

Trust, Leniency, and Deterrence[†]

Maria Bigoni
University of Bologna

Sven-Olof Fridolfsson
Research Institute of Industrial Economics (IFN)

Chloé Le Coq
SITE

Giancarlo Spagnolo
University of Rome Tor Vergata, SITE & EIEF

This article presents results from a laboratory experiment studying the channels through which different law enforcement strategies deter cartel formation. With leniency policies offering immunity to the first reporting party, a high fine is the main determinant of deterrence, having a strong effect even when the probability of exogenous detection is zero. Deterrence appears to be mainly driven by “distrust”; here, the fear of partners deviating and reporting. Absent leniency, the probability of detection and the expected fine matter more, and low fines are exploited to punish defections. The results appear relevant to several other forms of crimes that share cartels’ strategic features, including corruption and financial fraud. (*JEL* C92, D03, K21, K42, L41.)

1. Introduction

When several parties want to pursue a joint illegal activity, they need to trust each other. This article presents experimental evidence that taking this into account is crucial to the optimal design of law enforcement institutions. In particular, we focus on deterrence of collusion among

[†]Many thanks to Tore Ellingsen, Dirk Engelmann, Joe Harrington, Benedikt Herrmann, Magnus Johannesson, Prasad Krishnamurthy, Dorothea Kübler, Massimo Motta, Hans-Theo Normann, Henrik Orzen, Nicola Persico, and Georg Weizsaker for discussions and advice related to this project, and to audiences in Amsterdam (ACLE Conference, 2010), Berlin (WZB/RNIC Conference, 2008), Berkeley (CELS 2014), Bologna, Gothenburg (Second Nordic Workshop in Behavioral and Experimental Economics, 2007), Mannheim (MaCCI Inaugural Conference, 2012), Milan (EEA/ESEM Meeting, 2008), Mölle (SNEE, 2010), Princeton (ALEA Meeting, 2010), Rome (World ESA Meeting, 2007; IAREP/SABE World Meeting, 2008), Stockholm (EARIE 2011 and Konkurrensverket), Tilburg, and Vancouver (IIOC Meeting, 2010). Finally, we would like to thank the Editor Jennifer Reinganum and three anonymous referees for their comments and suggestions. All remaining errors are ours. All views expressed in this article are those of the authors and do not represent those of Konkurrensverket.

oligopolistic firms. Cartels, like most other forms of organized economic crime (corruption, fraud, smuggling, etc.), require effective cooperation between multiple wrongdoers. Being profitable in expectation is therefore not sufficient for them to be viable. Illegal contracts between wrongdoers cannot be enforced, so the collusive agreement must be self-enforcing, sustainable in equilibrium: each wrongdoer must prefer to respect the agreement rather than unilaterally deviate from it by “running away with the money” (Stigler 1964). Moreover, the wrongdoers must coordinate on the cooperative equilibrium and trust that their partners, for the entire period of collaboration and afterwards, will not get “cold feet” and report to law enforcers. A further peculiarity of crimes like these is that there are always “witnesses”: cooperating wrongdoers typically end up having information about each other that could be elicited through suitably designed incentives to report.

These features imply that deterrence can be achieved through a number of channels. While individual crimes must be deterred by large enough expected sanctions (Becker 1968)—so that the individual Participation Constraint (PC) is violated—crimes like cartels can be deterred also:

- by ensuring that at least one co-offender’s Incentive Compatibility Constraint (ICC) is violated, so that the crime—although profitable in expectation—is not an equilibrium; or
- by worsening the “Trust Problem” (TP), that is, each wrongdoer’s fear that his partners will not stick to the illegal agreement, even if it is an equilibrium.¹

This article reports results from an experiment investigating the cartel deterrence effects of changes in the level of fines and in the probability of detection by the competition authority, with and without a leniency policy that offers immunity to the first party to report the cartel. Beyond the novelty and direct policy relevance of the questions themselves, the answers may help shed light on which deterrence channels are more relevant in different legal environments.

We simulate a cartel formation game in the laboratory in which subjects play a repeated duopoly with uncertain end, and can choose whether or not to communicate illegally to fix prices. If players choose to communicate, they are considered to have formed an illegal conspiracy and fall liable to fines.² We adopt the same basic environment developed in

1. Spagnolo (2004) identifies and analyzes this channel theoretically. See also Harrington (2013), where the fear of being betrayed may also induce reporting because cartel members, unlike in our environment, observe private signals on the probability of being detected by a random audit. That criminal organizations require trust to function and pursue their illegal endeavors has also been noted by several legal scholars (see e.g., Leslie 2004; Von Lampe and Johansen 2004 and references therein).

2. The subjects in principle could collude tacitly and thereby reap the gains from cooperation without running the risk of being caught. As discussed below however, our data suggest that they were unable to do so.

Bigoni et al. (2012), which investigates the deterrence and price effects of offering either leniency or rewards to the first party reporting the cartel to law enforcers. Because of constant levels of the fine (F) and the probability of detection through a random audit (α), that study cannot answer the research questions we address here. To try to identify the Trust channel of deterrence, we use data from four benchmark treatments in Bigoni et al. (2012) and run four additional treatments with different levels of α and F . The resulting eight treatments included in the current study differ: (i) in the presence and size of F ; (ii) in the level of α ; (iii) in the possibility and consequences of betraying partners by reporting information to the authority; and (iv) in the possibility to obtain a lenient treatment by reporting.

In line with (our and others') previous experimental studies, we find that schemes granting leniency to the first wrongdoer that spontaneously "turns in" his partners before any investigation is opened, strongly increases deterrence. The novel finding of this work is that the size of the fine *per se* plays a critical role in cartel deterrence, independent of the probability of detection by the competition authority, in particular when leniency is available. In turn, this suggests that the presence of leniency tends to shift the balance between the different deterrence channels in favor of the TP. We find that under standard law enforcement policies, that is, absent leniency, deterrence is more sensitive to changes in the *minimum expected fine* (αF),³ as predicted by classic "Beckerian" law enforcement theory, than in the *actual fine* (F) itself (crime becomes less profitable in expectation and the PC is tightened). With leniency in place, the minimum expected fine also affects deterrence, but the actual fine F becomes even more important in driving behavior. Wrongdoers react less to changes in the probability of detection α when leniency is in place, most likely because they are more worried about the probability of being instead betrayed by fellow cartelists. Most strikingly, we observe a significant direct deterrence effect of F , *even when α equals zero*, that becomes much stronger when leniency is in place. This result stands in open contrast with the classic theory on deterrence of individual crimes (Becker 1968), which asserts that law enforcement produces no deterrence whatever the level of the fine if α equals zero.⁴

Our findings suggest that the possibility of being reported increases the fear and cost of being betrayed: the TP becomes more salient than other considerations. This is especially the case with leniency policies, because they increase the incentive to report. We also find that, in the absence of

3. αF must be lower than the effective expected fine as perceived by the subjects who may also believe that other cartel members may report. Hence, the qualification minimum.

4. Note that more recent models, where subjects may also report in equilibrium to preempt others from doing so before them because of private signals, starting with Harrington (2013), would also predict no deterrence effects with α equal zero. Here and in Spagnolo (2004), instead, a fine has deterrence effects even when α equals zero because it increases strategic risk through the cost of being betrayed. This affects equilibrium selection in favor of a safer equilibrium, not colluding.

leniency, low fines and the possibility to report to law enforcers are used as a costly punishment to discipline cartel deviations; for sufficiently low fines deterrence falls with the fine even if the minimum expected fine does not.

To the extent that these results are confirmed in future studies and apply outside the laboratory, they have important policy implications. They suggest that leniency policies limited to the first party that spontaneously reports, as in the United States before the 1993 reform, could significantly increase the efficiency of law enforcement if coupled with sufficiently robust sanctions. The results also point to the importance of complementing leniency-based revelation schemes with sufficiently severe sanctions rather than with a high probability of detection. The possibility that the recent and numerous leniency applications could undermine cartel deterrence by keeping competition authorities too busy to undertake independent audits [reducing α , see Riley (2007) and Chang and Harrington (2010)] may in turn be less worrisome if sanctions are sufficiently robust and/or can be further strengthened. Our results also suggest that a legal system without leniency could be strategically exploited to punish deviations and stabilize—rather than deter—cartels and similar crimes in jurisdictions where sanctions are relatively low. Finally, the smaller but positive deterrence effect observed in the absence of leniency when α equals zero contradicts our theoretical predictions, and suggests that fines may have a symbolic effect by stressing the illegal nature of cartels. On this aspect there is clearly scope and need for additional theoretical and experimental studies.

