maximal cartel price p^A is bounded from above by the highest cartel price under no antitrust enforcement (or equivalently $\beta(\cdot) = 0$ and/or $k(\cdot) = 0$), denoted as p^C, which is the solution of $\max_{p \in [p^N, p^M]} p$ under $\lambda(p) \geq 1 - \delta$, as illustrated in Figure 1. The following proposition summarizes these results.

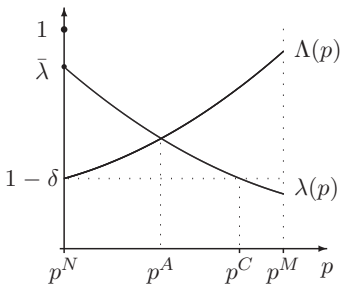

Figure 1: p^A is the largest p such that $\lambda(p) \geq \Lambda(p)$.

Proposition 1 *Under the antitrust policy without leniency, the maximal cartel price p^A has the following properties*:

1. *whenever $p^A \in (p^N, p^M)$, eq. [3] is binding at p^A;*
2. *any $p \in [p^N, p^A]$ is also sustainable by the cartel;*
3. *p^A is non-decreasing in δ and non-increasing in $\beta(\cdot)$ and $k(\cdot)$.*

The comparative statics of the maximal cartel price with respect to the policy parameters and industry characteristics are very intuitive. Firms that are more patient put more value to future illegal gains, which shifts $\Lambda(\cdot)$ downward, and consequently, the firms will be able to sustain a higher maximal cartel price. Similarly, higher $\beta(\cdot)$ and $k(\cdot)$ reduce the cartel's profitability, shift $\Lambda(\cdot)$ upward, and strengthen ICC [3].

10 Note that ICC [3] can also be written in terms of the conventional threshold on the discount factor. To see this, define $\delta^*(p) = 1 - \lambda(p)$ to be the critical discount factor for collusion to be sustainable in the absence of antitrust enforcement, i.e., when $\beta(p) = 0$. Then, for $p \in [p^N, p^M]$, eq. [3] can be rewritten as $\delta \geq \frac{\delta^*(p) + (1-\delta^*(p))\beta(p)k(p)}{1-\gamma\beta(p)}$, which is larger than $\delta^*(p)$ when $\beta(p) > 0$. We thank an anonymous referee for pointing this out. However, we take δ as an exogenous primitive of the model, while p and, hence, the threshold on δ are endogenous.

Next, we illustrate our findings with an example of a simple homogeneous Bertrand oligopoly model with linear demand, which will also be reconsidered to show the effects of ex-ante and ex-post leniency programs in Sections 4 and 5, respectively.

Example 1 *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* $k(p) = k$ *for all* $p \in (p^N, p^M] = (0, 1]$, *where* $k\beta p \leq k\beta < 1$, *and* $\beta(p^N) = k(p^N) = 0$. *Note that, for all* $p \in (p^N, p^M]$, $\pi^{\text{opt}}(p) = n\pi(p)$ *implies* $\lambda(p) = \frac{1}{n}$ *and* $\bar{\lambda} = \frac{1}{n}$. *If* $\delta \leq 1 - \frac{1}{n}$, *then we trivially obtain that collusion on any price* $p > p^N$ *is not sustainable. Otherwise, when* $\delta > 1 - \frac{1}{n}$, *ICC [3] becomes*

$$\frac{1}{n} \geq \frac{1 - \delta + \gamma\delta\beta p}{1 - k\beta p} \Rightarrow p^A = \frac{1 - n(1 - \delta)}{(n\gamma\delta + k)\beta} > 0.$$

The upper bound on p is the maximal cartel price p^A *whenever it is less than 1, otherwise* $p^A = p^M = 1$. *Note that the maximal cartel price* p^A *given above has all the intuitive comparative statics with respect to the policy parameters.*

Note that for homogeneous Bertrand competition, the function $\lambda(p)$ is constant. Alternatively, in differentiated products price competition setting this function will be decreasing.[11] Following these observations, it should be stressed that if the constituent Bertrand game is homogeneous, we need policy parameters as a function of p in order to have any impact of the enforcement policy on cartel prices, that is, we need an upward sloping $\Lambda(p)$ as in the example. Conversely, when $\lambda(p)$ is decreasing in p, even a $\Lambda(p)$ that is flat is sufficient to obtain the maximum price consistent with the ICC and that varies with the (non-price contingent) policy parameters, that affect the level of $\Lambda(p)$. Moreover, more generally one may add that price contingent policy parameters that make the expected fine monotonically increasing in the cartel price increase the "virtual" marginal costs compared to the true flat marginal cost. Then, the intersection of the marginal revenue and marginal costs occurs for a higher cartel price than with non-contingent fines.

The model of this section can be seen as a repeated game interpretation of the model in Harrington (2004, 2005) in which each convicted firm pays private damages that are increasing in profits, $k(p)\pi(p)$ in terms of our model. Harrington (2004, 2005) shows that, for a large enough discount factor, the ICC is nonbinding in the cartel's profit-maximization problem and, then, the

11 Note that under Cournot competition the ratio $\lambda(q) = \pi(q)/\pi^{\text{opt}}(q)$ should be expressed in terms of quantities. Then $\lambda(q)$ is an increasing function of quantity chosen. Hence, this case will be similar to the differentiated Bertrand case.

profit-maximizing steady-state cartel price is decreasing in the discount factor. Furthermore, this steady-state cartel price lies below the monopoly price. The essential difference of our study is that we investigate the cartel price for which the ICC is binding on the entire range of discount factors.[12] In contrast to Harrington (2004, 2005), Proposition 1 states that our variable of interest, the maximal cartel price, is increasing in the discount factor, which is confirmed in Example 1. Similar to these references, Figure 1 implies that the maximal cartel price is lower than the monopoly price. Also, both the profit-maximizing steady-state cartel price and the maximal cartel price respond similar to the policy instruments $\beta(\cdot)$ and $k(\cdot)$.

Finally, it is easy to assess the impact of prosecuting price-deviating firms, which we ruled out by assumption. In that case, the term $\pi^{\text{opt}}(p)$ on the right-hand side of ICC [3] would be reduced, and this relaxes the ICC. In the graphical analysis of Figure 1, this would shift the curve of the properly modified $\lambda(p)$ upward and, hence, increase the maximal cartel price. So, in our general setting, antitrust enforcement would be most effective if the incentives for price-deviating firms are optimal, and this requires that price-deviating firms are never fined. This extends the insights obtained for profit-maximizing cartels in Spagnolo (2004) and Rey (2003) to the maximal cartel price. As Spagnolo (2004) puts it, the "[l]aw enforcement agency should carefully avoid prosecuting cartel members who unilaterally defected from collusive strategies, and should make this policy of public domain." Our motivation to adopt this optimal policy is to investigate the role leniency might have augmenting such optimal antitrust enforcement.

It should also be stressed that the assumption of (not) punishing price-deviating firms appears to be essential for our results and results in related articles. This assumption varies among different theoretical approaches to model leniency and antitrust rules in general.[13] For example, Spagnolo (2004), Cyrenne (1999), Rey (2003), Chen and Harrington (2007), Chen and Rey (2013), and Jensen and Sorgard (2014) all allow for the possibility of punishing price-

12 For results on profit maximization for the entire range of discount factors in the model of this section, we refer to Houba, Motchenkova, and Wen (2012).

13 Under antitrust policy with leniency, which does not punish a single price-deviating firm only if it reports, it is optimal for a deviating firm to report. As indicated in Jensen and Sorgard (2014), this will give rise to the additional stationary strategy "price deviate and report" that will dominate the strategy systematic "collusion and reporting." According to Harrington (2008), such model in a repeated game setting can explain the "Deviator Amnesty effect" but not the "Cartel Amnesty effect." In contrast, Motta and Polo (2003) adopt the opposite assumption and can only explain the latter effect but not the former effect. So far, only Harrington (2008) obtained both effects in a single model.

deviating firms, while Motta and Polo (2003) assume the opposite. As a result, our treatment of leniency programs is very similar to Motta and Polo (2003) and, hence, rests on the analysis of three relevant stationary strategies that are called in their terminology: deviation strategy (no collusion in our case), collude and never report strategy (our silent cartel), and collude and report strategy (our systematically reporting cartel).[14]

4 Ex-Ante Leniency Programs

In this section, we analyze the maximal cartel price under antitrust policy with an ex-ante leniency program, where colluding firms may apply for leniency only before the AA starts its investigation.

The ex-ante leniency program is modeled as follows. Let $\alpha(p, s_1, \ldots, s_n) \leq k(p)$ be an individual firm's reduced one-time fine rate in case the cartel price is $p \in (p^N, p^M]$ and the firms' reporting decisions are $s_1, \ldots, s_n \in \{R, N\}$, where R (N) stands for (not) Reporting. Since we mostly deal with symmetric outcomes, we only need two reduced-fine rates:

1. If only one firm reports, the reduced-fine rate to the only self-reporting firm is $\alpha(p, N) \equiv \alpha(p, R, N, \ldots, N)$ for simplicity.
2. If *all* firms report, then a reduced-fine rate $\alpha(p, R) \equiv \alpha(p, R, \ldots, R)$ applies to every reporting firm.

