On the Optimal Design of Leniency Programmes

by

Panayiotis Agisilaou

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctorate (Ph.D.) in Economics

University of East Anglia
Norwich, UK

Date of submission: 20 February 2013
Words: 60.117 (WinEdit v.5)

Acknowledgements

Completing my PhD degree in Economics is probably the most challenging and enjoyable activity of my life hitherto. The road to Ithaka was *"long, full of adventure, full of knowledge ... and I feel rich with all I have gained on the way"*. The best and worst moments of my doctoral endeavour have been shared with many people to whom I owe my debt of gratitude.

First and foremost, my supervisors, Bruce Lyons and Morten Hviid, for their wise guidance and unflagging encouragement. With their insightful and instructive comments, they motivate and guide the direction of my research. Certainly, their confidence and enthusiasm facilitated the completion of this project.

I am also indebted to Edward Ball, Subhasish M. Chowdhury, Korina-Kyriaki Cleanthous, Stephen Davies, Peter Dijkstra, Spyros Galanis, Alessandro Iaria, Ioannis Katsoulacos, Constantinos Kouloumas, Michalis Ktori, Demetra Kalli, Maria Kyriacou, Gina Neff, Evgenia Motchenkova, Arnold Polansky, Stefania Sitzia, Andreas Stephan, Scott Summers, Nicolaos Vettas, David Ulph, Frederick Wandschneider, Peter Whelan, Jidong Zhou, Daniel Zizzo, and all the postgraduate students of the School of Economics at the University of East Anglia (UEA). I thank them for their endless patience with me and time spent on my work. They have all been extremely supportive and helpful.

I would like to make a special citation to Robert Sugden and David Rojo Arjona. Robert has taught me advance modelling and introduced me to the methodology of economics. The numerous discussions with David during the coffee breaks provided me with additional confidence and knowledge. I thank both of them for their constant, sincere and multifarious support.

As a member of both, the School of Economics and the Centre for Competition Policy (CCP) at UEA, I have been surrounded by wonderful colleagues. Both communities have provided a warm and prolific research environment. CCP has provided an excellent platform for valuable discussions and training during my post-graduate studies.

My research was funded by a studentship from the Social Sciences Graduate School at the UEA. Financial support from the CCP is also gratefully acknowledged.

Finally, this thesis would not have been possible without the love and support of my wife, Alexandra, to whom I would like to express my deepest and sincerest gratitude. Alexandra has being there for me from the beginning rekindling my dreams and sharing my passions. She has always encouraged and supported my academic aspirations and has surrounded me with unwavering love. Without her all these years would have been colorless and fruitless.

Abstract

This thesis comprises of a collection of essays that aim at enhancing our understanding of the underlying mechanics of leniency policies in antitrust.

In Chapter 1, we provide a systematic overview of the most influential contributions to the literature on collusion and leniency policies, with a focus on antitrust law. The survey elucidates the effects of leniency programmes on cartel formation and cartel implementation.

In Chapter 2, we provide a model to investigate the impact of a leniency programme on collusive firms' incentives to keep or destroy hard incriminating evidence. We show that firms may willfully keep the hard evidence to facilitate the implementation of the cartel. Firms are more inclined to keep the hard evidence when a leniency programme is available. Finally, firms are more likely to destroy the hard evidence when the collusive profits-fine ratio increases.

In Chapter 3, we study the strategic interaction between a cartel and an antitrust authority whose evidence against the cartel is private information. Within the framework of a signalling game, we explore the antitrust authority's incentives to reveal the strength of its evidence, before committing to its prosecutorial effort.
We show that, despite its potentially feeble evidence, the antitrust authority can exploit its informational lead and induce the cartel to self-report at an earlier stage of the prosecutorial process. The more generous the leniency programme, the easier it is to induce self-reporting by the cartel.

In Chapter 4, we provide a model to characterize the optimal leniency programme when colluding firms can invest resources to avoid detection. We show that the optimal fine discount rate depends positively on the severity of the fine and negatively on the probability of investigation and the cost of avoidance activities. A leniency programme that ignores firms' efforts to avoid detection may result in under-deterrence.

Contents

List of Figures

Chapter 1: Collusion in Industrial Economics and Optimally Designed Leniency

1.1 Introduction

The first fundamental theorem of welfare economics crystallizes that a system of competitive equilibria is Pareto optimal. No redistribution of resources could make anyone strictly better-off, without simultaneously making someone else worse-off. Nonetheless, there are situations where free markets fail to efficiently allocate the scarce resources of an economy. One blatant example of market failure is collusion among firms. In his famous book on the '*Wealth of Nations*' Adam Smith[1] acknowledged that the market system inherently provides a fertile ground for the development of orchestrated behaviour by firms.[2] The inimical effects of collusion to the welfare of consumers calls for government intervention; antitrust policy is one such manifestation of this. Its goal is twofold; dissolve existing cartels and deter new cartels to form in the future.