Our work contributes to a recent experimental literature evaluating the hard-to-measure deterrence effects of differently designed leniency policies against cartels, which includes Apesteguia et al. (2007), Hamaguchi et al. (2007, 2009), Hinloopen and Soetevent (2008), Krajčová and Ortman (2008), Bigoni et al. (2012), Chowdhury and Wandschneider (2014), and Hinloopen and Onderstal (2014) among others. See Marvão and Spagnolo (2014) for an extensive survey.⁵ These studies are based on the theoretical literature on leniency policies in antitrust, which extends to multi-agent conspiracies Kaplow and Shavell (1994)'s seminal analysis of self-reporting for individual crimes.⁶ To our knowledge, ours is the first

5. Our work is also related to experiments on collusion and oligopoly like Huck *et al.* (1999, 2004), Offerman et al. (2002), Engelmann and Müller (2011), and Potters (2009). A recent experimental study by Schildberg-Hörisch and Strassmair (2012) also deals with deterrence, but focuses on a single decision of a single individual who can steal from another, an environment where the strategic aspects at the core of our paper are not present, while important distributional concerns emerge.

6. See Spagnolo (2004) for a theoretical study close to our experimental set-up. This literature, initiated by Motta and Polo (2003), highlights several possible reasons behind the apparent success of such policies but also some potential counterproductive effects, generating a number of questions hard to address empirically. See Rey (2003) and Spagnolo (2008) for surveys, and Harrington (2013) and Chen and Rey (2013) for recent

experiment to consider different levels of fines and probabilities of apprehension along with leniency policies. It is also novel in trying to disentangle the role of distrust from other possible channels through which law enforcement instruments may deter collaborative crimes.

Furthermore, our work relates to the large experimental literature on trust, surveyed in Fehr (2009). This literature suggests that trust is determined by various factors, including social preferences, fairness, guilt aversion, and beliefs about others' trustworthiness (see e.g., Berg et al. 1995; Fehr and List 2004; Kosfeld et al. 2005; Charness and Dufwenberg 2006; Falk and Kosfeld 2006; Guiso et al. 2008, among others). The concept typically has a positive connotation, since the focus of most studies is on pro-social forms of cooperation (see Gambetta 1988; Knack and Zak 2003). In our context, trust is instead costly for society, and is intended as "trust as beliefs" (Fehr 2009; Sapienza et al. 2013) in that it defines the perceived likelihood that a partner wrongdoer sticks to the criminal plan rather than betrays the conspiracy.⁷

The remainder of the article proceeds as follows: The experimental design and procedures are described in Section 2. Section 3 (and the Appendix) derives theoretical predictions that form the benchmark for our tests. Section 4 reports the results and Section 5 concludes, discussing policy implications and avenues for future research. Supplementary material, containing details about the experimental sessions, our empirical methodology, and the instructions for one of our leniency treatments, complements the article.⁸

2. Experimental Design

The purpose of our experiment is to test the cartel deterrence effects of the fines and of the probability of detection by the competition authority. We adopt the same basic environment developed in Bigoni et al. (2012) and add treatments with new levels of the fine and the probability of detection.

Subjects were randomly matched in pairs, playing in(de-)finitely repeated duopoly games.⁹ Each stage-game consisted of four major steps: communication, price choices, self-reporting, and detection (Figure 1).

contributions. Brenner (2009) and Miller (2009) attempt to empirically identify the deterrence effects of leniency policies by looking at changes in the rate of detected cartels after the introduction of such policies. See also Chang and Harrington (2010), who introduce an alternative methodology based on observed changes in duration of detected cartels.

7. The concept of trust as beliefs in our context is isomorphic to that of "strategic risk", recently developed in Blonski et al. (2011) and Blonski and Spagnolo (2014) to capture the risk of miscoordination in generic repeated games.

8. The Supplementary material is available at *JLEORG* online.

9. A duopoly market contributed to simplifying an already complex strategic environment, hopefully helping the subjects to focus on the variables of interest for our research questions. This design choice involved a risk, however, as tacit collusion is generally easier to achieve in duopolies than in larger experimental markets (Huck et al. 2004). One reason is

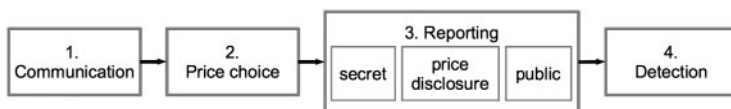


Figure 1. Steps of the Stage Game.

First, subjects could form a cartel by choosing to communicate on prices. If so, they had 30 s to agree on a non-binding price, corresponding to the minimum between the last two prices proposed by the subjects.¹⁰ This restricted protocol, which is similar to the one adopted by Hinloopen and Soeteven (2008), kept the communication phase reasonably short and sped up play, a critical design concern given the length and complexity of our stage game.

Second, subjects had 30 s to simultaneously choose a price in the range $0, \dots, 12$ yielding the payoffs in Table 1; the default price induced zero profit for subjects failing to choose a price. The payoff table, also used in Bigoni *et al.* 2012, is derived from a standard price game with differentiated products and has a unique Bertrand equilibrium, with each firm charging a price of 3 and receiving a profit of 100. The joint profit maximizing price is 9, yielding profits of 180. We adopt differentiated-goods Bertrand competition because we are interested in policy and want to avoid the highly unrealistic discontinuities of the homogeneous-good Bertrand game and the ensuing paradox, where a deviation implies zero profits—and in some previous experiments zero fines—for all other firms.

Third, cartel members were given two opportunities to report the cartels they formed in the current period, or those formed in previous periods that had not yet been detected. A first opportunity to “secretly” report was given right after the price choice, but before this price choice was disclosed to the competitor. Combining a price deviation with such a report constituted an optimal deviation under leniency, as it eliminated the risk of being detected exogenously by the competition authority or through the

probably the “reduced ability to punish a deviant competitor in large-number situations” (Holt 1995: 419). This problem is particularly severe in laboratory experiments, because of the lack of geographical dispersion of simulated markets, and of the possibility that subjects care for other non-defecting subjects and therefore refrain from punishing a defector so as not to harm non-defectors. But since punishments play a central role in our study, this argument also provides a reason for our duopoly choice. And most importantly, the experimental data validate our choice as subjects appeared unable to collude tacitly (see Section 4).

10. Subjects first stated simultaneously a minimum acceptable price in the range $\{0, \dots, 12\}$. When both had stated a price, they chose again a price from the new range $\{p_{\min}, \dots, 12\}$, where p_{\min} equaled the minimum of the two previously chosen prices. This communication continued until time was up. In this way, no subject was forced to collude on a price considered too high. And, they agreed on a minimum acceptable price in a reasonably short lapse of time: in 47.9% of the instances, the two subjects ended up proposing exactly the same price.

Table 1. Profits in the Bertrand Game

	Your competitor's price												
	0	1	2	3	4	5	6	7	8	9	10	11	12
Your price													
0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	29	38	47	56	64	68	68	68	68	68	68	68	68
2	36	53	71	89	107	124	128	128	128	128	128	128	128
3	20	47	73	100	127	153	180	180	180	180	180	180	180
4	0	18	53	89	124	160	196	224	224	224	224	224	224
5	0	0	11	56	100	144	189	233	260	260	260	260	260
6	0	0	0	0	53	107	160	213	267	288	288	288	288
7	0	0	0	0	0	47	109	171	233	296	308	308	308
8	0	0	0	0	0	0	36	107	178	249	320	320	320
9	0	0	0	0	0	0	0	20	100	180	260	324	324
10	0	0	0	0	0	0	0	0	0	89	178	267	320
11	0	0	0	0	0	0	0	0	0	0	73	171	269
12	0	0	0	0	0	0	0	0	0	0	0	53	160

Note: In the Bertrand equilibrium, each firm charges a price of 3 and receives a profit of 100. The joint profit maximizing price is 9, yielding profits of 180.

second reporting opportunity.¹¹ We refer to this second reporting opportunity as “public” because it arose after price choices (and thus possible price deviations) became public information, and was only available if no cartel member had previously reported their cartel secretly. This public report opportunity, if available, could then be used as a punishment against a deviator that did not report.

Finally, cartels that had not yet been reported could be detected and convicted by a competition authority, with probability α . At the end of each period the screen displayed the agreed price, the two players’ price choices, the number of reports, eventual fines, and net profits.

A session lasted for at least 20 periods in total, and potentially included several supergames (or matches). At the end of each stage game there was an 85% probability that the match continued for an additional period, and a 15% probability that the match ended. If the match ended, all pairs of subjects were dismantled and cartels formed in previous periods were

11. The ability to both secretly undercut the collusive price and to self report before the secret price cut becomes public is a crucial feature of leniency programs that sometimes has been overlooked, in theory and in experiments. This ability generates deterrence because it increases incentives to both deviate and report, an effect known as the “protection from fines effect” (Spagnolo 2004). Differently from Hinloopen and Soetevent (2008), in the reporting phase we chose not to account for the order in which subjects push the “report” button if they both report in the same 30-s period, as the two different stages at which a subject can report, a main novel feature of our design, provide a more precise indication of priority in reporting. Previous experiments had only one stage where subjects could report, after prices become public, and therefore could not cleanly distinguish reports linked to the “protection from fines effect” from “punitive” reports induced by subjects reporting to punish the price deviation.