We assume that both $\alpha(p, N)$ and $\alpha(p, R)$ are non-decreasing, continuous in $p \in (p^N, p^M]$, and have well-defined limits at p^N. According to current leniency guidelines, we also impose that $0 \leq \alpha(p, N) \leq \alpha(p, R) \leq k(p)$. Note that requiring $\alpha(p, N) \geq 0$ excludes the possibility to reward any self-reporting firm.

Our setup is quite flexible, because it includes the possibility of $\alpha(p, R) = k(p)$ for repeat offenders. In practice leniency discounts are designed to vary with respect of the ordering of arrivals (first versus later reporters) and not with the number of reporters. In the US, for example, full amnesty is granted to the first reporter and no reduction at all to the subsequent reporters. In that respect the sequence of application ideally should be taken into account. In our stationary environment the US system can be approximated by "orchestrated race to the court house," in which firms randomly determine who applies for leniency as the first applicant. In that case $\alpha(p, R)$ is equivalent to $\frac{1}{n}\alpha(p, N) + \frac{n-1}{n}k(p)$ and has a stochastic interpretation.

14 Consequently, we are able to capture the "Cartel Amnesty effect," while the "Deviator Amnesty effect" is absent in our paper.

Given that the traditional antitrust policy alone may not be sufficient to eradicate all cartel prices, the question is then if and when the antitrust policy with such an ex-ante leniency program is effective in limiting cartel market power. This means that, for given $\mu(\cdot)$, $\theta(\cdot)$, and $k(\cdot)$, we analyze what reduced-fine rates $\alpha(\cdot,\cdot)$ can accomplish in terms of deterring cartels or lowering the maximal cartel price. As mentioned above, the current literature on leniency program mainly focuses on two types of stationary subgame perfect equilibria, where the colluding firms either never report or always report to the AA.

4.1 Silent Cartels under Ex-Ante Leniency

In this subsection, our analysis focuses on a stationary subgame perfect equilibrium where the firms always collude but choose never to report to the AA. "Not report" is part of the firms' cartel agreement. In other words, the cartel operates silently. A silent cartel still dissolves deterministically after some firms deviate in setting their price and/or deviate by reporting to the AA, or probabilistically after the cartel is detected and convicted by the AA. More specifically, consider the following modified stationary grim-trigger strategy profile to sustain a cartel price of $p > p^N$:

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 is no deviation. Any deviation by some of the firms leads to 3.
2. As long as there is no deviation in price setting or reporting, every time the cartel is detected and convicted by the AA firms continue under 1 with probability $1-\gamma$ and go to 3 with probability γ.
3. All firms set the static Nash equilibrium price p^N in every period.

According to this stationary strategy profile, the present value of an individual firm's expected profits is still $V(p)$ as given by eq. [2]. Under ex-ante leniency, a firm has two types of actions to take along the equilibrium path, hence there are two ICCs. First, eq. [3] continues to ensure that no firm has any incentive to undercut the cartel price p. Second, a firm should have no incentive to report to the AA in the presence of the ex-ante leniency program. By the one-stage deviation principle, these two ICCs imply that a firm would have no incentive to undercut the cartel price and to report to the AA.

We now formulate the second constraint that ensures no firm will report to the AA. A firm's expected continuation profit from reporting to the AA consists of paying the reduced fine $\alpha(p,N)\pi(p)$ in the current period followed by the

continuation profits from setting the static Nash equilibrium price p^N forever after. Therefore, a colluding firm will not report if and only if

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

This constraint is nontrivial for $\alpha(p,N)<1$, and simplifies under this condition to

$$\frac{1}{1-\alpha(p,N)} \geq \Lambda(p). \qquad [5]$$

The left-hand side of eq. [5] has the same properties as $\alpha(p,N)$, namely it is non-decreasing in p and has a well-defined right limit at p^N.

With the two incentive constraints [3] and [5] under the ex-ante leniency program, the maximal sustainable price by a silent cartel is then given by

$$p^S = \max_{p\in[p^N,p^M]} p, \quad s.t.\ (3) \text{ and } (5). \qquad [6]$$

Since both constraints [3] and [5] involve weak inequalities between two continuous functions of $p \in [p^N,p^M]$, Program [6] is well defined. Since $\alpha(p,N)$ is non-decreasing in p, eq. [5] may not restrict p to a well-behaved compact interval as eq. [3] does. However, $\alpha(p,N) \in [0,1)$ implies that $\frac{1}{1-\alpha(p,N)} \geq 1 \geq \lambda(p)$ for all $p \in [p^N,p^M]$. For all $\alpha(p,N) \in [0,k(p)]$, ICC [5] will never be binding in Program [6]. Note that even full amnesty to a single-reporting firm is insufficient to be effective. Consequently, the maximal cartel price sustained by a silent cartel under ex-ante leniency program is given by $p^S = p^A$, as illustrated in Figure 2.

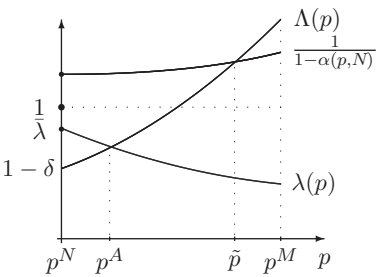

Figure 2: The maximal cartel price set by silent cartels is p^A, because ICC [5] is nonbinding in Program [6].

Proposition 2 *Under ex-ante leniency, the maximal price p^S sustainable by a silent cartel is the same as the maximal cartel price under antitrust without leniency p^A, so it has the same comparative statics as p^A stated in Proposition 1.*

The implication of Proposition 2 is that if the cartel can sustain the cartel price $p \in (p^N, p^A]$ under antitrust enforcement, then introducing an ex-ante leniency program allows the cartel to maintain its illegal activity with the same cartel price by operating silently. Hence, such a program is not effective in reducing cartel prices. Proposition 2 relates to the findings in Motta and Polo (2003), Spagnolo (2004), and Rey (2003), who all stress the ineffectiveness of ex-ante leniency programs without rewards in destabilizing collusion on the monopoly price.[15] Our framework with a continuum of possible collusive prices is more general. Treating the cartel price as a continuous variable allows us to employ a more powerful technique and derive richer results on cartel pricing. Given the low expected fine $\beta(p)k(p) < 1$, any fine reduction for the first reporting firm under an ex-ante leniency program would not be sufficient either to reduce the maximal cartel price or to block cartel formation. The reason is that the additional ICC introduced by the ex-ante leniency program is never binding. Intuitively, self-reporting is always less attractive than a price deviation for any firm. By a price deviation, a firm increases its profit in the current period and also it will not be fined afterward. On the other hand, by self-reporting a firm does not enjoy such an increase in profits in the current period and only faces a reduced fine.

These results are illustrated in the linear Bertrand oligopoly of Example 1.

Example 2 *Reconsider the homogeneous Bertrand oligopoly model of Example 1 and the linear reduced-fine function* $\alpha(p,N) = \alpha_N p$ *for* $\alpha_N \in [0,1)$. *Program [6] becomes*

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

Because $\frac{1}{1-\alpha_N p} \geq 1 > \frac{1}{n}$, *the second constraint is never binding. The first constraint can be rewritten as* $p \leq \frac{1-n(1-\delta)}{(n\gamma\delta+k)\beta}$, *which is the maximal cartel price* p^S *under silent cartels. Hence, we have* $p^S = p^A = \frac{1-n(1-\delta)}{(n\gamma\delta+k)\beta}$, *confirming Proposition 2.*

4.2 Systematically Reporting Cartels under Ex-Ante Leniency

As discussed in the introduction, many researchers also studied stationary subgame perfect equilibria, in which some cartels may exploit the ex-ante

15 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 minimal reward needed to upset cartel prices by allowing $\alpha(\cdot, N) < 0$. For a detailed analysis we refer to Houba, Motchenkova, and Wen (2009), an earlier version of the current study.

leniency program by systemically colluding and reporting. Such equilibria have also attracted a fair amount of attention when studying both ex-ante and ex-post leniency programs.[16] In this subsection, we focus on the maximal cartel price sustained by a cartel where its members systematically report. In this setting, reporting to the AA is part of the cartel agreement.

Consider the following modified stationary grim-trigger strategy profile to sustain a cartel price of $p > p^N$:

1. Firms set price $p > p^N$ and report in the first period and continue to do so as long as there is no *price* deviation. Any price deviation will lead to 3.
2. A price-deviating firm does not report in the period of deviation.
3. All firms set the static Nash equilibrium price p^N in every period.