Because of the clandestine nature of cartels, Antitrust Authorities, which are public agencies designated to enforce antitrust law, face two major challenges.[3] The first transpires because of the limited resources they have to combat cartels. The stringency of an antitrust authority's budget unequivocally restrains its investigative competency. The second is due to the asymmetry of information between an antitrust authority and cartel members. Contrary to the antitrust authority, a firm knows whether it participates or not in a cartel. Given the high standard of evidentiary proofs required by courts, an antitrust authority cannot prosecute cartels unless it discovers incriminating evidence.[4] Importantly, reliable and cogent incriminating evidence is not costlessly obtainable.

A powerful policy tool that antitrust authorities have at their disposal in order to overcome the informational disadvantages *vis-à-vis* cartels, is leniency pro-

---

[1] Adam Smith (1723-1790) was a Scottish moral philosopher. He is considered to be the ideological father of the freedom of choice of individuals.

[2] Smith [1776] lucidly observed that "*people of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the public, or in some contrivance to raise prices*".

[3] As Hovenkamp and Leslie [2011] put it: "*Firms that engage in price fixing may try to reduce their probability of antitrust liability in a number of ways. First, members of a price-fixing conspiracy go to great lengths to conceal their illegal activities from antitrust enforcers.*"

[4] Theoretically, the decision of a firm as to whether to participate or not in a cartel is endogenous. In real life, however, things are more complex and less straightforward than in theory. An illuminating study case is *Graphite Electrodes*. For a dicussion of the concerns that surfaced in that particular case see Hviid and Stephan [2009]. Petrinent also is the discussion on whether tacit collusion should be considered as illegal. For a discussion, see Turner [1962], Vaska [1985], Feuerstein [2005], Odudu [2006] and Buccirossi and Spagnolo [2006a].

grammes. Essentially, leniency programmes are information revelation mechanisms aiming to incentivise colluding firms to come forward and denounce their unlawful conduct. More formally, a leniency programme specifies a set of rules for granting reduction of fines (or even offering rewards or bounties), to firms or individuals, for actively collaborating with the antitrust authority, either before or after the commencement of an investigation procedure.[5] Depending on whom the beneficiary is, leniency programmes are distinguished between those addressed to firms and those addressed to individuals.

By incentivising the instability of cartels, such programs can trigger the collapse of existing cartels (*desistance effect*). More importantly, they may render cartels less profitable, and thus less likely to form (*deterrence effect*). However, leniency programmes do not unambiguously reinforce the efficacy of antitrust law enforcement; poorly designed leniency programmes may exacerbate cartel behaviour and can potentially provoke cartel formation.[6] A sound understanding of the functioning of cartels is essential to gain a thorough insight into leniency programmes and their effect on firms' incentives to collude. Prominently, the literature on leniency programmes in antitrust draws insights from the economic literature on generic law enforcement. An overview of this literature can provide useful background knowledge such that the advancements in the literature on leniency programmes in antitrust can be fully realized.

The economic theory of antitrust law enforcement builds on the pioneering work of Becker [1968].[7] Becker was the first to explicitly use rational choice theory to analyze the supply side of crimes. Since Becker [1968], the study of how criminals respond to incentives has flourished. Importantly, Becker's seminal work led to the development of a branch of economics known as the '*Economic Theory of Public Enforcement of Law*'.[8] The analysis of the effects of self-reporting schemes for individual crimes advanced, more than two decades after Becker's contribution, with the ground-breaking research by Malik [1990] and Kaplow and Shavell [1994]. This literature has been the lodestar for a recently developed literature on the optimal design of leniency programmes in antitrust law enforcement. Remarkably, this literature forged 25 years after the implementation of the Corporate Leniency

---

[5]It is worth noting that the provision of lenient treatment (or amnesty) to wrongdoers is not a novelty of antitrust law. Lenient treatment has a long history of use in various fields of law. For instance, bounties against wanted criminals before capture but after detection, tax evasion, unreturned library books, parking tickets, draft evasion and illegal immigration. See Marceau and Mongrain [2000].

[6]For instance, see Ellis and Wislon [2003] and Buccirossi and Spagnolo [2006a].

[7]As Becker [1968, 1993] himself concedes, the discussion has its origins to the representatives of the enlightenment era in the 18th century, and, in particular, to philosophers Beccaria [1986] and Bentham [1931].

[8]For a comprehensive survey of the economic literature on the enforcement of law see Polinsky and Shavell [2000] and Garoupa [1997].

Policy in the US in 1978, with the highly influential papers by Motta and Polo [2003] and Spagnolo [2004].[9]

This paper surveys the literature on collusion and the optimal design of leniency programmes. It also reviews the literature on the economics of law enforcement, with a focus on self-reporting schemes, both for individuals and groups of individuals. The objective is to encapsulate the most influential contributions, highlighting their main findings and limitations, in a systematic and useful and pedagogical way for the reader. Moreover, the paper intends to pin down possible shortcomings and areas that require further research.