Table 2. Treatments

Treatment	Fine (F)	Probability of detection (α)	Report	Report's effect
Fine	200 ^a	0.10 ^a	Yes	Pay the full fine
	1000	0.02		
	1000	0		
Leniency	200 ^a	0.10 ^a	Yes	No fine (half the fine if both report)
	1000	0.02		
	1000	0		
L-Faire	0 ^a	0 ^a	No	—
NoReport	200 ^a	0.10 ^a	No	—

^aData from these sessions were also used in Bigoni *et al.* (2012).

no longer liable for fines. Subjects were then randomly matched again into pairs and played a new supergame, unless 20 periods or more had passed. At that point the experiment ended.¹² This procedure allowed us to observe the subjects' behavior in several repeated games, and how behavior evolved with experience.¹³

We ran eight treatments (summarized in Table 2) differing in the probability of detection, α , the level of the fine, F ,¹⁴ and in the possibility to report. In all treatments, both cartel members paid the fine F if detected by the competition authority. The effect of a report depended instead on the law enforcement institution. In the FINE treatments, meant to capture traditional law enforcement, the fine F was paid by both cartel members (including the reporting one). In the LENIENCY treatments, if only one member reported the cartel, this member did not pay the fine whereas the other member paid the full fine. Both cartel members paid a reduced fine if instead both reported the cartel simultaneously; this reduced fine equaled $F/2$, corresponding to the expected fine if only the first reporting member was granted (full) immunity and both cartel members were equally likely to report first.

We ran three treatments for each of these law enforcement institutions. Two treatments had the same minimum expected fine, $\alpha F = 20$, the fine being high ($F = 1000$) and the probability of detection low ($\alpha = 0.02$) in one treatment whereas $F = 200$ and $\alpha = 0.1$ in the other treatment. The third treatment had a high fine $F = 1000$ but a 0 probability of detection so

12. To pin down expectations on very long realizations, subjects were also informed that the game would end after 2 h and 30 min. This possibility was unlikely and never occurred.

13. All subjects in a session could in principle compete with each other. Each session therefore constitutes an independent observation. To account for possible dependencies across observations we adopt a parametric approach in the data analysis (see section II in the Supplementary material).

14. We chose to keep the fine F constant *within* each treatment, because we wanted to study the effect of changing the level of the fine *across* treatments, and it is important to be sure that subjects fully understand what this level was.

that the minimum expected fine αF equaled 0. Besides these six treatments we ran a benchmark treatment, L-FAIRE, corresponding to a *laissez-faire* regime where $\alpha = F = 0$. Cartels were thus allowed (but not legally enforced) in L-FAIRE and, to simplify instructions, subjects were not given the opportunity to report.¹⁵ Finally, we ran NoREPORT, a FINE treatment with $\alpha = 0.1$ and $F = 200$ but where cartel members could not report their cartel, neither secretly nor publicly.

A total of 256 students from Tor Vergata University (Rome, Italy) participated voluntarily in this experiment. In each session, 32 subjects interacted anonymously via computer using the software Z-Tree (Fischbacher 2007). Subjects participated in only one treatment and were paid in private at the end of each session. Subjects started with an initial endowment of 1000 points in order to reduce the likelihood of bankruptcy. At the end of the experiment, subjects were paid an amount equal to their cumulated earnings (including the initial endowment) plus a show-up fee of €7. The conversion rate was 200 points for €1. The average payment in the main game was €24.18, with a minimum (maximum) of €11 (€34). Additional information on the experimental sessions is provided in Table I.a of the Supplementary material.

3. Theoretical Predictions and Hypotheses

Our design ensures that forming a cartel by communicating on prices is an equilibrium in all treatments (see Appendix A.1). However, the incentives to participate and to sustain collusion vary under different reporting regimes, as does the cost of being betrayed by a trusted partner.

3.1 Standard Equilibrium Conditions and Deterrence

The PC and the ICC, two necessary conditions for the existence of a collusive equilibrium, provide valuable insights about possible effects of law enforcement institutions. All else equal, the PCs in FINE and LENIENCY treatments are identically tighter than in L-FAIRE due to the expected fine payment. Moreover, the ICCs are tighter in LENIENCY than in FINE or in L-FAIRE, since a deviation in LENIENCY is optimally combined with a secret report providing protection against the fine.¹⁶

The ICCs presume that agents are perfectly able to coordinate on the collusive equilibrium. Even if cooperation constitutes an equilibrium, agents could however be discouraged from forming a cartel by the fear of miscoordination, and even more by the fear of being “cheated” by the opponent. Recent theoretical and experimental work has highlighted that the fear of being cheated and receiving the “sucker’s payoff” constitutes a

15. The instructions for the L-FAIRE treatment did not mention that communication was illegal, unlike the instructions for the FINE and LENIENCY treatments (see the Supplementary material).

16. Appendix A.1 provides a formal analysis underlying these claims. See also Spagnolo (2004) for an in-depth discussion.

critical determinant of subjects' decisions to cooperate (Dreber et al. 2008; Blonski et al. 2011; Dal Bó and Fréchette 2011; Blonski and Spagnolo 2014).

Indeed, we show next, in the spirit of Spagnolo (2004), that the demand for trust required to enter an illegal price-fixing conspiracy varies across law enforcement regimes.¹⁷

3.2 Demand for Trust and Deterrence

Assume that a subject believes that, following communication on the collusive price, his opponent will deviate by undercutting the agreed price with some probability $(1 - \beta)$. The complementary probability β can be viewed as the agent's "belief component of trust" in a partner conspirator (see e.g., Fehr 2009; Sapienza et al. 2013). The *minimum level of trust*, β_K , required to make price-fixing collusion profitable and sustainable in treatment $K \in \{L - \text{Faïre}, \text{Fine}, \text{Len}\}$ can then be viewed as a measure of the "demand for trust" in this treatment. Collusion is then sustainable if $\beta \geq \beta_K$.

Let V_K^{ss} (V_K^{ds}) denote the values of sticking to (deviating from) the collusive agreement in treatment K , assuming the opponent is trustworthy (i.e., sticks to the agreement). Similarly, let V_K^{sd} and V_K^{dd} denote these values, assuming instead that the opponent is not trustworthy (i.e., undercuts the agreed price). Then, β_K is defined by the equality $\beta_K V_K^{ss} + (1 - \beta_K) V_K^{sd} = \beta_K V_K^{ds} + (1 - \beta_K) V_K^{dd}$, or equivalently

$$\beta_K = \frac{(V_K^{dd} - V_K^{sd})}{(V_K^{dd} - V_K^{sd}) + (V_K^{ss} - V_K^{ds})}. \quad (1)$$

β_K is thus determined by two components, $V_K^{ss} - V_K^{ds}$ and $V_K^{dd} - V_K^{sd}$. This measure corresponds to the "basin of attraction" or "resistance" of the cooperative strategy as defined in evolutionary game theory (see Myerson 1991, Section 7.11) and to the measure of strategic "riskiness" of cooperation developed in Blonski and Spagnolo (2014) and Blonski et al. (2011).¹⁸ Presumably subjects are less willing to form cartels when the demand for trust increases. A reasonable conjecture is thus that deterrence increases as β_K increases (as the basin of attraction of cooperation shrinks, or the strategic risk of cooperating increases).

Appendix A.2 provides a formal expression for β_K and characterizes for each treatment the minimum level of trust, showing that $\beta_{L-\text{Faïre}} = \beta_{\text{Fine}} < \beta_{\text{Len}}$. The amount of trust required by the price-fixing

17. An alternative approach capturing deterrence driven by the fear that others report, already mentioned in the introduction and which we did not follow here, would be to allow subjects to have private information on the probability of exogenous detection (α), as in Harrington (2013).

18. Blonski et al. (2011) and Dal Bó and Fréchette (2011) provide experimental evidence in support of the effective predictive power of these measures in repeated games.

conspiracy is thus higher in LENIENCY (but not in FINE) than in L-FAIRE. The reason is that an optimal deviation is combined with a simultaneous report only under LENIENCY, and this increases the cost of being betrayed relative to FINE.

3.3 Hypotheses

Under the assumption that stricter equilibrium conditions make it harder to sustain the equilibrium, deterrence should be stronger in treatments where the PC and ICC are tighter and the demand for trust is higher. This implies that, given α and F , deterrence is lowest in L-FAIRE, followed in order of magnitude by FINE and LENIENCY. The deterrence effect of FINE relative to L-FAIRE is driven only by different PCs. Both the ICC and the minimum level of trust drive the higher deterrence effect of LENIENCY relative to FINE.