Note that the cartel still breaks down if a firm undercuts the cartel price, but not if a firm refrains from reporting. According to this strategy profile, the cartel will re-establish in every period. The present value of a firm's profit, denoted as $V^R(p)$, is determined by the following recursive dynamics:

$$V^R(p) = \pi(p) - \alpha(p,R)\pi(p) + \delta V^R(p) \Rightarrow V^R(p) = \frac{1-\alpha(p,R)}{1-\delta}\pi(p). \qquad [7]$$

As in Section 4.1, there are two types of deviations from the cartel agreement an individual firm may take: undercut the price and not report to the AA. It is easy to see that given all the other firms will report, not reporting to the AA is not optimal for any individual firm because not report will only reduce a firm's profit in the current period, and it does not change the continuation equilibrium. The incentive constraint to ensure no firm will undercut the cartel price is given by

$$\pi^{\text{opt}}(p) + \frac{\delta}{1-\delta}\pi(p^N) \leq V^R(p) \Rightarrow [1-\alpha(p,R)]\lambda(p) \geq 1-\delta.$$

Because the right-hand side of the latter inequality is always positive, we continue with $\alpha(p,R) < 1$ and obtain

$$\lambda(p) \geq \frac{1-\delta}{1-\alpha(p,R)}. \qquad [8]$$

16 Note that feasibility of such strategy rests on the assumption that the antitrust policy is stationary and leniency is available for both first-time offenders and repeat offenders, which corresponds to the structure of the current EU Leniency Guideline, see, e.g., EC (2006) and Wils (2008). Moreover, this strategy and underlying assumption have been adopted in Motta and Polo (2003), Spagnolo (2004), Chen and Harrington (2007), Chen and Rey (2013), and observed in experimental settings in Hinloopen and Soetevent (2008).

The properties of $\alpha(\cdot, R)$ carry over to the expression $\frac{1-\delta}{1-\alpha(p,R)}$. Hence, the right-hand side of eq. [8] is non-decreasing in p and has a well-defined right limit at p^N. Also, this right-hand side is decreasing in δ and increasing in $\alpha(\cdot, R)$.[17]

The maximal cartel price when the colluding firms always report can be formulated as

$$p^R = \max_{p \in [p^N, p^M]} p, \quad s.t.\ (8). \tag{9}$$

This program is well defined because $p \in [p^N, p^M]$ and eq. [8] induces a closed subinterval of $[p^N, p^M]$ that contains p^N. Given the monotonicity of $\lambda(\cdot)$ and $\frac{1-\delta}{1-\alpha(\cdot,R)}$, condition [8] alone determines the maximal cartel price set by a systematically reporting cartel. These functions and the interval of sustainable cartel prices $[p^N, p^R]$ are illustrated in Figure 3, where an interior solution $p^R \in (p^N, p^M)$ is depicted. The boundary solutions are $p^R = p^M$ if $\lambda(p^M) \geq \frac{1-\delta}{1-\alpha(p^M,R)}$ and $p^R = p^N$ if $\bar{\lambda} < \frac{1-\delta}{1-\underline{\alpha}(R)}$, where $\underline{\alpha}(R)$ denotes the right limit of $\alpha(\cdot, R)$. The condition for $p^R = p^N$ can be rewritten as

$$\underline{\alpha}(R) > 1 - \frac{1-\delta}{\bar{\lambda}}, \tag{10}$$

which separates the leniency policy from the industry characteristics. Note that the right-hand side of eq. [10] is non-negative for $\delta \geq 1 - \bar{\lambda}$.

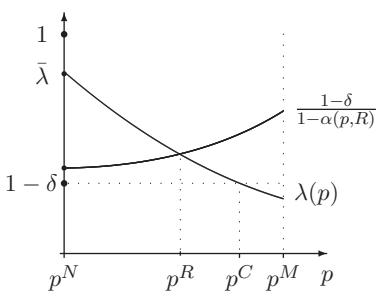

Figure 3: The maximal cartel price p^R is the largest p such that $\frac{1-\delta}{1-\alpha(p,R)} \leq \lambda(p)$.

Shifting the curve $\alpha(\cdot, R)$ upward would tighten ICC [8] and reduce the maximal cartel price. Hence, effective ex-ante leniency programs in lowering the maximal cartel price sustained by a systematically reporting cartel would be to set

17 In terms of the conventional threshold on the discount factor discussed in Footnote 10, we obtain $\delta \geq \delta^*(p) + (1 - \delta^*(p))\alpha(p, R)$. As before, contrary to the conventional threshold, in our setting δ is exogenous, while p and this threshold are endogenous.

$$\alpha(p,R) > 1 - \frac{1-\delta}{\lambda(p)}, \quad \text{for all } p \in \left(p^N, p^M\right]. \tag{11}$$

So, combined with the trivial case, the cartel cannot sustain the cartel price $p > p^N$ if $\alpha(p,R)$ lies between $1 - \frac{1-\delta}{\lambda(p)}$ and $k(p)$. It is evident that any such reduced-fine rate $\alpha(\cdot,R)$ can deter the cartel from exploiting leniency by systematic collusion and reporting. All these rates $\alpha(\cdot,R)$ implement $p^R = p^N$ and, hence, are both price and welfare equivalent. However, the upper bound $\alpha(\cdot,R) = k(\cdot)$ would be the easiest to implement in practice as it does not require any additional information about the industry characteristics. Also, to ensure the effectiveness in reducing the maximal cartel price in as many industries as possible, the AA should set $\alpha(\cdot,R)$ at the maximal permissible rate, $\alpha(\cdot,R) = k(\cdot)$.[18] For these reasons, we call this rate the most effective ex-ante leniency program.

We now summarize these results in the following proposition:

Proposition 3 *Under ex-ante leniency, the maximal price p^R sustained by a systematically reporting cartel has the following properties*:

1. *whenever $p^R \in (p^N, p^M)$, eq. [8] is binding at p^R*;
2. *p^R is non-decreasing in δ and non-increasing in $\alpha(\cdot,R)$*;
3. *for systematically reporting cartels, the most effective ex-ante leniency program sets $\alpha(p,R) = k(p)$ for all $p \in (p^N, p^M]$ and achieves $p^R = p^N$.*

This proposition implies that effective ex-ante leniency programs, which reduce the maximal cartel price sustained by systematically reporting cartels, should impose reduced penalty rates that are not too generous. It is evident from part 3 of Proposition 3 that a properly designed ex-ante leniency program can deter cartels from exploiting leniency by systematic collusion and reporting. For that the design of leniency guidelines should avoid fine reductions in case of multiple reporting. By setting $\alpha(\cdot,R)$ at the maximal permissible rate, $\alpha(\cdot,R) = k(\cdot)$, the AA ensures effectiveness in reducing the maximal cartel price in as many industries as possible when facing the possibility of systematically reporting cartels. In contrast to exploitability considerations, it is not the class of strategies in which the cartel systematically colludes and reports, but rather silent cartel strategies that cause the inefficiency in designing optimal ex-ante leniency programs. Silent collusion cannot be deterred by ex-ante leniency as it induces

18 Based on the interpretation of an orchestrated race to the courthouse, where $\alpha(p,R) = \frac{1}{n}\alpha(p,N) + \frac{n-1}{n}k(p) \leq k(p)$, the logic of the arguments in this subsection are a reason to set $\alpha(p,R) = k(p)$ by setting $\alpha(p,N) = k(p)$, which does not alter the results of Proposition 2.

$p^S = p^A > p^N$ independent of $\alpha(\cdot, N) \geq 0$, while systematic colluding and reporting can be deterred as it induces $p^R = p^N$ when $\alpha(\cdot, R)$ is sufficiently high.

Next, we revisit the Bertrand oligopoly example with linear demand.

Example 3 *Reconsider Example 1 under the ex-ante leniency program with linear reduced-fine function $\alpha(p, R) = \alpha_R p$, $\alpha_R > 0$, for all $p \in [p^N, p^M]$. Note that the right-hand side of eq. [11] is positive for all $p \in [p^N, p^M]$ when $\delta > 1 - \frac{1}{n}$. Then, Program [9] becomes*

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

where $p = p^N = 0$ is feasible in the constraint. Proceeding similar as in Example 1, we rewrite the constraint as

$$p \leq \frac{1 - n(1-\delta)}{\alpha_R},$$

where the upper bound is always positive and equal to the maximal cartel price $p^R < p^M = 1$ whenever $\alpha_R > 1 - n(1-\delta)$. Otherwise, $p^R = p^M$.

We conclude this subsection by discussing the plausibility of the systematic collusion and reporting strategy profile and the practical evidence that supports our analysis. Recall that feasibility of such a strategy profile rests on the assumption that the antitrust policy is stationary and the leniency program treats both first-time offenders as well as repeat offenders the same. This assumption obviously corresponds to the structure of the current EU Leniency Guidelines, which do not prohibit repeat offenders from amnesty (see, e.g., EC (2006) and Wils (2008)). Furthermore, Wils (2008)[19] also puts forward a number of arguments against excluding repeat offenders from leniency: not allowing repeat offenders to apply for leniency would restrict the leniency program only to a limited subset of existing cartels and over time will make leniency programs inapplicable for the majority of cartels. The analysis in Chen and Rey (2013) validates this concern and suggests that the AA should be cautious before refusing to grant leniency to repeated offenders, unless it can deter exposed cartels from returning to collusion. They formally discuss optimality of the policy that grants leniency to repeat offenders as well as to first-time offenders. Chen and Rey (2013) conclude that leniency policy that gives discounts only once, meaning that repeat offenders cannot receive immunity, is suboptimal. Their recommendation is not to prohibit leniency for repeat offenders.