In particular, the paper considers: 1) the economic literature on collusion (how cartels form, what challenges they confront, their effects on consumers, the empirical estimation of overcharges, the challenges an antitrust authority encounters in enforcing antitrust law, the policy tools available to combat illegal cartels), 2) the literature on public enforcement of (generic) law, with a focus on the effects of self-reporting schemes, both for individual violators and groups of violators, 3) the literature on antitrust law enforcement, 3) the literature on the optimal design of leniency programmes, with an emphasis put on the impact of such programs on firms' incentives to collude and, 4) the empirical literature on the effectiveness of leniency programmes.

The paper develops as follows: section 2.2 examines the economics of collusion with a focus on cartel formation and cartel sustainability; section 2.3 provides a systematic, yet non-exhaustive, review of the literature on optimal law enforcement; section 2.4 reviews the economic literature on leniency programmes in the realm of antitrust policy and finally; section 4.5 concludes.

1.2 Collusion in Industrial Economics

The overarching priority of antitrust authorities in various jurisdictions across the globe is the detection of existing cartels, and, ultimately, the deterrence of cartel formation in the future.[10] Collusive agreements among firms constitute the most

---

[9]See also the earlier working papers by Motta and Polo [1999] and Spagnolo [2000].

[10]For instance, see OECD, Hard Core Cartels: Third Report on the Implementation of the 1998 Recommendations, 2006; EU Competition Commissioner Mario Monti, Speech/00/295, 11/9/2000; Mario Monti, Fighting cartels: Why and How? Why should we be concerned with cartels and collusive behavior, 3rd Nordic Competition Policy Conference, Stockholm, 11-12/9/2000; Nelly Kroes, Speech/09/375, European Economic and Social Committee, Internal market section, Brussels, 9/9/2009; Report on Competition Policy, 2009, Brussels 3/6/2010, COM(2010) 282; US DoJ Assistant Attorney General R.H. Pate, Speech at the Third annual conference on international and comparative law, London, UK, 16/5/2003; Makan Delrahim, Deputy Assistant Attorney General, Antitrust enforcement priorities and efforts towards international cooperation at the US Department of Justice, 15/11/2004, Taipei, Taiwan.

serious and egregious form of anticompetitive behaviour. Such agreements may adversely affect consumers and the economy as a whole. On account of their pernicious effects to consumers, collusive agreements have been aptly portrayed as a '*major drain*'[11] and '*carcinoma*'[12] on the free market economy.[13] Apart from the tangible malign effects to welfare, such agreements have also intangible and non-measurable effects, i.e., ethical issues. In the words of Smith [1776], almost 230 years ago, a collusive agreement constitutes a conspiracy[14] against consumers.[15]

From an economic point of view, collusion describes a situation where prices are higher than some competitive benchmark.[16] The benchmark is usually considered to be the equilibrium of a strategic scenario where firms meet only once in the marketplace (the stage game equilibrium) and at which collusion is infeasible. Essentially, every subgame perfect equilibrium that is distinct from a Nash equilibrium of the stage game is considered as collusive.[17] The same holds for finitely repeated supergames.[18]

---

[11] OECD Report, Hard Core Cartels, [2001], p. 7.

[12] Mario Monti, Fighting cartels: Why and How? Why should we be concerned with cartels and collusive behavior, 3rd Nordic Competition Policy Conference, Stockholm, 11-12/9/2000.

[13] Some authors argue that collusion, under some conditions, may be beneficial for consumers and/or society as a whole. For instance, see Fershtman and Pakes [2000], Deltas et al. [2009], Doyle and Martijn [2005], Bos and Pot [2012].

[14] The word '*conspiracy*' is used in Section 1 of the Sherman Act: "*Every contract, combination in the form of trust or otherwise, or conspiracy, in restrain of trade or commerce among the several states, or within foreign nations, is declared to be illegal*".

[15] The statement by James Randall (then CEO of ADM - one of the firms participating in the notorious *Lysine cartel*) is enlightening: "*We are gonna get manipulated by these God damn buyers. [...]. They can be smarter than us if we let them be smarter [...] They are not your friend. They are not my friend. And we gotta have 'em. Thank God we gotta have 'em, but they are not my friends. You are my friends. I wanna be closer to your than I am to any customer 'cause you can make us [...] money*".

[16] The enduring temptation of firms to coordinate their conduct in the market is not new for economists. Almost 230 years ago Smith [1776] observed that "*people of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the public, or in some contrivance to raise prices*".

[17] For instance, in a homogeneous good market with price (quantity) competition, a collusive outcome exists whenever prices are higher than the one-shot Bertrand (Cournot) equilibrium.