To disentangle the effects of the ICCs and the minimum level of trust, we explore the deterrence effects of changes in α and F , taking the policy as given. An increase in the (per period) *minimum expected fine* αF increases the *sum of discounted minimum expected fine payments* (DEF) and thereby tightens the PC for all policy treatments. The effects on the ICC and on β_K , however, depend on whether the policy includes leniency. Absent leniency, the change has no effect, either on the ICC or on β_{FINE} since DEF is the same under FINE whether one, two, or no cartel member undercuts the agreed price. By contrast, the ICC is tightened under LENIENCY, since a deviation combined with a secret report protects against the increase in DEF. For the same reason, β_{Len} also increases. These observations are discussed formally in Appendix A.3 and lead to our first hypothesis.

Hypothesis 1 (increased minimum expected fine): An increase in the per period minimum expected fine

H1_L: increases deterrence under LENIENCY;

H1_F: increases deterrence under FINE but less than under LENIENCY.

To understand if an increase in F per se affects deterrence when a leniency program is present, consider first an increase in F compensated by a fall in α , so as to keep αF constant. Accounting for the possibility of being fined in several periods, such a change tightens the PC slightly in all policy treatments (see Appendix A.4). This dynamic effect is subtle, difficult to compute, and not very intuitive, as DEF increases despite the fact that the per period minimum expected fine αF is constant. One may therefore expect that subjects perceived this as a neutral change. By contrast the change also tightens the ICC in LENIENCY and increases β_{Len} . The effect on β_{Len} may be particularly strong as the increase in F per se lowers the sucker's payoff by increasing the cost of being betrayed (since a defecting subject also reports the cartel, which increases $V_{\text{Len}}^{\text{dd}} - V_{\text{Len}}^{\text{sd}}$). These

observations, of which we give a more formal treatment in Appendix A.4, lead to our second hypothesis.

Hypothesis 2 (constant minimum expected fine): An increase in F compensated by a fall in α so as to keep the per period minimum expected fine constant

H2_L: increases deterrence under LENIENCY;

H2_F: increases deterrence (weakly) under FINE but less than under LENIENCY.

3.4 Main Experimental Hypothesis

Let us now turn to our central experimental hypothesis. The treatments when $\alpha = 0$ (but $F > 0$) are particularly useful in disentangling the different potential deterrence channels discussed so far. The subtle dynamic effect discussed in connection to Hypothesis 2 is not present as both DEF and αF equal zero. According to the PC and the ICC, FINE and LENIENCY should therefore have no deterrence effect relative to L-FAIRE. Instead, the sucker's payoff is much worse under LENIENCY only and therefore increases the demand for trust in that treatment only. This motivates our last hypothesis (see also Appendix A.5).

Main Hypothesis (zero minimum expected fine): With a zero probability of detection, a positive fine generates deterrence under LENIENCY, but not under FINE.

There is an additional reason why this hypothesis is the core of our paper. It stands in sharp contrast with the standard intuition of most previous work in law and economics which, starting from Becker (1968), focuses mostly on individual crime. One of the classic results of this literature is that the first best cannot be achieved even with infinite fines because α must be positive for fines to have any effect and a strictly positive α constitutes a deadweight loss for society (as resources must be devoted to detecting the criminal activity). Our Main Hypothesis therefore marks a stark qualitative difference with this large body of previous work. If supported, it suggests that with organized crime, the first best can instead be achieved with an appropriate combination of well-designed leniency policies and robust (finite) sanctions.

4. Results

Figure 2 provides an overview of how the legal framework affected deterrence. In our analyses, deterrence is measured as the reduction in subjects' decisions to communicate (given that a cartel was not yet formed), that is, in their attempts to form cartels.¹⁹ Its most striking feature is probably

19. The focus of this article is on the channels through which alternative enforcement strategies induce individuals to comply with antitrust law or not. For this reason, as an empirical measure of deterrence, we take the *individual* decision to communicate. If not

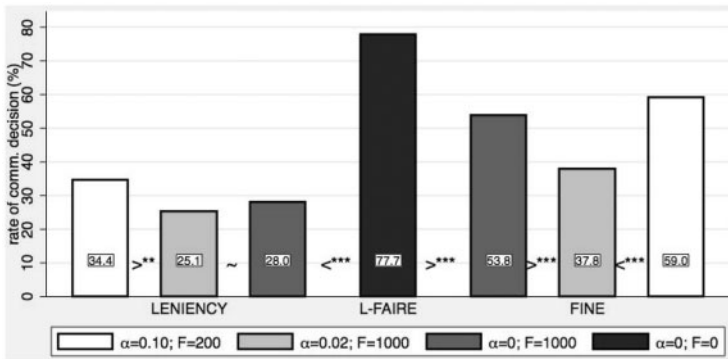


Figure 2. Rates of Communication Decision.

Note: The symbols ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

that the introduction of a positive and substantial fine ($F=1000$) produced considerable deterrence—that is, significantly reduced the rates of communication attempts—even with a probability of detection α equal to 0, particularly when a leniency program was in place.

The introduction of a positive but small probability of detection ($\alpha = 0.02$) increased deterrence primarily in the FINE treatment. A reduction in the fine ($F=200$), compensated by an increase in the probability of detection so as to keep the minimum expected fine constant ($\alpha F = 20$), decreased deterrence both with and without leniency.

The figure also reports statistical tests on the differences in the rates of communication decisions across treatments. These tests are based on *logit* regressions with three-level random effects, to account for potential correlations among observations from the same subject and from the same duopoly. Regression results are reported in Table 3, (a)–(c).

We will return to Figure 2 throughout the discussion of our results. Before doing so, two remarks are warranted. First, the policy effects seem consistent with the theoretical predictions; given α and F , FINE increases deterrence relative to L-FAIRE but less so than LENIENCY.

Second, the subjects appeared unable to collude tacitly. In line with our assumption in Section 3 and consistently with the large experimental literature on communication and coordination (see e.g., Crawford 1998), actual communication was indeed critical for sustaining high prices. In all treatments, prices on average were close to the non-cooperative Bertrand prices absent communication and were systematically larger with communication (see Table 4).²⁰ We now turn to an in-depth discussion of our data

otherwise stated, however, all our results are robust for considering actual communication instead.

20. These price patterns are consistent also with our results on post-conviction prices in Bigoni et al. (2012) where we noted that subjects were unable to sustain high prices after

Table 3. Differences in Deterrence across Treatments

(a) L-Faire versus [$\alpha = 0$; $F = 1000$]		
	Fine	Leniency
$\alpha = 0$; $F = 1000$	-1.042*** (0.295)	-2.927*** (0.332)
Log-likelihood	-567.350	-677.705
N	1108	1350
(b) [$\alpha = 0$; $F = 1000$] versus [$\alpha = 0.02$; $F = 1000$]		
	Fine	Leniency
$\alpha = 0.02$; $F = 1000$	-1.409*** (0.474)	-0.501 (0.401)
Log-likelihood	-406.766	-649.543
N	838	1414
(c) [$\alpha = 0.02$; $F = 1000$] versus [$\alpha = 0.10$; $F = 200$]		
	Fine	Leniency
$\alpha = 0.10$; $F = 200$	1.604*** (0.448)	0.813** (0.371)
Log-likelihood	-511.666	-775.124
N	1010	1552
(d) Antitrust versus leniency		
		$\alpha = 0$; $F = 1000$
Leniency		-2.297*** (0.413)
Log-likelihood		-502.324
N		986

Notes: Results from logit regressions with three-level random effects. Each regression compares pairs of treatments to assess the size and significance of the impact that a change in the size of the actual and expected fine has on deterrence, with and without leniency programs. The dependent variable is the binary decision whether to communicate or not when subjects would not otherwise be liable for collusion. The main independent variable is a dummy taking value 1 for the treatment of interest, and 0 for the reference treatment. Standard errors are robust for heteroschedasticity. See the Supplementary material for further details. The symbols ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

conviction unless they re-formed a cartel by communicating. Note also that the LENIENCY treatments appear to improve welfare relative to the FINE treatments by reducing prices on average. And yet the welfare effects of standard policies—with and without leniency—are ambiguous; prices on average fall in some FINE and LENIENCY treatments relative to L-FAIRE, whereas in others they increase. Explaining these price patterns was at the heart of our previous paper, where we investigated why under both standard antitrust policies and leniency programs the total number of cartels decreases, but existing cartels stabilize and sustain higher prices. In this article we focus mainly on deterrence of cartel formation, and not on cartel effectiveness. These are also more likely to apply to other forms of cooperative crimes (fraud, corruption, and smuggling) than are price effects, which are specific to oligopolies.