For a richer model with an ex-ante leniency program that excludes repeat offenders from fine discounts, Proposition 3 describes the subgame perfect

19 See Wils (2008, footnote 139, p. 138).

equilibrium in the subgame where the cartel reported at least once. As Example 3 illustrates for linear fine rates, repeat offenders are still able to exploit leniency and the maximum cartel price lies above the non-cooperative price p^N. Furthermore, we must also consider that the cartel may replace the systematic collusion and reporting strategy profile with a reporting once and never again strategy profile as was indicated in Chen and Rey (2013). However, their analysis shows that "[t]his form of collusion may actually be more robust than 'collude and report'; therefore, ruling out leniency for repeated offenders may actually weaken antitrust enforcement." In particular, prohibiting leniency for repeated offenders creates robust alternative collusion strategies: by reporting once, cartel members can make sure that no one has an incentive to report afterward, which thus stabilizes normal collusion in the future.[20] This also contributes to stabilizing collusion in the first period.

Introducing sufficient re-incidence penalties,[21] which would deter convicted cartels from returning to collusion, could be a solution to mitigate the negative effect of both strategies (systematic "collude and report" and "collude and report once"). And in practice penalties for repeat offenders are normally higher than for the first-time offenders. However, as practice shows, the detection probabilities and fine rates for repeat offenders are still far below the level that would deter cartel formation and discovered cartels are tempted to continue collusion even in face of higher expected penalties for repeat offenders.[22]

4.3 The Maximal Cartel Price under Ex-Ante Leniency

In this section, we characterize and illustrate the maximal cartel price for both stationary strategy profiles we considered in the previous two subsections. We conclude by relating our results to the literature.

20 In our setting we can reach the same conclusion, see Section 4.3 of Houba, Motchenkova, and Wen (2009), where leniency is available for a final number of times. Although Houba, Motchenkova, and Wen (2009) report that the maximal cartel price under "report once and never again" strategy can be lower than under "systematic collusion and reporting," the exploitability concerns cannot be eliminated.

21 The stationarity of our model implies that first-time offenders and repeat offenders are treated equally and pay the same fine rate. For an analysis of a richer model, in which fines for repeat offenders increase and monitoring of an industry that is caught at least once intensifies, a proper dynamic game analysis with a state variable should be introduced.

22 Insufficient ceilings on antitrust fines have been analyzed by Bos and Schinkel (2006), Wils (2008), Bageri, Katsoulacos, and Spagnolo (2013), or Katsoulacos and Ulph (2013). In a number of related empirical studies, Connor and Lande (2012) also argue that the existing US and EU penalties for cartel violations are too low resulting in high cartel overcharges.

The cartel can always switch between the equilibrium stationary strategy profiles considered so far. Therefore, the set of sustainable cartel prices is equal to the union of the subintervals of sustainable cartel prices derived in Sections 4.1 and 4.2. Formally, we obtain that $\left[p^N, p^S\right] \cup \left[p^N, p^R\right] = \left[p^N, \max\left\{p^S, p^R\right\}\right]$, where we denote p^L as $\max\left\{p^S, p^R\right\}$. Under the conditions of Proposition 2 and by Proposition 3, we have $p^S = p^A$ and $p^R = p^N$. Then, the following result is immediate and summarizes the maximum cartel price for effective ex-ante leniency programs.

Proposition 4 *Under the most effective ex-ante leniency program and modified grim-trigger strategy profiles, the maximal cartel price is given by* $p^L = \max\left\{p^A, p^N\right\} = p^A$.

This proposition implies that ex-ante leniency cannot deliver any improvement over simple antitrust enforcement in reducing the maximal cartel price. The rationale is given in Section 4.1, where we showed that, for $\alpha(\cdot, N) \geq 0$, the additional constraint imposed by any ex-ante leniency program is redundant. In other words, effective (or properly designed) ex-ante leniency programs that avoid exploitability by systematic collusion and reporting still pose no threats to silent cartels.

Next, we illustrate the maximal cartel price for *arbitrary* ex-ante leniency programs in the (p, α)-space in Figure 4. The question then becomes, by what strategy profile can we sustain the cartel price p (if sustainable) if the reduced-fine rate is $\alpha(p, R)$. For every combination of p and $\alpha(p, R)$ in region C, only the silent cartel strategy profile supports cartel price p. For combinations in region B, only the systematic collusion and reporting strategy profile supports cartel price p. For every combination of cartel prices p and reduced-fine rates $\alpha(p, R)$ in region A, both strategy profiles support cartel price p. Finally, all combinations of p and $\alpha(p, R)$ in region D cannot be sustained by the cartel. More importantly, the boundaries of regions B and C illustrate the maximal cartel price in the (p, α)-space. To see this, if for example we would consider the constant reduced-fine rate $\alpha(p, R) = \bar{\alpha}_R$, the maximal cartel price would be p^A, and for $\bar{\alpha}'_R < \bar{\alpha}_R$ the maximal cartel price is determined by the intersection of the horizontal line $\bar{\alpha}'_R$ with the curve $1 - \frac{1-\delta}{\lambda(p)}$.[23]

The following example combines the previous examples and illustrates Proposition 4.

Example 4 *Recall the homogeneous Bertrand oligopoly model of Example 1 and the linear reduced-fine rates* $\alpha(p, N) = \alpha_N p$ *and* $\alpha(p, R) = \alpha_R p$, *of Example 2 and 3. Taking into account the results of Examples 2 and 3, we have that the maximal cartel price is given by*

23 For $\alpha(p, R) = \bar{\alpha}_R$, the maximal cartel price p^R is characterized by $\lambda^{-1}\left(\frac{1-\delta}{1-\bar{\alpha}_R}\right)$.

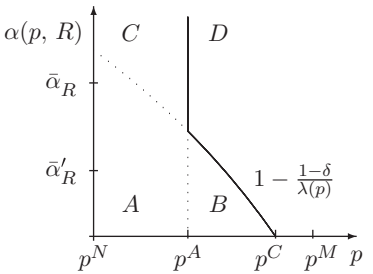

Figure 4: The curve of maximum cartel prices (solid curve) in the (p, α)-space.

$$p^L = \max\left\{\frac{1-n(1-\delta)}{(n\gamma\delta+k)\beta}, \frac{1-n(1-\delta)}{\alpha_R}\right\}.$$

The highest of the two terms in brackets is the solution for the maximal cartel price under ex-ante leniency. Note also that if $\alpha_R \geq (n\gamma\delta+k)\beta$, *then* $p^L = p^S = p^A$. *Otherwise, if* $\alpha_R < (n\gamma\delta+k)\beta$, *then* $p^L = p^R > p^A$, *which can be illustrated in terms of region B in Figure 4. This implies that the ex-ante leniency program may enhance collusion on higher cartel prices when fine reductions in case of multiple reporting are too generous as it also allows to sustain cartel prices in* $(p^A, p^R]$.

To summarize, ineffective or wrongly designed ex-ante leniency programs give too generous fine reductions in case of multiple reporting. We show how introducing ineffective ex-ante leniency programs may improve the stability of higher cartel prices. This adverse effect is represented by region *B*. Previous contributions, which stressed possible adverse effects of leniency programs, mostly analyzed their impact on cartel formation, not on pricing.

A striking observation is that regions *A* and *B* tell a richer story than Proposition 4 in Motta and Polo (2003) that applies for ex-ante leniency programs. In their setting, ex-ante leniency programs only allow for either silent cartels or deterrence of cartels, as the strategy of silent collusion (CNR) dominates the strategy of systematic collusion and reporting (CR).[24] The explanation is twofold. First, cartels maximize profits in Motta and Polo (2003) and then the cartel only has to contemplate either to operate silently or to set the non-cooperative price p^N. For profit-maximizing cartels, the issue of sustainability of CR

24 This dominance may be due to the fact that, in Motta and Polo (2003), reporting forces the cartel to reverse its price to the competitive price p^N for two periods starting in the current one while it still will be punished for its intension to set its cartel price at p. In our setting, the cartel price is set and irreversible at the time the firms report, which makes CR more attractive in our model.

strategies is therefore irrelevant in terms of their setting. Second, if a cartel price cannot be sustained by a CNR strategy in their setting, it cannot be sustained by a CR strategy either. So, their equivalent of our region B is empty. The irrelevance of ex-ante leniency identified in Motta and Polo (2003) is confirmed in regions C and D, while in regions A and B cartel prices are sustainable and can be supported by a CR strategy. Moreover, the maximal cartel price goes above the maximal cartel price under traditional antitrust in region B with ex-ante leniency. These results are different from those of Motta and Polo (2003).