Motta [2004] contends that in economics collusion coincides with an outcome (sufficiently high prices) and not with the specific form whereby that outcome is attained. See Motta [2004], p. 138.

[18] Supergames are games for which the strategic form is endogenous (the strategic form or the beliefs over the strategic form evolve), i.e. durable goods, inventories, switching costs etc. For these type of games any equilibrium which is distinct from a *Markov Perfect Equilibrium* (MPE) is considered as '*collusive*'.

A MPE is a subgame perfect equilibrium in which all players use Markov strategies. Informally, a Markov strategy depends only on payoff-relevant past events. This type of strategies were introduced by Maskin and Tirole [2001].

A collusive agreement can manifest in various ways. For instance, fixing prices, restricting production, allocating (geographically or otherwise) market shares or consumers, and rigging bids in auctions.

Collusive agreements are classified into two broad categories: explicit and tacit. The former involves an explicit communication between firms and typically takes place in the framework of an organized cartel, while the latter involves collusion in a non-cooperative manner without any explicit communication between firms. For instance, as Rotemberg and Saloner [1990] show, in an oligopoly firms may arrive at a viable coordination without explicit cooperation or agreement by tuning their prices to that of the largest firm in the market. Coordinating to a commonly acceptable collusive equilibrium, when firms tacitly collude, normally postulates costly experimentation in the market. On the contrary, explicitly colluding firms evade this cost. Therefore, the core of the distinction between tacit and explicit collusion is centred on the way whereby an equilibrium is reached.

Regardless of the type of collusion, tacit or explicit, its sustainability depends on the existence of a compliance monitoring mechanism and a mechanism to punish unilateral defections by colluding firms.[19] Importantly, the explicit coordination does not necessarily guarantee the sustainability of the collusion, since firms communication might simply be 'cheap talk'.[20] It has to be remarked though that the motive to deviate from the collusive equilibrium exists both with explicit and tacit collusion. Hence, the underlying economics is the same for the two types of collusion (tacit and explicit). The crucial point with either type of collusion however, is that the collusive agreement must be self-enforcing in order to be sustainable.

The (tacitly or implicitly) coordinating firms will ideally aim to reproduce, in an artificial manner, the outcome of a monopoly. However, this does not necessarily mean that the welfare implications of a cartel are identical to that of a monopoly. Such a statement would erroneously imply that a cartel is equivalent to a merger that results in a monopoly. Indeed, it is totally different to have one or more firms with market power as a result of competition on the merits (i.e., successful product or process innovation and/or exploitation of economies of scale), and firms with market power gained as a result of a conspiracy to impair and contravene free market competition. The subversion of the competitive process may result, among others, in the survival of productively less efficient firms in the market[21] or the

---

[19]Stigler [1964] was the first to argue that cartels are intrinsically unstable because of the individual cartel member's incentive to benefit from defecting from the agreement by means of secret price cuts. Therefore, apart from being profitable, a cartel must also be feasible. The latter requires a mechanism to monitor firms' compliance to the agreement and a credible punishment mechanism for defecting firms.

[20]In Game Theory 'cheap talk' is the communication between players which does not directly influence the payoff of the game. See Farrell and Rabin [1996] and Crawford and Sobel [1982].

[21]Competition, among others, can be described as a selection process whereby only the most

existence of unexploited economies of scale. Moreover, it can result in less product variety, and thus freedom of choice, both for firms and consumers.

Taking the above into consideration, a starting point to describe the welfare implications of a cartel is the standard microeconomics textbook analysis of a monopoly.[22] To begin with, a cartel causes allocative inefficiency. The consumers that continue to purchase the cartel product see the margin between their maximum willingness to pay and the price they actually pay squeezed. As a consequence their welfare diminishes. Although transferred from consumers to cartel members, it is claimed that these resources should be considered as a deadweight welfare loss. The heart of the argument is that those resources are wasted by cartel members in rent seeking activities, which do not generate, either directly or indirectly, any value for society.[23] On the other hand, there are also consumers who stop purchasing the product because the collusive price is greater than their maximum willingness to pay. These consumers divert to other products to which they attribute a lower value. Thus, given that the prices cease to send the right signals to economic agents (consumers and firms), the market fails to allocate resources to their highest and best use (*allocative inefficiency*).

Moreover, the lack of competitive pressure, as an outcome of the formation of a cartel, ultimately results in organizational and/or managerial slack.[24] Protected from the competitive forces and having secured supra-normal profits, colluding firms may fail to produce at the lowest possible cost (*'x-inefficiency'*).[25] Finally, firms may experience less pressure to improve their competitive position in the market by investing in new products and/or new and more advanced production processes. As a result, the dynamic efficiency may also be unfavorably affected by the existence of collusion.

efficient firm(s) survive in the market against the less productive and efficient firms (Darwinian *'survival-of-the-fittest'*). For more on *'concepts of competition'* see Vickers [1995].