Table 4. Average Price With and Without Communication

Treatment	Average price					
	Without communication		With communication		Overall	
	Mean	95% CI	Mean	95% CI	Mean	95% CI
Fine						
$F = 200, \alpha = 0.10$	3.41	[3.25; 3.57]	5.72	[5.47; 5.97]	4.34	[4.18; 4.50]
$F = 1000, \alpha = 0.02$	3.29	[3.20; 3.38]	5.99	[5.71; 6.27]	4.15	[4.01; 4.29]
$F = 1000, \alpha = 0.00$	3.83	[3.67; 3.99]	6.76	[6.52; 7.00]	5.19	[5.02; 5.36]
Leniency						
$F = 200, \alpha = 0.10$	3.54	[3.43; 3.65]	5.73	[5.37; 6.09]	3.93	[3.80; 4.06]
$F = 1000, \alpha = 0.02$	3.48	[3.38; 3.58]	7.22	[6.64; 7.80]	3.77	[3.64; 3.90]
$F = 1000, \alpha = 0.00$	3.85	[3.72; 3.98]	7.93	[7.63; 8.23]	4.70	[4.53; 4.87]
L-Faire	3.25	[3.08; 3.42]	4.87	[4.67; 5.07]	4.27	[4.12; 4.42]
Total	3.53	[3.48; 3.58]	5.99	[5.88; 6.10]	4.32	[4.26; 4.38]

on deterrence. We present first our main result, then discuss the results from the LENIENCY treatments. Finally, we focus on the deterrence effects in the FINE treatments and compare them with those in the LENIENCY (and the NoREPORT) treatments.

4.1 Main Result

Theories of law enforcement that ignore the TP do not predict that leniency programs will deter cartel formation due to the fear of being reported. The reporting behavior in the LENIENCY treatments suggests that these theories may miss an important deterrence channel; the majority of subjects undercutting the agreed price in the LENIENCY treatments did indeed simultaneously report the cartel (see Table 5).²¹

Our main result corroborates the failure of the standard theories in explaining our experimental data, but raises also novel questions. The experimental data are consistent with the first but not the second part of the Main Hypothesis.

Main Result (zero minimum expected fine): *With a zero probability of detection, a positive fine increases deterrence more under LENIENCY than under FINE.*

In line with our theoretical predictions, the experimental data indicate a large drop in the rate of communication decisions in the LENIENCY treatment where $\alpha = 0$ and $F = 1000$ relative to L-FAIRE. The rate of communication decisions decreases by nearly 50% and the effect is highly significant (see Figure 2 and Table 3, (a)). This finding suggests that the

21. In the LENIENCY treatments 64% of the secret reports, and 44% of the “public” ones, were raised simultaneously by both members of the cartel. In the FINE treatments, it never happened that the two cartel members reported the cartel simultaneously.

Table 5. Rate of Reports

Treatment	“Secret” reports ^a		“Public” reports ^b	
	Fine	Leniency	Fine	Leniency
$\alpha = 0; F = 1000$	0.005	1.000	0.125	—
Number of observations	186	20	64	0
$\alpha = 0.02; F = 1000$	0.000	0.538	0.027	1.000
Number of observations	127	13	37	3
$\alpha = 0.10; F = 200$	0.000	0.577	0.197	0.500
Number of observations	211	71	61	10

^aGiven own price deviation.

^bGiven that only the opponent deviated.

threat of being reported in the *LENIENCY* treatment may have undermined trust among the subjects, reducing their willingness to form cartels. In *L-FAIRE*, however, the instructions may have fostered cartel formation by not mentioning fines and the illegal nature of cartels. Such a framing effect could thus explain the observed drop in subjects’ willingness to form a cartel in the *LENIENCY* treatment. However, this effect should have been equally powerful in the *FINE* treatment. As significantly larger deterrence effects emerge in the *LENIENCY* treatment relative to the *FINE* treatment—the rates of communication decisions are 25.8% lower in the former treatment, and the difference is significant at the 1% level (Table 3, (d))—we conclude that the framing effect can explain only part of the difference in cartel formation between *LENIENCY* and *L-FAIRE*. The TP appears to be the most likely explanation for the remaining part.

4.2 More on Deterrence with Leniency

The potentially important policy implication of our main result is that leniency programs may have a non-trivial deterrence effect even when no resources are devoted to crime detection. And, if fines are more efficient in deterring crime than are costly audits and dawn raids, it could be optimal to divert most enforcement resources to other tasks than crime detection (e.g., to the prosecution of reported cartels). To shed more light on this issue, we now investigate the effects of changes in the probability of detection and in the fine on the subjects’ propensity to form cartels.

Result *R1_L*: *An increase in the minimum expected fine does not increase deterrence under *LENIENCY*.*

Our experimental data are not consistent with Hypothesis **H1_L**. The comparison between the *LENIENCY* treatments with the same large fine ($F=1000$) but different probabilities of detection ($\alpha=0$ and $\alpha=0.02$) indicates that an increase in the minimum expected fine through an increase in the probability of detection does not increase deterrence. The rate of communication attempts falls by 2.9% but the effect is not

statistically significant (see Figure 2 and Table 3, (b)).²² The relatively small increase in deterrence induced by the increase in α suggests that the tightening of the PC and the ICC plays a minor role when leniency programs are in place. The worsening of the TP, linked to a large fine F and the risk of being betrayed and reported, seems to be a more important channel through which leniency generates deterrence.

The comparison between the **LENIENCY** treatments with the same minimum expected fine but different mixes of α and F (i.e., the treatments where $(\alpha, F) = (0.02, 1000)$ and $(\alpha, F) = (0.1, 200)$) appears to corroborate this interpretation.

Result R2_L: *Under LENIENCY an increase in the fine F , compensated by a reduction in the detection probability α so as to keep αF constant, significantly increases deterrence.*

Result R2_L is consistent with **Hypothesis H2_L**: the large reduction in the fine from 1000 to 200 substantially reduces deterrence, despite the simultaneous increase in the probability of detection. The rate of communication attempts increases by 9.3% and the effect is significant at the 5% level (see Figure 2 and Table 3, (c)). The fine thus appears to be the most efficient deterrence tool with leniency.

4.3 More on Deterrence under Traditional Law Enforcement

This section focuses on the **FINE** treatments. Our purpose is to identify channels through which deterrence works absent leniency and to contrast these channels with those at work under leniency. Our first finding sheds light on the role of the minimum expected fine in law enforcement environments without leniency.

Result R1_F: *An increase in the minimum expected fine αF induced by an increase in the probability of detection α increases deterrence significantly in FINE; this effect is larger than in LENIENCY.*

Result R1_F is consistent only with the first part of **Hypothesis H1_F**. It reflects the substantial decrease in the rate of communication decisions when we compare the **FINE** treatment where $\alpha = 0.02$ and $F = 1000$ with the same treatment where $\alpha = 0$ and $F = 1000$. The rate of communication decisions decreases by 16% and the effect is highly significant (see Figure 2 and Table 3, (a)). Without leniency, deterrence thus increases with *the minimum expected fine αF* as predicted by classic law enforcement theory (crime becomes less profitable in expectation and the PC is tightened).

Our data, however, are not consistent with the second part of **Hypothesis H1_F**, as the deterrence effect is stronger in **FINE** than in

22. The effect of the introduction of a small probability of detection, however, turns out to be significant when we consider actual communication—rather than the individual decision to communicate—as the independent variable. The rate of actual communication drops from 9.1% to 4%, and the difference is significant at the 1% level according to a two-level random effect logit regression.

LENIENCY (see Result **R1_L**). The increase in the minimum expected fine was driven by an increase in α keeping F constant, suggesting that subjects were more concerned with this change in FINE than in LENIENCY. A possible explanation is that in LENIENCY the risk of being cheated and reported is more salient, limiting the impact of changes in α . This explanation would be consistent with our general idea that leniency programs may be altering the mechanism through which deterrence takes place in favor of the risk of being betrayed, the size of the fine, and the impact on the demand for trust.

Our next finding suggests that low fines may have perverse effects.

Result R2_F: *An increase in the fine F compensated by a reduction in the probability of detection α , so as to keep the minimum expected fine αF constant, significantly increases deterrence in FINE; this effect is larger than in LENIENCY.*

The first part of Result **R2_F** is in line with Hypothesis **H2_F**, and reflects the sharp increase in the rate of communication decisions in the FINE treatment where $\alpha = 0.1$ and $F = 200$ relative to the treatment where $\alpha = 0.02$ and $F = 1000$. The rate of communication decisions increases by 21.2% and the effect is highly significant (see Figure 2 and Table 3, (c)). The second part of the result is instead inconsistent with Hypothesis **H2_F**, which predicted a larger effect in LENIENCY (Result **R2_L**).