To be more precise, Figure 4 indicates that at the monopoly price p^M there does not exist a CR equilibrium, which confirms the result of Motta and Polo (2003). However, when one considers the entire range of possible cartel prices, then CR equilibria also exist under ex-ante leniency programs in regions A and B. Two other differences between Motta and Polo (2003) and our setting stand out: whether penalties are exogenous and the possibility of substitutability of cartel prices. In Motta and Polo (2003), the reduced (and full) penalty is fixed for all industries and firms decide whether to collude on a particular cartel price, i.e., the monopoly price. In our setting, collusion involves substitutability within a range of sustainable cartel prices and varying such prices implies varying the reduced-fine rate in case of reporting, which also may increase the profitability of the CR strategy for some lower ranges of prices compared to the setting in Motta and Polo (2003).

Finally, in our model, sustaining cartel price $p \in \left[p^N, p^M\right]$ by systematic collusion and reporting is more attractive than a silent collusion in terms of expected cartel profits whenever $\alpha(p,R) < 1 - \frac{1-\delta}{\Lambda(p)}$, which reduces to $\alpha(p,R) < \beta(p)k(p)$ for $\gamma = 0$. This upper bound on $\alpha(p,R)$ extends a similar result in Kaplow and Shavell (1994) that implies that, for static situations and individual violators, a reduced fine for self-reporting below the expected fine of remaining silent induces self-reporting. The condition under which systematic collusion and reporting is more attractive than silent collusion extends the notion of exploitability in Spagnolo (2004). Hence, we obtain that such exploitability not only impacts cartel stability through increasing the value of collusion, it also has the additional effect of increasing the range of sustainable cartel prices, which has not been identified in previous literature.

5 Ex-Post Leniency Programs

In this section, we examine the maximal cartel price sustained by the two types of stationary subgame perfect equilibria under antitrust policy with an ex-post leniency program. Under ex-post leniency, the firms that colluded may

apply for leniency even after AA has launched its investigation. In each period, firms first collude in setting price at $p > p^N$, then the AA launches its investigation with probability $\mu(p)$. After knowing whether they are investigated or not, the firms then decide whether to apply for leniency. If the firms do not apply after an investigation started, their cartel will be convicted with probability $\theta(p)$.

Similar to Section 4, the ex-post leniency program is modeled as a pair of reduced-fine rates $\hat{\alpha}(p, N)$ and $\hat{\alpha}(p, R)$, which have the properties of $\alpha(p, N)$ and $\alpha(p, R)$, respectively. The aim of our analysis is for given $\mu(\cdot)$, $\theta(\cdot)$, and $k(\cdot)$, we analyze what role the reduced-fine rates $\hat{\alpha}(\cdot, \cdot)$ have in combatting cartels. As in Section 4, we focus on two types of stationary subgame perfect equilibria, where the colluding firms either never report or always report to the AA.

5.1 Silent Cartels under Ex-Post Leniency

In this subsection, we focus on a stationary subgame perfect equilibrium where the firms always collude and never apply for leniency. Recall that assumptions about behavior of silent cartels imply that a silent cartel dissolves either deterministically after some firms deviate in setting their price and/or deviate by reporting to the AA, or probabilistically after the cartel is detected and convicted by the AA. According to the stationary strategy profile of the silent cartel, the present value of an individual firm's stream of expected profits is still $V(p)$ by eq. [2]. Under ex-post leniency, any firm has two types of actions to take along the equilibrium path, hence, two incentive constraints. First, eq. [3] continues to ensure that no firm has any incentive to undercut the cartel price p. Second, a firm should have no incentive to report to the AA after it finds out it is investigated. By the one-step deviation principle, these two ICCs imply that a firm would have no incentive to undercut the cartel price and to report to the AA only if being investigated.

According to the second incentive constraint a colluding firm will not report if and only if

$$-\hat{\alpha}(p, N)\pi(p) \leq \theta(p)[-k(p)\pi(p) + (1-\gamma)\delta V(p)] + (1-\theta(p))\delta V(p). \qquad [12]$$

After multiplying both sides with $\mu(p)$, we may rewrite ICC [12] as

$$[1-\mu(p)\hat{\alpha}(p, N)]\pi(p) \leq [1-(1-\mu(p))\delta]V(p).$$

This constraint is nontrivial for $\hat{\alpha}(p, N) < 1/\mu(p)$, and under this condition it can be rewritten as

$$\frac{1-[1-\mu(p)]\delta}{1-\mu(p)\hat{\alpha}(p,N)} \geq \Lambda(p). \quad [13]$$

The left-hand side of eq. [13] is non-decreasing in p and has a well-defined right limit at p^N above $1-\delta$, it decreases with respect to δ and it increases with respect to $\hat{\alpha}(\cdot,N)$ and $\mu(\cdot)$.

With the two incentive constraints [3] and [13] under the ex-post leniency program, the maximal sustainable price of a silent cartel is given by

$$\hat{p}^S = \max_{p\in[p^N,p^M]} p, \quad s.t.\ (3)\ and\ (13). \quad [14]$$

Although eq. [13] may not restrict p to an interval as eq. [3] does, Program [14] is well defined. In contrast to ex-ante leniency, eq. [13] can be more restrictive than eq. [3], in which case ex-post leniency would be effective in lowering the maximal cartel price sustained by a silent cartel.

We now provide the condition under which the maximal cartel price of antitrust enforcement without leniency, p^A, fails [13]:

$$\frac{1-\left[1-\mu(p^A)\right]\delta}{1-\mu(p^A)\hat{\alpha}(p^A,N)} < \Lambda(p^A) = \lambda\left(p^A\right). \quad [15]$$

Then, by continuity of all policy functions on the open interval $\left(p^N,p^M\right)$, we have that there must exist a price $\hat{p}<p^A$ such that each $p\in\left(\hat{p},p^A\right]$ also fails eq. [13].

Decreasing $\hat{\alpha}(\cdot,N)$ would tighten ICC [13]. Hence, the most effective ex-post leniency program in lowering the maximal cartel price sustained by a silent cartel should grant full amnesty to any single-reporting firm: $\hat{\alpha}(p,N)=0$ for all $p\in\left[p^N,p^M\right]$. Then, condition [15], which ensures ex-post leniency is effective in reducing the maximal cartel price, simplifies to

$$\lambda\left(p^A\right)+\left[1-\mu(p^A)\right]\delta>1.$$

In this condition, industry characteristics $\lambda(\cdot)$ and δ are related to the given antitrust policy $\mu(\cdot)$, $\theta(\cdot)$, and $k(\cdot)$, where the latter two only operate implicitly through p^A. To summarize these results, we have

Proposition 5 *Under the ex-post leniency program,*

1. *the maximal cartel price sustained by silent cartels is bounded from above by* p^A;
2. *for silent cartels, the most effective ex-post leniency program is to grant full amnesty to any single-reporting firm, i.e.,* $\hat{\alpha}(p,N)=0$ *for all* $p\in[p^N,p^M]$;
3. *if* $\lambda\left(p^A\right)+\left[1-\mu(p^A)\right]\delta>1$, *under the most effective ex-post leniency program, we have* $\hat{p}^S<p^A$.

This proposition identifies the design of the most effective ex-post leniency program, which should give full amnesty to the single-reporting firm. It also identifies the conditions under which the maximal cartel price sustained by a silent cartel, $\hat{p}^S$, can be reduced below the maximal cartel price under antitrust without leniency, p^A. This effectively implies that ex-post leniency programs can be more successful than ex-ante leniency programs in reducing the maximal cartel price sustained by silent cartels. Recall under ex-ante leniency programs $p^S = p^A$. So, the ex-post leniency program can target the weakest link (destabilizing the silent cartel strategy profile) and reduce the harm compared to the ex-ante leniency program.

Similar to Motta and Polo (2003), the probability of monitoring plays a key role in our results. However, assessing the effects of changes in this parameter is less straightforward than in Motta and Polo (2003), where collusion is possible only at the monopoly price and effectively μ influences only deterrence not the collusive price. In our setting, $\mu(\cdot)$ influences both deterrence and the collusive price. An increase in $\mu(\cdot)$ will raise both sides of eq. [13] and, hence, the overall effect on the maximal cartel price can be ambiguous.[25] Further, an increase of the monitoring probability decreases the maximal cartel price of antitrust enforcement without leniency, p^A, this affects the condition under part 3 of Proposition 5 and makes it more difficult for ex-post leniency programs to improve upon this reduced price p^A. Note also that this effect cannot be observed in Motta and Polo (2003) due to the above-mentioned restrictions.

We now revisit Example 1 under an ex-post leniency program.