[22]For an elegant exposition of the welfare effects of market power see the trade-off model by Williamson [1968].

[23]These resources can be used by cartel members to secure or even to increase their market power. For instance, lobbying and bribing, detection avoidance activities, prosecution litigation, and developing (and maintaining) mechanisms to administer and monitor the compliance of cartel members to the collusive agreement.

[24]Adam Smith maintains that "*Monopoly ... is the greatest enemy of good management, which can never be universally established but in consequence of the free and universal competition which forces every body to have recourse to it [i.e. good management] for the sake of self-defence*". See Smith [1776], pp. 163-164.

[25]The term *'x-inefficiency'* has its origin to Leibenstein [1966]. Hicks [1935] vividly declared that "*the best of all monopoly profits is quiet life*".

1.2.1 Empirical estimation of cartels overcharges

The empirical assessment of the welfare effects of cartels exclusively focuses on the quantitative estimation of overcharges to consumers. It has to be emphasized that this method may not always lead to an accurate estimation of the total harm caused to society by cartels. If, for example, the demand is insensitive to price changes, then an increase in the price may lead to a transfer from consumers to cartel members, which from a social welfare point of view is not inefficient. Another remark is that the estimation of the overcharges is based on data from detected cartels. Thus, if the most successful cartels are those who remain undetected, then the empirical estimation may underrate the overcharges of the total population of cartels.

The relevant literature comprises of two different methodologies:[26] the first estimates the overcharge of each individual cartel; the second uses a meta-analysis[27] to estimate the average overcharge for various cartels in different time periods and across a range of geographical areas. Regardless of their differences, the existing studies conclude that many cartels were effective in achieving significant overcharges.[28] Clearly, some cartels were more successful than others at raising prices.[29]

For instance, according to Connor [2001] US citizens would have paid \$65-134 million less in the absence of the notorious *Lysine cartel* (1992-1995). Connor and Lande [2008] also estimate the actual overcharge[30] of the *vitamins cartel* (1990-1999) to \$11.5 billion globally and \$3.6 billion for the US economy. For the *graphite*

---

[26]Although there are several studies that investigate and estimate the welfare loss of monopolies (for instance, see Harberger [1954], Schwartzman [1960], Worcester [1973], Kamerschen [1966] and Jenny and Weber [1983], Cowling and Mueller [1978]), the estimation of the corresponding welfare loss due to the existence of cartels is not as straightforward. This is mainly ascribed to the ignorance of the number of cartels that actually exist in markets. What could be potentially estimated empirically is the deadweight welfare loss caused by the detected cartels. This typically requires the estimation of the Hicksian demand functions.

[27]A *meta-analysis* is a statistical technique for combining the findings from independent studies.

[28]The differences of these papers are mainly the sample selection method, the approach that is used to estimate the overcharge, and the specification of the econometric model employed.

[29]The empirical literature concentrates on the welfare effects of detected cartels. Apparently, there is a data selection bias, as we do not have information with regard to the non-detected cartels. Therefore, measuring the welfare effects of all existing cartels (detected and undetected) is a rather complex, if not impractical, task. This is a semantic difference between the empirical literature on cartels and mergers. Unlike cartels, potentially merged firms are obliged to notify to the antitrust authority their plans to merge, and, therefore, data for all proposed mergers exists (approved and blocked). This allows a more thorough estimation of the welfare effects of mergers, compared to cartels.

[30]The actual overcharge takes also into consideration the overcharge of other firms that supply the same product but do not participate in the cartel.

electrode cartel (1992-1997), the overcharge is estimated to 50% in the EU and 65% in the US.[31] Moreover, the overcharge for the *citric acid cartel* is estimated to 30%, both for US and globally, which corresponds to $1.5 billion.[32]

Other studies focus on the estimation of overcharges in bid rigging auctions and other procurement contests. For instance, Porter and Zona [1999] estimate the overcharge in the procurement of milk to southwest Ohio schools to be, on average, 6.5%, however, in many areas the overcharges soar to 49%. Froeb et al. [1993] estimate the average overcharge paid by the US Defence Department for the procurement of frozen perch to be in the range of 23-30%. Kwoka [1999] estimates the undercharge[33] to the real estate auctions in Washington DC. to 32%. Similarly, Howard and Kaserman [2004] estimate the overcharge for the public procurements of sewer construction projects to be approximately 36%.