The strength of the reduction in deterrence under FINE is also puzzling, because the dynamic effect driving it is subtle. An alternative explanation is that the subjects may have perceived the use of costly reports as a credible threat against deviations only when the fine was moderate and equal to 200. The cost of self-reporting when the fine was high and equal to 1000 relative to the cost imposed by a cheating opponent—at most 160^{23} —may have been perceived as too high for reporting to be considered a credible threat without leniency. A first piece of evidence consistent with this explanation is that public reports were used more often to punish cheating opponents in the FINE treatment with the low fine than in the FINE treatment with the large fine (Table 5).

The NoREPORT treatment used in our companion paper Bigoni et al. (2012), a FINE treatment with $(\alpha, F) = (0.1, 200)$ and where reporting was impossible by design, sheds light on this potential explanation. Removing the possibility to report substantially increased deterrence; communication rates dropped from 59% to 31.8% and the effect was highly significant (see Table 6). This suggests that the possibility to report was indeed perceived as an additional credible (albeit costly) punishment tool against defections when the fine was low enough, which increased cartel formation.

23. This number equals the cost imposed on the cheated party if the monopoly price is the agreed price and the cheating subject undercuts optimally (see Table 1).

Table 6. Deterrence in the “NoReport” Treatment

	$\alpha = 0.1, F = 200$ versus NoReport	$\alpha = 0.02, F = 1000$ versus NoReport
NoReport	-2.146*** (0.336)	-0.682* (0.389)
Log-likelihood	-671.612	-582.433
N	1218	1140

Notes: Results from logit regressions with three-level random effects. The dependent variable is the binary decision whether to communicate or not when subjects would not otherwise be liable for collusion. The main independent variable is a dummy taking value 1 for the treatment of interest, and 0 for the reference treatment. Standard errors are robust for heteroschedasticity. The symbols ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

5. Conclusion

Our laboratory experiment explores how the fine and the probability of exogenous detection by a law enforcement agency affect cartels and analogous cooperative crimes in which wrongdoers are exposed to the risk of being betrayed and denounced by partner wrongdoers. The results suggest that leniency policies restricted to the first party to report spontaneously, without being subject to an investigation, can produce a considerable increase in deterrence. With leniency, deterrence appears to be mainly driven by an increased fear of being betrayed and reported. By increasing both the incentive to betray and the cost of being betrayed by a partner, leniency seems to generate a higher demand for trust among criminals, hence less crime for any given level of trust. Then, a high absolute level of the fine is the most important determinant of deterrence, effective even if there is no risk of detection by the law enforcement agency without a report. Traditional deterrence channels appear more important in the absence of leniency policies: then the probability of detection and the minimum expected fines have a stronger influence on behavior. Low fines are not advisable, however, because besides having limited deterrence effects, these may be used strategically as punishments to stabilize rather than deter cartels.

To the extent that our results apply to real-world settings, they have important policy implications. They suggest that the benefits of tough sanctions may have been underestimated in the case of cartels and similar crimes. When reporting is an option, tough sanctions have a direct effect on deterrence, independent of their well-known effect on the level of the “Beckerian” expected sanction. This effect is particularly strong when well-designed leniency policies make betrayal and self-reporting highly attractive and likely. The smaller deterrence effect we still observe when α equals zero even in the absence of leniency suggests that a framing effect might be in place, as the mere presence of a legal ban on cartels may induce subjects to abstain from collusion even if the ban is not actively enforced. Our results also highlight the high complementarity between modern leniency policies and traditional sanctions. They also suggest that recent concerns that competition authorities are being overburdened by an excess of leniency applications should be addressed by increasing fine levels.

Interesting avenues for future research include robustness checks, such as introducing fines that are functions of accumulated cartel profits; changing the parameterization and the framing of the experiment [although recent work by Krajčová and Ortmann (2008) suggests already that our results should be robust]; understanding whether and how less structured forms of communications influence our findings (e.g., Harrington et al. 2013); and studying whether the structure of criminal organizations reacts and adapts to the introduction of novel law enforcement methods, a possibility suggested by recent theoretical work (e.g., Garoupa 2007; Baccara and Bar-Isaac 2008). It would also be interesting to attempt to identify and quantify experimentally the possible role played by “betrayal aversion” (Bohnet et al. 2008). An additional perceived cost of being betrayed by a peer relative to that of being discovered and fined by a more neutral “law enforcement agency” may indeed have contributed to the strong deterrence effect of leniency in our experiment.

Funding

We gratefully acknowledge research funding from Konkurrensverket (the Swedish Competition Authority) and the Wallander and Hedelius foundation that made this research possible.

Supplementary material

Supplementary material is available at *Journal of Law, Economics, & Organization* online.

Conflict of interest statement. None declared.

Appendix

A.1 Existence of Collusive Equilibria

Our experimental design implements a discounted repeated (uncertain horizon) price game embedded in different antitrust law enforcement institutions. Experimental evidence shows that communication helps subjects coordinating on cooperation (see Crawford 1998). In line with these findings, the simple analysis below presumes communication (i.e., cartel formation) to be a prerequisite for successful cooperation (collusion). Its purpose is to reach sensible testable hypotheses, not to derive the whole equilibrium set.

For simplicity we assume throughout this section that the subjects must communicate once to establish successful collusion, but are able to collude tacitly following a detection by the competition authority.²⁴ Cartel members thus risk to be fined once on the collusive path. Given a per period

24. This assumption implies that the subsequent expressions are relevant mainly for decisions to form cartels given that subjects are not currently members of a successful cartel.

probability of detection α , a fine F , and a discount factor δ (the probability of being re-matched with the same competitor in the next period), the per period minimum expected fine is given by αF and the discounted minimum expected fine payment by $DEF = \alpha F + (1 - \alpha)\delta DEF$, or equivalently

$$DEF = \frac{\alpha F}{1 - (1 - \alpha)\delta}. \quad (\text{E - Fine})$$

A.1.1 The PC. The PC states that the gains from collusion should be larger than the expected cost. Assuming that across periods and treatments, cartels charge the same price on the collusive path, the PCs in L-FAIRE and in the policy treatments can then be expressed as

$$\frac{\pi^c - \pi^b}{1 - \delta} \geq 0 \text{ and } \frac{\pi^c - \pi^b}{1 - \delta} \geq DEF, \quad (\text{PC})$$

where π^b denotes the profits in the competitive Bertrand equilibrium and π^c the profits on the collusive path. Given α and F , the PCs are the same in FINE and LENIENCY treatments, and are tighter in the policy treatments than in L-FAIRE due to the discounted minimum expected fine payment, DEF.

A.1.2 The ICCs. The ICC states that sticking to an agreement is preferred over a unilateral price deviation followed by a punishment. Punishments are assumed to take the standard form of a price war.²⁵ In addition, cartels are assumed not to re-form once they have been dismantled following a price deviation. This assumption implies that the present value in the beginning of the punishment phase (net of potential fine payments), V^p , can be generated by optimal symmetric punishments (given the above stated assumptions). Alternatively, V^p can be viewed as resulting from some weaker form of punishment, which by assumption is the same across treatments.

All else equal, the ICCs can then be expressed as:

$$\frac{\pi^c}{1 - \delta} \geq \pi^d + \delta V^p, \quad (\text{ICC - L - Faire})$$

$$\frac{\pi^c}{1 - \delta} - DEF \geq \pi^d - DEF + \delta V^p, \quad (\text{ICC - Fine})$$

25. We also assume that reports are not used on the punishment path. Public reports as punishments against a price deviation can however be credible in the FINE treatments. In fact, we show in Bigoni et al. (2009), a working paper version of Bigoni et al. (2012), that optimal punishments involve public reports. Subjects nevertheless appear not to use such strong punishments [see Bigoni et al. (2012) for details] and therefore we disregard these when stating our theoretical predictions.

$$\frac{\pi^c}{1 - \delta} - \text{DEF} \geq \pi^d + \delta V^p, \tag{ICC - Leniency}$$

where π^d denotes the deviation profit. Following a deviation, a player risks to be fined in FINE only, since an optimal deviation in LENIENCY is combined with a simultaneous secret report. After reporting the defecting player is protected against the fine, not only because the risk of being detected by the competition authority is eliminated, but also because the competitor cannot use the public report to punish. Note in (ICC-Fine) that DEF appears on both sides of the inequality, since dismantled cartels are assumed not to re-form, either on the collusive path or on the punishment path. Thus, the ICCs are (i) the same in L-FAIRE and FINE treatments and (ii) all else equal, tighter in LENIENCY than in FINE treatments (since a deviation combined with a secret report provides protection against the fine, DEF).