Example 5 *Reconsider the homogeneous Bertrand oligopoly model of Example 1. To be consistent with $\beta(p) = \beta p$, we specify $\mu(p) = \mu$ and $\theta(p) = \theta p$ for all $p \in (p^N, p^M] = (0,1]$. The reduced-fine function $\hat{\alpha}(p,N) = 0$ for all $p \in (0,1]$ grants full amnesty. Program [14] becomes*

$$\hat{p}^S = \max_{p\in[0,1]} p, \quad s.t. \frac{1}{n} \geq \frac{1-\delta+\gamma\delta\mu\theta p}{1-k\mu\theta p} \quad \text{and} \quad 1-(1-\mu)\delta \geq \frac{1-\delta+\gamma\delta\mu\theta p}{1-k\mu\theta p}.$$

As in Example 1, the first constraint can be rewritten as $p \leq \frac{1-n(1-\delta)}{(n\gamma\delta+k)\mu\theta} = p^A$. The second constraint can be rewritten as $p \leq \frac{\delta}{[\gamma\delta+(1-[1-\mu]\delta)k]\theta}$. Combining them together, we obtain

$$\min\left\{\frac{1-n(1-\delta)}{(n\gamma\delta+k)\mu\theta}, \frac{\delta}{(\gamma\delta+(1-[1-\mu]\delta)k)\theta}\right\},$$

25 In the example of homogeneous Bertrand below, we identify a price-reducing effect.

where the minimum is the maximal cartel price $\hat{p}^S$ whenever it is smaller than 1, otherwise it is p^M. More important, we have $\hat{p}^S \leq p^A$, confirming Proposition 5. Note that $\hat{p}^S < p^A$ if and only if

$$\frac{1}{n} > 1 - (1-\mu)\delta \Leftrightarrow \delta > \frac{1}{1-\mu} \cdot \frac{n-1}{n}.$$

Furthermore, we investigate the effect of an increase in the monitoring probability μ on the cartel price. An increase in μ will shift both curves for p^A and $\frac{\delta}{(\gamma\delta+(1-[1-\mu]\delta)k)\theta}$ downward in Figure 5. One can see from these expressions that such shift will decrease the maximal cartel price $\hat{p}^S$, because this is the lower envelope of these two price curves. Of course, the range $\delta > \frac{1}{1-\mu} \cdot \frac{n-1}{n}$ for which a price below p^A can be obtained will also shrink. Nevertheless, the overall effect on the maximal cartel price is positive, which can be illustrated in Figure 5.

As evident from Figure 5, introducing ex-post leniency does not affect the range of discount factors for which silent collusion is sustainable. But, the maximal cartel price sustained by silent cartels is reduced from p^A to $\hat{p}^S$ for $1 > \delta > \frac{1}{1-\mu} \cdot \frac{n-1}{n}$. So, even though full deterrence is still not feasible, reducing the maximal cartel price below p^A is possible for industries with a large enough δ.

This also highlights the difference between our example and the results in Motta and Polo (2003), who show that ex-post leniency may destabilize the monopoly price. We extend their results and specify by how much the price is reduced below the cartel price p^A (or p^M) and also for which discount factors this price reduction is achieved.

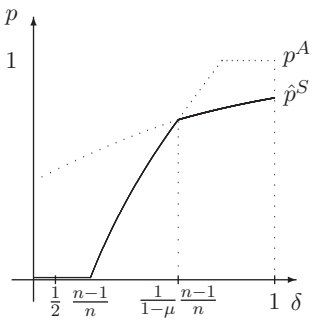

Figure 5: The maximal cartel price of Example 5 is the lower envelope of the depicted curves, where the curve for p^A only becomes positive for $\delta > \frac{n-1}{n}$.

5.2 Systematically Reporting Cartels under Ex-Post Leniency

In this subsection, we look into the stationary equilibrium where the colluding firms systematically report to the AA only after the AA has started its investigation in every period. In such a stationary equilibrium, the cartel will re-establish in every period as long as there is no price deviation. Once the AA starts its investigation, all cartel members will report to the AA and the cartel will be convicted with certainty. Since reporting to the AA is part of the cartel agreement, only the reduced-fine rate when all firms report plays a role in our analysis.

The present value of a firm's expected profit from such a stationary equilibrium, denoted by $V^R(p)$, is determined by the following recursive dynamics:

$$V^R(p) = (1-\mu(p))\left[\pi(p) + \delta V^R(p)\right] + \mu(p)\left[\pi(p) - \hat{\alpha}(p,R)\pi(p) + \delta V^R(p)\right],$$

which gives us

$$V^R(p) = \frac{1-\mu(p)\hat{\alpha}(p,R)}{1-\delta}\pi(p).$$

As before, there are two types of deviations from the cartel agreement an individual firm may take: undercut the price and do not report to the AA. It is easy to see that given all the other firms report, refraining from reporting to the AA is not optimal for any individual firm. To see this, observe that not reporting only reduces a firm's profit in the current period, and it does not change the continuation equilibrium profits. The incentive constraint to ensure no firm will undercut the cartel price is given by

$$\pi^{\text{opt}}(p) + \frac{\delta}{1-\delta}\pi(p^N) = \pi^{\text{opt}}(p) \leq V^R(p).$$

Similar as in Section 4.2, we restrict attention to $\hat{\alpha}(p,R) < 1$ and obtain

$$\lambda(p) \geq \frac{1-\delta}{1-\mu(p)\hat{\alpha}(p,R)}. \qquad [16]$$

Recall that $\lambda(p)$ is non-increasing in p. The properties of $\hat{\alpha}(\cdot,R)$ and $\mu(\cdot)$ carry over to the right-hand side of eq. [16]. So, it is non-decreasing in p and has a well-defined right limit at p^N. Obviously, this fraction is decreasing in δ and increasing in $\mu(\cdot)$ and $\hat{\alpha}(\cdot,R)$.

The maximal cartel price in this case can be formulated as

$$\hat{p}^R = \max_{p\in[p^N,p^M]} p, \quad \text{s.t. (16)} \qquad [17]$$

This program is well defined because $p \in [p^N, p^M]$ and eq. [16] induces a closed interval of $[p^N, p^M]$ that contains p^N. Given the monotonicity of $\lambda(\cdot)$ and $\frac{1-\delta}{1-\mu(\cdot)\hat{\alpha}(\cdot,R)}$, condition [16] alone determines the maximal cartel price under systematic collusion and reporting. At an interior solution $\hat{p}^R \in (p^N, p^M)$, eq. [16] holds with equality.

Given $\mu(\cdot)$, shifting the curve $\hat{\alpha}(\cdot, R)$ upward would tighten ICC [16]. Similarly, given $\hat{\alpha}(\cdot, R)$, an increase in $\mu(\cdot)$ tightens ICC [16] and reduces the maximal cartel price for systematically reporting cartels. Recall that this improvement is not feasible under ex-ante leniency programs as ICC [8] is not influenced by the probability of investigation. Hence, effective ex-post leniency programs that reduce the maximal cartel price of a systematically reporting cartel to p^N require

$$\mu(p)\hat{\alpha}(p,R) > 1 - \frac{1-\delta}{\lambda(p)}, \quad \text{for all } p \in (p^N, p^M].$$

It is evident that any such reduced-fine rate $\hat{\alpha}(\cdot, R)$ can deter the cartel from exploiting leniency by systematic collusion and reporting. All these rates $\hat{\alpha}(\cdot, R)$ implement $\hat{p}^R = p^N$ and, hence, are both price and welfare equivalent. However, the upper bound $\hat{\alpha}(\cdot, R) = k(\cdot)$ would be the easiest to implement in practice as it does not require any additional information about the industry characteristics. Also, by setting $\hat{\alpha}(\cdot, R)$ at the maximal permissible rate, $\hat{\alpha}(\cdot, R) = k(p)$, the AA ensures the effectiveness in reducing the maximal cartel price in as many industries as possible.[26] For these reasons, we call this rate the most effective ex-post leniency program.

Since all feasible curves $\hat{\alpha}(\cdot, R)$ and $\mu(\cdot)$ are non-decreasing and $\lambda(\cdot)$ non-increasing, this condition is met whenever at the right limit

$$\underline{\hat{\alpha}}(R) > \frac{1}{\underline{\mu}}\left[1 - \frac{1-\delta}{\bar{\lambda}}\right], \qquad [18]$$

where $\underline{\hat{\alpha}}(R)$ and $\underline{\mu}$ denote the right limit of $\hat{\alpha}(\cdot, R)$ and $\mu(\cdot)$, respectively. Then, $\hat{p}^R = p^N$. Note that the right-hand side of eq. [18] is non-negative for $\delta \geq 1 - \bar{\lambda}$. We now summarize these results in the following proposition:

Proposition 6 *Under ex-post leniency, the maximal price $\hat{p}^R$ sustained by a systematically reporting cartel has the following properties*:

26 Note that the trivial solution $\hat{\alpha}(p, R) = k(p)$ is inconsistent with both the expected fine reduction interpretation of $\hat{\alpha}(p, R)$ and the orchestrated race to the courthouse interpretation, where $\hat{\alpha}(p, R) = \frac{1}{n}\hat{\alpha}(p, N) + \frac{n-1}{n}k(p) < k(p)$. Then the lower bound on $\mu(p)\hat{\alpha}(p, R)$ requires a sufficiently large $\hat{\alpha}(p, N)$, which might require a positive $\hat{\alpha}(p, N)$, contradicting Proposition 5. We leave further investigation of this issue to future research.

1. *whenever* $\hat{p}^R \in (p^N, p^M)$*, eq. [16] is binding at* $\hat{p}^R$;
2. $\hat{p}^R$ *is non-decreasing in* δ *and non-increasing in* $\mu(\cdot)$ *and* $\hat{\alpha}(\cdot, R)$;
3. *for systematically reporting cartels, the most effective ex-post leniency program sets* $\hat{\alpha}(p, R) = k(p)$ *for all* $p \in \left(p^N, p^M\right]$ *and achieves* $\hat{p}^R = p^N$.