As has already been alluded to, the relevant literature encompasses numerous meta-analysis surveys. In one of his several studies, Connor [2005] uses a comprehensive sample, which includes around 1000 cartels, to estimate an average overcharge of 25%. The same study estimates the average overcharge for US to 19%, while for international cartels the average overcharge is significantly higher and amounts to 32%. Connor and Lande [2008] estimate the average overcharge to 27.8%, which corresponds to $711 million, while the median[34] overcharge is estimated to 27%, which corresponds to $63.5 million. On an individual cartel level, the range of the overcharge is $0-14,256 million, while the overcharge as a percentage of the sales ranges from 0 to 79.5%. Similarly, Bolotova [2009] estimates the average overcharge for all cartels in the period 1770-2006 (legal, illegal, US and international) to 21.88%, while the median overcharge amounts to 20%. On an individual cartel level, the range of the overcharge is -5.3-81.8%.[35] By using a subsample of 157 modern international cartels in the period 1991-2005, covering more than 80 different markets, the same study estimates the overcharge to 27.81% and the median overcharge to 28%. The range of the overcharge on individual cartel

---

[31] OECD, Report, Hard Core Cartels, 2000, OECD, Report - Fighting Hard Core Cartels: Harm, Effective Sanctions and Leniency Programmes, 2002 and OECD, Hard Core Cartels – Recent progress and challenges ahead, 2003.

[32] OECD, Report, Hard Core Cartels, 2000.

[33] This is because bidders collude to lower, rather than to increase, their bids.

[34] In general, the statistical average can severely influenced by the existence of extreme values (outliers) in the sample and the selection of a non-representative set of observations. When the distribution exhibits asymmetry, it is more appropriate to use a measure which is based on the median of the sample of observations. The median can estimate with greater accuracy the central tendency of the distribution rather than the average.

For the empirical estimation of cartel overcharges, the statistical average can be sometimes misleading, especially when there is a small number of cartels that were extremely successful in raising prices. That is why, by convention, both, the average and median, are calculated.

[35] Evidently, some cartels were not successful at setting positive overcharges for their products.

level is 0-90.9%.

More recently, Boyer and Kotchoni [2011] conduct a refined meta-analysis and show that the overcharges are significantly lower than previously estimated in the literature. In particular, the authors estimate, after correcting biases in the estimation method and the publication source of data, an average overcharge of 17.5%, with median overcharge of 14.1%.[36]

Functioning of cartels   Before exploring the effects of leniency programmes, it may be worthwhile to analyze the microfoundations of cartels (i.e., comprehend how a cartel operates and what constraints, intrinsic or extrinsic, it copes with). For this the focus initially is to delineate the underlying rationale for firms' participation to cartels. Subsequently, the spotlight is turned to the factors that potentially facilitate or hamper cartels' sustainability.

According to economic theory, the restrictions that a cartel encounters boil down to two fundamental equations: the *participation constraint* and the *incentive compatibility constraint.* The former determines the incentives to form or to join a cartel. A profit maximizing firm will join a cartel if the (expected) benefits exceed the (expected) costs associated with its participation to the cartel. The benefit is the additional (supra-competitive) profits due to the coordinated behavior, while the expected cost is the fine that a firm has to pay in case of detection and conviction by an antitrust authority.[37] Noticeably, the level of the expected cost of the cartel, contrary to the expected benefit, is influenced by the level of the fine and the probability of detection; two policy instruments. Nonetheless, a cartel's sustainability in the long-run is precarious because of the innate temptation of its members to defect from the illegal agreement. This will be the case if the deviation profits exceed the loss entailed from the ensuing retaliatory measures by other cartel members and the concomitant deprivation of future supra-normal profits. The relationship between the compliance and non-compliance profits is encapsulated by the incentive compatibility constraint.

Correspondingly, the economic literature on collusion consists of two strands. The first deals with the procedures of cartel formation and examines cartel stability. It uses the static game theory approach, and thus by its very nature leaves unexplained why some cartels are more successful than others (i.e., last for a longer period of time). In other words, it concerns explicit rather than tacit collusion. No-

---

[36]For other empirical studies estimating cartels overcharges see, Werdem and Simon [1987], Griffin [1989], Cohen and Scheffman [1989], Gallo et al. [2000], Connor [2001], Posner [2001], Connor [2003], Werden [2003], Connor and Bolotova [2006], Levenstein and Suslow [2006], Connor and Lande [2008], Bolotova et al. [2007] and Miller and Connor [2010].

[37]Realistically, the cost of a cartel also takes into account the potential 1) private damages actions, 2) individual sanctions (e.g. risk of criminal proceedings, disqualification of directors, personal damages) and 3) reputational damages because of adverse publicity.

tably, this branch of the literature is closely related to the one on non-cooperative games of coalition formation with payoff externalities.[38] The second focuses on the enforcement of the collusive behavior. In particular, it uses the supergame-theoretic approach to explore the factors that reinforce or hinder the sustainability of cartels. Furthermore, it investigates the factors that facilitate the adoption of a commonly acceptable collusive equilibrium. Clearly, this literature concentrates on firms' incentive compatibility constraints, leaving unexplained the participation constraints.