Finally, collusive equilibria exist if the PC and the ICC hold. Note from the PCs and ICCs that a collusive price is sustainable in all treatments if it is sustainable in the LENIENCY treatment with the largest DEF. Thus, let $\alpha = 0.02$ and $F = 1000$ (as in the treatment with the largest DEF) and consider a collusive equilibrium sustained through grim trigger strategies where the collusive price equals 9. The re-matching procedure implies for risk neutral subjects that $\delta = 0.85$. Moreover, $\pi^b = 100$, $\pi^c = 180$ and $\pi^d = 296$. Then, $\text{DEF} = 119.76$ and $V^p = \pi^b / (1 - \delta) = 666.67$ so that both (PC) and (ICC-Leniency) hold with strict inequality. Thus, the joint profit-maximizing price is sustainable in all treatments.

A.2 The Determinants of the Minimum Level of Trust

This appendix offers a formal comparison of the *minimum level of trust* across treatments.²⁶ We assume symmetric punishment strategies. That is, the payoff on the punishment path is given by V^p regardless of whether one or both subjects defect, and is the same for defecting and cheated subjects. We get

$$V_{L-Faire}^{ss} - V_{L-Faire}^{ds} = \frac{\pi^c}{1 - \delta} - (\pi^{d1} + \delta V^p), \tag{A.1}$$

$$V_{Fine}^{ss} - V_{Fine}^{ds} = \frac{\pi^c}{1 - \delta} - \text{DEF} - (\pi^{d1} - \text{DEF} + \delta V^p), \tag{A.2}$$

$$V_{Len}^{ss} - V_{Len}^{ds} = \frac{\pi^c}{1 - \delta} - \text{DEF} - (\pi^{d1} + \delta V^p), \tag{A.3}$$

26. The comparisons between treatments do not depend on the exact deviation strategy considered. It is however important to assume that subjects undercut by the same amount (and attempt to collude on the same price) across treatments.

where π^{d1} denotes the one period payoff from a unilateral price deviation, and

$$V_{L-Faire}^{dd} - V_{L-Faire}^{sd} = \pi^{d2} + \delta V^p - (\pi^s + \delta V^p), \quad (A.4)$$

$$V_{Fine}^{dd} - V_{Fine}^{sd} = \pi^{d2} - DEF + \delta V^p - (\pi^s - DEF + \delta V^p), \quad (A.5)$$

$$V_{Len}^{dd} - V_{Len}^{sd} = \pi^{d2} - \frac{F}{2} + \delta V^p - (\pi^s - F + \delta V^p), \quad (A.6)$$

where π^{d2} denotes the deviation payoff if both players undercut and π^s the “sucker’s payoff” following a unilateral deviation by the opponent. It can be easily verified that $V_{L-Faire}^{ss} - V_{L-Faire}^{ds} = V_{Fine}^{ss} - V_{Fine}^{ds} > V_{Len}^{ss} - V_{Len}^{ds}$ and $V_{L-Faire}^{dd} - V_{L-Faire}^{sd} = V_{Fine}^{dd} - V_{Fine}^{sd} < V_{Len}^{dd} - V_{Len}^{sd}$. Hence, $\beta_{L-Faire} = \beta_{Fine} < \beta_{Len}$.

Note that the ICC (as defined in equation A.1) affects the demand for trust through $V_K^{ss} - V_K^{ds}$: the basin of attraction of sticking to the cooperative strategy expands as the ICC gets looser (since β_K decreases as $V_K^{ss} - V_K^{ds}$ increases). Yet there is also a notable difference between the expressions for $V_K^{ss} - V_K^{ds}$ and the ICCs: π^{d1} replaces π^d in $V_K^{ss} - V_K^{ds}$. This difference stems from the fact that the size of an optimal price deviation must be (weakly) larger if the defecting subject believes that the opponent also undercuts with some positive probability. As a result, the payoff following a unilateral deviation ranges from the payoff resulting from a “safe” Bertrand price (when the opponent chooses the collusive price) and the payoff from an optimal unilateral defection, π^d . Hence $\pi^b < \pi^{d1} \leq \pi^d$ and $\pi^b \leq \pi^{d2} < \pi^{d1}$.²⁷

Note also that β_K increases with $V_K^{dd} - V_K^{sd}$: the basin of attraction of sticking to the cooperative strategy shrinks as $V_K^{dd} - V_K^{sd}$ increases (i.e., since the gains from defecting relative to sticking to the agreement, given that the opponent is not trustworthy, increase).

A.3 Increased Minimum Expected Fine

This appendix motivates Hypothesis 2. αF increases here through an increase in α , keeping F constant, and therefore affects the PC, the ICC, and β_K through its impact on DEF only, not through F . This increase in αF tightens the PC under FINE (since DEF increases) but has no effect on the ICC or on β_{Fine} (as DEF cancels out in (ICC-Fine) as well as in equations (A.2) and (A.5)). Similarly, the increase in αF tightens the PC under LENIENCY. In this case, however, the increase in αF also tightens the

27. The gains from a unilateral deviation are thus (weakly) lower than those indicated by the ICCs, since the defecting subject may find it profitable to undercut the agreed price by a larger amount. Conditional on all other assumptions, however, this fact does not affect the ranking of the ICCs across treatments.

ICC (as DEF increases) and increases β_{Len} (since $V_{Len}^{ss} - V_{Len}^{ds}$ decreases as DEF increases).

A.4 Constant Minimum Expected Fine

This appendix motivates Hypothesis 3. An increased F compensated by a reduced α so as to keep αF constant increases DEF. Therefore, the PC is tightened under FINE while both the ICC and β_{Fine} are unaffected by the change (as DEF does not enter the relevant expressions). Similarly, the change tightens the PC under LENIENCY. In addition, the increase in DEF also tightens the ICC under LENIENCY and thereby also increases β_{Len} . The effect on β_{Len} is exacerbated since $V_{Len}^{dd} - V_{Len}^{sd}$ increases in F .

A.5 Zero Minimum Expected Fine

This appendix motivates Hypothesis 4. Based on the PCs and the ICCs, neither FINE nor LENIENCY should have a deterrence effect relative to L-FAIRE when $\alpha F = 0$. Note also that FINE does not require more trust than L-FAIRE as $\beta_{L-Faire} = \beta_{Fine}$. Therefore, only LENIENCY should have a deterrence effect when $\alpha F = 0$ (and $F > 0$) as it requires more trust in the sense that $\beta_{Fine} < \beta_{Len}$ (since $V_{Fine}^{dd} - V_{Fine}^{sd} < V_{Len}^{dd} - V_{Len}^{sd}$).

A.6 Robustness

Two assumptions underlying the above analysis are worth emphasizing. First, subjects collude tacitly following an exogenous detection on the collusive path and, second, cartels are not re-formed on the punishment path. Provided the cartel is not reported following a deviation (as it is under LENIENCY) the discounted minimum expected fine payment, DEF, is therefore the same on the collusive and the punishment paths.

These assumptions are not innocuous. Suppose that successful collusion requires cartels to be re-formed on the collusive path, even after an exogenous detection by the competition authority. All else equal, this alternative assumption introduces additional deterrence channels. Under FINE, the ICC is tightened (and thereby β_{Fine} also increases) since discounted minimum expected fine payments on the collusive path, given by $\alpha F / (1 - \delta)$, are larger than the corresponding fine payment on the punishment path, DEF. The ICC is also tightened under LENIENCY, as the secret report (associated with a price deviation) provides protection against the larger discounted minimum expected fine payments, $\alpha F / (1 - \delta)$. Most hypotheses nevertheless remain unchanged. The exception is Hypothesis 3 as an increase in F , compensated by a fall in α so as to keep the per period minimum expected fine constant, leaves $\alpha F / (1 - \delta)$ (but not DEF) unchanged. Thereby such a change in the mix of α and F should have a deterrence effect only under LENIENCY, since the increase in F per se worsens the sucker's payoff and thereby increases the demand for trust.

Consider next the assumption that cartels are not re-formed on the punishment path. Presumably it holds if the punishment is carried out through a grim trigger strategy. By contrast, a stick and carrot type of

punishment probably requires cartels to be formed during the “carrot” phase, and possibly also during the “stick” phase. Relaxing the assumption would alter the analysis in two ways. First, it would strengthen the punishment in the policy treatments (though not in L-FAIRE) as subjects run the risk of being fined also on the punishment path. Second, it would affect the scope for punishing defectors, particularly in the LENIENCY treatments because the deviation incentives (from the punishment path) would be magnified by the possibility to report. A formal treatment of these complicating factors is beyond the scope of this experimental paper.