The analysis above shows that properly designed feasible ex-post leniency programs that are not too generous, i.e., such that $k(p) \geq \hat{\alpha}(p, R) > \frac{1}{\mu(p)}\left[1 - \frac{1-\delta}{\lambda(p)}\right]$, can block cartel formation even when cartels may exploit leniency by systematic collusion and reporting. For that the design of leniency guidelines should avoid substantial fine reductions $\hat{\alpha}(\cdot, R)$ in case of multiple reporting, i.e., such fine reductions should be absent or moderate. Furthermore, the AA should set $\hat{\alpha}(\cdot, R)$ at the maximal permissible rate, $\hat{\alpha}(\cdot, R) = k(p)$, to ensure the effectiveness in reducing the maximal cartel price in as many industries as possible when facing the possibility of systematically reporting cartels. The latter resembles the US guidelines. These results are very similar to our results on ex-ante leniency programs. Also similarly to ex-ante leniency programs, it is the class of silent cartel strategies that cause the inefficiency in designing optimal ex-ante leniency programs and not the systematic reporting strategies. Silent collusion cannot be completely deterred by ex-post leniency as it induces $p^N < \hat{p}^S \leq p^A$ for all feasible $\hat{\alpha}(\cdot, N) \geq 0$, while systematic collusion and reporting can be deterred as it induces $\hat{p}^R = p^N$ when fine reductions are properly designed, i.e., $\underline{\hat{\alpha}}(R) > \frac{1}{\underline{\mu}}\left[1 - \frac{1-\delta}{\lambda}\right]$.

We now revisit the Bertrand oligopoly example with linear demand.

Example 6 *Reconsider Example 5 under the ex-post leniency program with linear reduced-fine function* $\hat{\alpha}(p, R) = \hat{\alpha}_R p$, $\hat{\alpha}_R > 0$*, for all* $p \in [p^N, p^M]$*. Recall that we assumed* $\mu(p) = \mu$*. Note that the right-hand side of eq. [18] is positive for all* $p \in [p^N, p^M]$ *when* $\delta > 1 - \frac{1}{n}$*. Then, Program [17] becomes*

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

Proceeding similar as in previous examples, we rewrite the constraint as

$$p \leq \frac{1 - n(1-\delta)}{\mu\hat{\alpha}_R},$$

where the upper bound is equal to the maximal cartel price $\hat{p}^R < p^M = 1$ *whenever* $\hat{\alpha}_R > [1 - n(1-\delta)]/\mu$*. Otherwise,* $\hat{p}^R = p^M$*. In the former case, an increase in the probability of monitoring* μ *reduces the maximal cartel price. Note that such improvement through increasing* μ *is not feasible under the ex-ante leniency of*

Example 3, where β (the combined probability of investigation, μ, and conviction, θ) does not influence the price-setting incentives of systematically reporting cartels.

Note also that in our model, sustaining cartel price $p \in (p^N, p^M]$ with a systematic collusion and reporting strategy profile is more attractive than a silent collusion strategy profile in terms of expected cartel profits whenever $\mu(p)\hat{\alpha}(p,R) < 1 - \frac{1-\delta}{\Lambda(p)}$, which can be interpreted in terms of Kaplow and Shavell (1994). Similar to Section 4.2, this inequality reflects the notion of exploitability as defined in Spagnolo (2004) and we show that it extends easily to ex-post leniency programs.

5.3 The Maximal Cartel Price under Ex-Post Leniency

In this subsection, we characterize the maximal cartel price for both stationary strategy profiles we considered so far in this section. This result will also be graphically illustrated. Finally, we compare both leniency programs and derive policy implications.

The set of sustainable cartel prices is obviously equal to the union of the intervals derived in Sections 5.1 and 5.2. Formally, the set of sustainable cartel prices is given by $\left[p^N, \hat{p}^S\right] \cup \left[p^N, \hat{p}^R\right] = \left[p^N, \max\left\{\hat{p}^S, \hat{p}^R\right\}\right]$. Since $\hat{p}^S < p^A$ by Proposition 5, the following result is immediate and summarizes the maximum cartel price for effective ex-post leniency programs.

Proposition 7 *Under the most effective ex-post leniency program and modified grim-trigger strategy profiles, the maximal cartel price is given by* $\max\left\{\hat{p}^S, \hat{p}^R\right\} = \max\left\{\hat{p}^S, p^N\right\} = \hat{p}^S < p^A$.

This proposition implies that a properly designed most effective ex-post leniency program, i.e., $\hat{\alpha}(\cdot, N) = 0$ and $\hat{\alpha}(\cdot, R) = k(\cdot)$, is more successful than the most effective ex-ante leniency program in reducing the maximal cartel price. It is evident from comparing Propositions 4 and 7 that the maximal cartel price under effective ex-post leniency is smaller than the maximal cartel price under ex-ante leniency, which cannot be reduced below p^A. This implies that the maximal cartel harm is reduced by an effective ex-post leniency program.

Next, as in Section 4.3, we illustrate the maximal cartel price for *arbitrary* ex-post leniency programs in the (p, α)-space in Figure 6 for the case $\hat{p}^S < p^A$.[27] For

27 The case $\hat{p}^S = p^A$ results in a figure similar to Figure 4, which we omit.

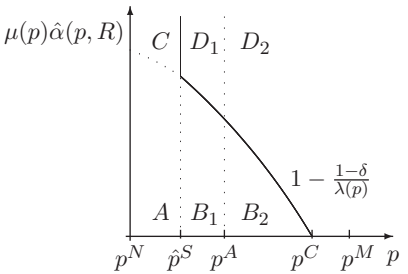

Figure 6: Maximum cartel prices in the (p,α)-space in case $\alpha(\cdot,R)=\mu(\cdot)\alpha(\cdot,R)$.

reasons of comparison with ex-ante leniency programs, we put the expected reduced-fine rate conditional on being inspected on the vertical axis, $\mu(\cdot)\hat{\alpha}(\cdot,R)$. As before, the question is what strategy profile supports cartel price p (if sustainable) if the expected reduced-fine rate is $\mu(p)\hat{\alpha}(p,R)$. This question is answered for a systematic collusion and reporting strategy profile whenever $\mu(p)\hat{\alpha}(p,R)\leq 1-\frac{1-\delta}{\lambda(p)}$, which are regions A, B_1, and B_2 in Figure 6, and similarly for a silent strategy profile for all $p\leq\hat{p}^S$ independent of the expected reduced-fine rate, which are regions A and C. So, for every combination of cartel price p and expected reduced-fine rate $\mu(p)\hat{\alpha}(p,R)$ in region A, both strategy profiles support cartel price p. However, for every combination of cartel price p and $\mu(p)\hat{\alpha}(p,R)$ in regions B_1 and B_2 only the systematic collusion and reporting strategy profile supports cartel price p. The union of B_1 and B_2 can be seen as the region with adverse effects of ex-post leniency in terms of Motta and Polo (2003). Similarly, for region C, only the silent cartel strategy profile supports cartel price p. For completeness, all combinations of p and $\alpha(p,R)$ in regions D_1 and D_2 cannot be sustained by the cartel. The boundaries of regions C, B_1, and B_2 illustrate the maximal cartel price in the (p,α)-space.

Combining Figures 4 and 6 enables to identify the differences between ex-ante and ex-post leniency programs. In case ex-post leniency is effective in reducing $\hat{p}^S$ below p^A, the vertical line at $\hat{p}^S$ under ex-post leniency lies to the left of the vertical line at p^A, which is the lowest possible maximal cartel price under ex-ante leniency. For arbitrary (in)effective leniency programs, we obtain the following result. By having $\alpha(\cdot,R)$ and $\mu(\cdot)\hat{\alpha}(\cdot,R)$ on the vertical axis, it is immediately clear that $p^R=\hat{p}^R$ whenever

$$\mu(p)\hat{\alpha}(p,R)=\alpha(p,R)=1-\frac{1-\delta}{\lambda(p)}\quad\text{for all}\,p\in\left[p^N,p^M\right].$$

Otherwise, it would follow from the combined figure that one of the two leniency programs would allow to sustain a larger maximal cartel price than the other.

So, in order to have the same maximal cartel price under systematic collusion and reporting strategy profile in both leniency regimes, it suffices to impose $\hat{\alpha}(\cdot, R) = \alpha(\cdot, R)/\mu(\cdot)$. Because $0 < \mu(\cdot) < 1$, this suggests that ex-post leniency should be less generous than ex-ante leniency in case of multiple reporting. This supports the current practice in the US and the EU. It also gives a clear-cut novel policy recommendation on how the effectiveness of leniency programs can be improved.

We conclude this section by revisiting the Bertrand oligopoly example with linear demand for the final time.