1.2.2 Cartel Formation

As mentioned above, the heart of the literature on cartel formation is based around the stability of cartels.[39] A cartel is said to be stable if it is individually rational for every firm participating in the cartel to remain in the cartel (internal stability) and for every firm not participating in the cartel to remain outside the cartel (external stability).[40]

A central remark of this literature is that a cartel is basically a public good. A firm that decides to stay an outsider cannot be excluded from the positive externality that is generated by the coordinated conduct of cartel members.[41] The public good properties of a cartel give rise to a free-riding problem. Invariably, the opportunistic behaviour undermines a cartel's fragile stability. However, this line of argument disregards the effect on the equilibrium price, and consequently the impact on firms profits, after a change in firms membership of the cartel.

The process of cartel formation was originally analyzed by Selten [1973]. The main conclusion of the study is that under Cournot competition, the process of forming a cartel in a market that comprises of more than five firms is quite difficult and most probably will be futile. With more than five rivals, it becomes more advantageous to stay out of the cartel formed by others. This is because the position of an outsider becomes more attractive as the number of competitors

---

[38]See Bloch [1996, 2005] and Yi [1997].

[39]A crucial assumption of this literature is that cartels are sustained with binding agreements or enforceable contracts. This implicitly entails that collusive firms' conduct does not change when a defection occurs in the market.

[40]The concepts of internal and external stability were originally introduced by d'Aspremont et al. [1983].

[41]The intrinsic difficulty in convincing firms to form a cartel was originally pointed out by Stigler [1950, 1983] in a discussion of horizontal mergers. Stigler [1950] maintained that "*If there are relatively few firms in the industry, the major difficulty in forming a merger is that it is more profitable to be outside a merger that to be a participant. The outsider sells at the same price but at much larger output at which marginal cost equals price.*" Stigler [1950] p. 25. It is noteworthy to remark the resemblance between firms' incentives to merge and form a cartel. In both scenarios firms have similar incentives to free ride. Thus, in principle, the two literatures, on cartel formation and mergers, are alike.

increases.[42] More precisely, the numerical analysis by Selten [1973] demonstrates that if there are four or less rivals, then a cartel forms with certainty. If there are five, the probability that all five collude is 0.22. If there are more than five rivals, this probability drops to approximately 0.01.[43]

Contrary to Selten [1973], who considers the cartel as another firm in a Cournot-competition game, d'Aspremont et al. [1983] focus on the formation of a dominant cartel. A dominant cartel acts as a price leader, whereas all firms outside the cartel remain price-takers. In a model of collusive price leadership with increasing marginal cost, d'Aspremont et al. [1983] show that there is always a stable cartel, provided that a finite number of firms exists. Moreover, the authors show that the fraction of the firms that participate in a stable cartel diminishes with the number of firms that are active in the market.[44] This is because, the smaller the size of the cartel (in relation to market), the poorer the effect of the marginal collusive firm on the equilibrium price. Similar conclusions apply when the strategic decision of firms is at the level of production (and not the price), as Donsimoni et al. [1986] show.

In another study, Donsimoni [1985] examines the effects of the existence of asymmetry in costs, assuming increasing marginal costs. The main result of the study is that heterogeneous stable cartels, i.e., cartels with different types of firms, are feasible. Stable cartels comprise of the most efficient firms, while the less efficient firms remain independent outsiders.

Shaffer [1995] demonstrates that when the strategic variable is the level of production, and when there is a Cournot fringe, the stable cartel comprises of 50% of the firms, regardless of the total number of firms in the market.[45] However, Escrihuela-Villar [2004] shows that when firms simultaneously make their strategic decisions, then 80% of them participate in the stable cartel.[46] The intuition is that

---

[42]This does not imply that it is more difficult to enforce the collusive agreement if there are more rivals in the market. Indeed, in Selten's [1973] model there is no room for cheating since cartels appear as non-cooperative Nash equilibria with perfect and complete information.

[43]For a simplified exposition of Selten's [1973] model, see Phlips [1995], pp. 23-28.

[44]The size of the market is determined by the number of active firms. Technically, this is due to the assumption that the demand function is homogeneous of degree one with respect to the number of firms in the market. This implies that i) the equilibrium price depends on the number of firms in the market and ii) the profits per firm, when the price is chosen optimally, depend on the fraction of firms in the dominant cartel. As d'Aspremont et al. [1983] admits it remains an open question whether this conclusion holds with more general demand and cost functions.

[45]This confirms previous findings by Selten [1973], d'Aspremont et al. [1983] and Donsimoni et al. [1986]. Moreover, it resembles to Kamien and Zang [1990] who show that by prohibiting mergers that result in firms acquiring more than 50% of the the firms in the market partial, monopolization is completely eliminated.

[46]Note the similarity of this finding with the *merger paradox* result by Salant et al. [1983]. The authors show that a merger (that entails no efficiency gains) between two or more firms in a Cournot oligopoly with homogeneous good is not profitable unless it accounts for more that

a cartel member appropriates a higher share of collusive profits when only a small number of firms participate in the cartel. At the same time, joining the cartel is more appealing since outside cartel competition is fierce. Conversely, if the number of cartel participants is large, cartel profits are distributed among many firms, and, therefore, abandoning the cartel becomes more tempting. The profits of the firms that remain outside the cartel are relatively high compared to cartel profits.