References

- Apestequia, J., M. Dufwenberg, and R. Selten. 2007. “Blowing the Whistle,” 31 *Economic Theory* 143–66.
- Baccara, M., and H. Bar-Isaac. 2008. “How to Organize Crime,” 75 *Review of Economic Studies* 1039–67.
- Becker, Gary S. 1968. “Crime and Punishment: An Economic Approach,” 76 *The Journal of Political Economy* 169–217.
- Berg, J., J. Dickhaut, and K. McCabe. 1995. “Trust, Reciprocity, and Social History,” 10 *Games and Economic Behavior* 122–42.
- Bigoni, M., S.-O. Fridolfsson, C. Le Coq, and G. Spagnolo. 2009. “Fines, Leniency and Rewards in Antitrust: an Experiment,” *CEPR Discussion Papers* 7417.
- . 2012. “Fines, Leniency, and Rewards in Antitrust,” 43 *The RAND Journal of Economics* 368–90.
- Blonski, M., P. Ockenfels, and G. Spagnolo. 2011. “Equilibrium Selection in the Repeated Prisoner’s Dilemma: Axiomatic Approach and Experimental Evidence,” 3 *American Economic Journal: Microeconomics* 164–92.
- Blonski, M., and G. Spagnolo. 2014. “Prisoners’ Other Dilemma,” *International Journal of Game Theory* 1–21.
- Bohnet, I., F. Greig, B. Herrmann, and R. Zeckhauser. 2008. “Betrayal Aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States,” 98 *American Economic Review* 294–310.
- Brenner, S. 2009. “An Empirical Study of the European Corporate Leniency Program,” 27 *International Journal of Industrial Organization* 639–45.
- Chang, M. H., and J. E. Harrington. 2010. *The Impact of a Corporate Leniency Program on Antitrust Enforcement and Cartelization*. Mimeo.
- Charness, G., and M. Dufwenberg. 2006. “Promises and Partnership,” 74 *Econometrica* 1579–601.
- Chen, Z., and P. Rey. 2013. “On the Design of Leniency Programs,” 56 *Journal of Law and Economics* 917–57.
- Chowdhury, S. M., and F. Wandschneider. 2014. *Anti-trust and the Beckerian Proposition: The Effects of Investigation and Fines on Cartels*. Mimeo.
- Crawford, V. 1998. “A Survey of Experiments on Communication via Cheap Talk,” 78 *Journal of Economic Theory* 286–98.
- Dal Bó, P., and G. R. Fréchette. 2011. “The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence,” 101 *American Economic Review* 411–29.
- Dreber, A., D. G. Rand, D. Fudenberg, and M. A. Nowak. 2008. “Winners Don’t Punish,” 452 *Nature* 348–51.
- Engelmann, D., and W. Müller. 2011. “Collusion through Price Ceilings? In Search of a Focal-point Effect,” 79 *Journal of Economic Behavior & Organization* 291–302.
- Falk, A., and M. Kosfeld. 2006. “The Hidden Costs of Control,” 96 *The American Economic Review* 1611–30.

- Fehr, E. 2009. "On the Economics and Biology of Trust," 7 *Journal of the European Economic Association* 235–66.
- Fehr, E., and J. A. List. 2004. "The Hidden Costs and Returns of Incentives—Trust and Trustworthiness Among CEOs," 2 *Journal of the European Economic Association* 743–71.
- Fischbacher, U. 2007. "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," 10 *Experimental Economics* 171–8.
- Gambetta, D. 1998. "Can We Trust Trust?," in D. Gambetta, ed., *Trust: Making and Breaking Cooperative Relations*. 213–37. Oxford: Basil Blackwell.
- Garoupa, N. 2007. "Optimal Law Enforcement and Criminal Organization," 63 *Journal of Economic Behavior & Organization* 461–74.
- Guiso, L., P. Sapienza, and L. Zingales. 2008. "Trusting the Stock Market," 63 *Journal of Finance* 2557–600.
- Hamaguchi, Y., T. Ishikawa, M. Ishimoto, Y. Kimura, and T. Tanno. 2007. "An Experimental Study of Procurement Auctions with Leniency Programs.," *Competition Policy Research Center Discussion Paper Series CRDP-24-E*.
- Hamaguchi, Y., T. Kawagoe, and A. Shibata. 2009. "Group Size Effects on Cartel Formation and the Enforcement Power of Leniency Programs," 27 *International Journal of Industrial Organization* 145–65.
- Harrington, J. E. 2013. "Corporate Leniency Programs when Firms have Private Information: The Push of Prosecution and the Pull of Pre-emption," 61 *The Journal of Industrial Economics* 1–27.
- Harrington, J. E., R. Hernan-Gonzalez, and P. Kujal. 2013. *The Relative Efficacy of Price Announcements and Express Communication for Collusion: Experimental Findings*. Mimeo.
- Hinloopen, J., and A. R. Soetevent. 2008. "Laboratory Evidence on the Effectiveness of Corporate Leniency Programs," 39 *RAND Journal of Economics* 607–16.
- Hinloopen, J., and S. Onderstal. 2014. "Going Once, Going Twice, Reported! Cartel Activity and the Effectiveness of Antitrust Policies in Experimental Auctions," 70 *European Economic Review* 317–36.
- Holt, C. A. 1995. "Industrial Organization: A Survey of Laboratory Research," in J. H. Kagel, and A. E. Roth, eds., *The Handbook of Experimental Economics*. 349–443. Princeton, New Jersey: Princeton University Press.
- Huck, S., H.-T. Normann, and J. Oechssler. 1999. "Learning in Cournot Oligopoly—An Experiment," 109 *Economic Journal* C80–95.
- . 2004. "Two Are Few and Four Are Many: Number Effects in Experimental Oligopolies," 53 *Journal of Economic Behavior & Organization* 435–46.
- Kaplow, L., and S. Shavell. 1994. "Optimal Law Enforcement with Self-Reporting of Behavior," 102 *Journal of Political Economy* 583–606.
- Knack, S., and P. J. Zak. 2003. "Building Trust: Public Policy, Interpersonal Trust, and Economic Development," 10 *Supreme Court Economic Review* 91–107.
- Kosfeld, M., M. Heinrichs, P. J. Zak, U. Fischbacher, and E. Fehr. 2005. "Oxytocin Increases Trust in Humans," 435 *Nature* 673–6.
- Krajčová, J., and A. Ortmann. 2008. "Testing Leniency Programs Experimentally: The Impact of "Natural" Framing.," *CERGE-EI Working Paper*.
- Leslie, C. R. 2004. "Trust, Distrust, and Antitrust," 82 *Texas Law Review* 515–680.
- Marvão, C. M. P., and G. Spagnolo. 2014. *What Do We Know about the Effectiveness of Leniency Policies? A Survey of the Empirical and Experimental Evidence*. Mimeo.
- Miller, N. H. 2009. "Strategic Leniency and Cartel Enforcement," 99 *American Economic Review* 750–68.
- Motta, M., and M. Polo. 2003. "Leniency Programs and Cartel Prosecution," 21 *International Journal of Industrial Organization* 347–79.
- Myerson, R. B. 1991. *Game Theory: Analysis of Conflict*. Cambridge, MA: Harvard University Press.

- Offerman, T., J. Potters, and J. Sonnemans. 2002. "Imitation and Belief Learning in an Oligopoly Experiment," 69 *Review of Economic Studies* 973–97.
- Potters, J. 2009. "Transparency about Past, Present and Future Conduct. Experimental Evidence on the Impact on Competitiveness," in J. Hinloopen, and H.-T. Normann, eds., *Experiments and Competition Policy*. Cambridge University Press.
- Rey, P. 2003. "Toward a Theory of Competition Policy," in M. Dewatripont, L. P. Hansen, and S. J. Turnovsky, eds., *Advances in Economics and Econometrics: Theory and Applications: Eighth World Congress*. Chap. 3. 82–132. Cambridge, UK: Cambridge University Press.
- Riley, A. 2007. "Developing Criminal Cartel Law: Dealing with the Growing Pains," 4 *Competition Law Review* 1–6.
- Sapienza, P., A. Toldra-Simats, and L. Zingales. 2013. "Understanding Trust," 123 *Economic Journal* 1313–32.
- Schildberg-Hörisch, H., and C. Strassmair. 2012. "An Experimental Test of the Deterrence Hypothesis," 28 *Journal of Law, Economics, and Organization* 447–59.
- Spagnolo, G. 2004. "Divide et Impera: Optimal Leniency Programs.," *CEPR Discussion Paper No 4840*.
- . 2008. "Leniency and Whistleblowers in Antitrust," in P. Buccirossi, ed., *Handbook of Antitrust Economics*. 259–304. Cambridge, MA: MIT Press.
- Stigler, G. J. 1964. "A Theory of Oligopoly," 72 *Journal of Political Economy* 44–61.
- Von Lampe, K., and P. O. Johansen. 2004. "Organized Crime and Trust: On the Conceptualization and Empirical Relevance of Trust in the Context of Criminal Networks," 6 *Global Crime* 159–84.