Example 7 *Reconsider Example 5 under the ex-post leniency program with linear reduced-fine functions. Recall from the previous examples that we have obtained*

$$p^R = \frac{1 - n(1-\delta)}{\alpha_R} \quad \text{and} \quad \hat{p}^R = \frac{1 - n(1-\delta)}{\mu\hat{\alpha}_R}.$$

When we compare the maximal cartel price $\hat{p}^R$ for the case $\hat{p}^R < p^M$ to the maximal cartel price under ex-ante leniency, p^R, derived in Example 3, we conclude that both are equal whenever $\hat{\alpha}_R = \alpha_R/\mu > \alpha_R$. This confirms that ex-post leniency should be less generous than ex-ante leniency and that the ratio between $\hat{\alpha}_R$ and α_R has to be aligned with $1/\mu$.

6 Concluding Remarks

In this paper, the maximal cartel price in infinitely repeated sequential games with two stages is studied as a proxy for both the set of sustainable cartel prices, and consumers' worst-case scenario of maximal damage. We first characterized the maximal cartel price under antitrust enforcement without leniency as a benchmark. In the presence of leniency programs, we analyze the three important decisions firms face: The concerted decision to set the cartel price, the unilateral decision to deviate from the cartel price, and the decision to report to the AA. Given the currently adopted policy rules, 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. Our characterization disentangles the effects of these two cartel strategies as well as the effects of traditional antitrust policies and both ex-ante and ex-post leniency programs.

We provide policy recommendations on how to improve the design of antitrust policy and leniency, how to eliminate adverse effects, and what is necessary to prevent cartel formation in the first place. Our policy recommendations support the current practice of leniency policies and extend the existing theoretical findings with explicit analysis of the impact on prices set by cartels. In particular, we obtain that individual fine reductions in case of multiple-reporting firms should be avoided in order to mitigate exploitability by systematic collusion and reporting in as many industries as possible. Both current ex-ante and ex-post 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. On the other hand, single (or the first-)reporting firm should be granted full immunity. Results of Section 5 support the policies in countries where ex-post leniency applications are allowed and ex-post leniency is designed to be less generous than ex-ante.

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.

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 nor 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.

Acknowledgement: We would like to thank Rene van den Brink, Riccardo Calcagno, Eric van Damme, Arantza Estevez-Fernandez, Dave Furth, Pieter Gautier, Joseph Harrington, Gerard van der Laan, Ines Lindner, Els Kleijn, Michele Polo, Robert Porter, Patrick Rey, Urs Schweizer, Lars Sorgard, Giancarlo Spagnolo, the anonymous referees, and the participants of ACLE (2009), EEA-ESEM (2009), EARIE (2010), and CRESSE (2010) conferences for stimulating discussions and valuable comments. This paper is based on a revised version of TI Discussion Paper 2009-081, "The Effects of Leniency on Maximal Cartel Pricing," by the same authors.

References

Apesteguia, J., M. Dufwenberg, and R. Selten. 2007. "Blowing the Whistle." *Economic Theory* 31 (1):143–66.

Bageri, V., Y. Katsoulacos, and G. Spagnolo. 2013. "The Distortive Effect of Antitrust Fines Based on Revenue." *Economic Journal* 123 (572):545–57.

Bigoni, M., S. Fridolfsson, C. L. Coq, and G. Spagnolo. 2012. "Fines, Leniency, and Rewards in Antitrust." *RAND Journal of Economics* 43:368–90.

Block, M., F. Nold, and G. Sidak. 1981. "The Deterrent Effect of Antitrust Enforcement." *Journal of Political Economy* 89:429–45.

Bos, I., and M. Schinkel. 2006. "On the Scope for the European Commission's 2006 Fining Guidelines under the Legal Maximum Fine." *Journal of Competition Law and Economics* 2 (4):673–82.

Brenner, S. 2009. "An Empirical Study of the European Corporate Leniency Program." *International Journal of Industrial Organization* 27 (6):639–45.

Brenner, S. 2011. "Self-Disclosure at International Cartels." *Journal of International Business Studies* 42 (2):221–34.

Bryant, P., and E. Eckard. 1991. "The Probability of Getting Caught." *Review of Economics and Statistics* 73:531–6.

Buccirossi, P., and G. Spagnolo. 2006. "Leniency Programs and Illegal Transactions." *Journal of Public Economics* 90 (6–7):1281–97.

Chen, J., and J. Harrington. 2007. "The Impact of the Corporate Leniency Program on Cartel Formation and the Cartel Price Path. In *The Political Economy of Antitrust*, edited by V. Ghosal and J. Stennek, p. 59–80. Elsevier.

Chen, Z., and P. Rey. 2013. "On the Design of Leniency Programs." *Journal of Law and Economics* 56:917–57.

Choi, J., and H. Gerlach. 2012. "Global Cartels, Leniency Programs and International Antitrust Cooperation." *International Journal of Industrial Organization* 30 (6):528–40.

Connor, J., and Y. Bolotova. 2006. "Cartel Overcharges: Survey and Meta-Analysis." *International Journal of Industrial Organization* 24 (6):1109–37.

Connor, J., and R. Lande. 2012. "Cartels as Rational Business Strategy: Crime Pays." *Cardozo Law Review* 34:427–90.

Cyrenne, P. 1999. "On Antitrust Enforcement and the Deterrence of Collusive Behaviour." *Review of Industrial Organization* 14:257–72.

D.O.J. 1993. "Us Corporate Leniency Policy." http://www.usdoj.gov/atr/public/guidelines/0091.htm. Accessed April 8, 2015.

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*. Cambridge, MA: MIT Press.

Hamaguchi, Y., T. Kawagoe, and A. Shibata. 2010. "Group Size Effects on Cartel Formation and the Enforcement Power of Leniency Programs." *International Journal of Industrial Organization* 27 (2):145–65.

Harrington, J. 2004. "Cartel Pricing Dynamics in the Presence of an Antitrust Authority." *The Rand Journal of Economics* 35:651–73.

Harrington, J. 2005. "Optimal Cartel Pricing in the Presence of an Antitrust Authority." *International Economic Review* 46:145–70.
Harrington, J. 2008. "Optimal Corporate Leniency Programs." *The Journal of Industrial Economics* LVI (2):215–46.
Harrington, J. 2011. "Corporate Leniency with Private Information: The Push of Prosecution and the Pull of Pre-Emption." January 2011, Working paper.
Harrington, J. 2014. "Penalties and the Deterrence of Unlawful Collusion." *Economics Letters* 124:33–6.
Harrington, J., and J. Chen. 2006. "Cartel Pricing Dynamics with Cost Variability and Endogenous Buyer Detection." *The International Journal of Industrial Organization* 24 (6):1185–212.
Hinloopen, J. 2003. "An Economic Analysis of Leniency Programs in Antitrust Law." *De Economist* 151:415–32.
Hinloopen, J. 2006. "Internal Cartel Stability with Time-Dependent Detection Probabilities." *The International Journal of Industrial Organization* 24 (6):1213–29.
Hinloopen, J., and A. Soetevent. 2008. "Laboratory Evidence on the Effectiveness of Corporate Leniency Programs." *RAND Journal of Economics* 39:607–16.
Houba, H., E. Motchenkova, and Q. Wen. 2009. "The Effects of Leniency on Maximal Cartel Pricing." Tinberegen Institute Discussion paper. 2009–081.
Houba, H., E. Motchenkova, and Q. Wen. 2010. "Antitrust Enforcement with Price-Dependent Fines and Detection Probabilities." *Economics Bulletin* 30 (3):2017–27.
Houba, H., E. Motchenkova, and Q. Wen. 2012. "Competitive Prices as Optimal Cartel Prices." *Economics Letters* 14 (1):39–42.
Jensen, S., and L. Sorgard. 2014. "Fine Schedule with Heterogeneous Cartels: Are the Wrong Cartels Deterred?" Mimeo, Institute for Research in Economics and Business Administration, Norway.
Kaplow, L., and S. Shavell. 1994. "Optimal Law Enforcement with Self-Reporting of Behavior." *Journal of Political Economy* 102 (3):583–606.
Katsoulacos, Y., and D. Ulph. 2013. "Antitrust Penalties and the Implications of Empirical Evidence on Cartel Overcharges." *Economic Journal* 123 (572):558–81.
Lefouili, Y., and C. Roux. 2012. "Leniency Programs for Multimarket Firms: The Effect of Amnesty Plus on Cartel Formation." *International Journal of Industrial Organization* 30 (6):624–40.
Miller, N. 2009. "Strategic Leniency and Cartel Enforcement." *American Economic Review* 99:750–68.
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* 21:347–79.
Rey, P. 2003. "Towards a Theory of Competition Policy. In *Advances in Economics and Econometrics: Theory and Applications*, edited by M. Dewatripont, L. Hansen, and S. Turnovsky, p. 82–132. Cambridge: 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?dpno4840 (revised 2008).
Spagnolo, G. 2008. "Leniency and Whistleblowers in Antitrust. In *Handbook of Antitrust Economics*, edited by P. Buccirossi, p. 259–305. Cambridge, MA: MIT Press.
Wen, Q. 2002. "A Folk Theorem for Repeated Sequential Games." *Review of Economic Studies* 69:493–512.
Wils, W. 2008. *Efficiency and Justice in European Antitrust Enforcement*. Portland, OR: Hart.