The role of private information on firms' incentives to form a cartel is examined by Hviid [1992]. The author assumes that the cartel operates as a quantity leader and focuses the analysis on an involuntary information transmission from the cartel to any outside firms. The information pertains to a common random parameter in the demand function. The main finding of the paper is that private information dampens firms' incentives to form a cartel. Apart from the existing free-riding effect, Hviid [1992] shows that outsiders also benefit from the correct inference of the cartel's output choice.

A significant contribution to the literature focusing more generally on coalition formation procedures is made by Bloch [1996, 2002]. The author illustrates the difference between simultaneous and sequential games of coalition formation, and games of open and exclusive memberships. In simultaneous procedures, the formation of a cartel is not feasible because firms have an incentive to free-ride on a cartel that is formed by other firms. Contrarily, in sequential move games the formation of a (partial) cartel is feasible because of the ability of the firms, and, in particular, those firms that move first, to credibly commit to stay out of the cartel.[47] Moreover, Bloch [2002] investigates the prospect of signing bilateral market sharing agreements between firms, whereby firms commit not to enter each other's market territory.[48]

1.2.3 Cartel Sustainability

In spite of the absence of an explicit agreement, firms can circumvent the competitive pressures that squeeze their profits by recognizing the interdependence of their strategic choices. This idea originates form Chamberlin [1929, 1933] and it has been formally developed by Friedman [1971]. Friedman [1971] shows that collusion can be sustained if firms condition their conduct in the market on the history of their strategic interaction.[49] In particular, the author demonstrates that firms are able to maintain a collusive equilibrium by adopting grim trigger strategies.

80% of the total number of firms in the market.

[47]The sequence of moves has a commitment value.

[48]For a selective survey on the major contributions of the literature on cartel formation see Bloch [2005].

[49]Essentially, the collusive outcome can be supported in equilibrium if firms are sufficiently patient (i.e., if the discount factor is sufficiently high).

This type of strategy calls for an eternal reversion to the *stage game equilibrium* in case of deviation from the collusive equilibrium.[50] A weakness of these strategies is that they preclude the possibility to adapt the level of punishment, if a deviation from the collusive equilibrium occurs.[51] Abreu [1986, 1988] remarks that if the payoff at the punishment phase is positive, then the severity of the punishment can be strengthened.[52] This would decrease the stream of post deviation profits, and, hence, facilitate the sustainability of the collusive equilibrium. Specifically, Abreu [1988] defines the concept of '*optimal penal codes*' as a particular trigger strategy that entails a punishment, which may lead firms to make losses for a number of periods (the '*stick*') in case of defection from the collusive equilibrium. The optimal penal code prescribes that firms will return to the collusive equilibrium (the '*carrot*') provided that all of them participate in the punishment phase for a given number of time periods; otherwise, firms will remain in the punishment phase.

The above suggests that an essential condition for the success of tacit collusion is repeated interaction of firms in a market. Without repeated interaction, firms will be confronted with a Prisoners' Dilemma situation; although all of them can make higher profits by coordinating their actions, each of them can expropriate more by unilaterally diverging from the collusive equilibrium. On the contrary, with repeated interaction, the firm that is tempted to renege from the collusive equilibrium, by imposing, for instance, a lower than the agreed price, acknowledges that its benefit from the deviation will be ephemeral if the other collusive firms are competent enough to trace and reciprocate such a behavior. Thus, another essential condition for the sustainability of collusion is the existence of sufficient transparency to enable or facilitate the detection of such deviations. Indeed, Stigler [1964] argues that the greatest challenge to the sustainability of a cartel is secret price cuts. The weakness or inability to monitor the compliance to the agreement may trigger or precipitate the occurrence of price wars among cartel members. According to Stigler [1964], price wars signal the collapse of the cartel.

However, Green and Porter [1984] argue that price wars do not always confirm Stigler's conclusion. In particular, Green and Porter [1984] maintain that occasional price wars are a property of the collusive equilibrium, even if no ac-

---

[50]This is because trigger strategies set an upper bound, which is defined by the Nash Equilibrium of the stage game, to the severity of the punishment when a defection is detected. Another weakness of grim trigger strategies is that they are not *renegotiation-proof*. That is, it is in the interests of firms to negotiate again in order to restart collusion, depriving the grim trigger of credibility, and thus effectiveness. For more on *renegotiation-proof equilibria* in repeated games see Farrell and Maskin [1989] and van Damme [1989].

[51]For other significant studies characterizing the optimal penal codes see Lambson [1987, 1994, 1995] and Häckner [1996].

[52]As we will see in section 2.4 the punishment strategies adopted by colluding firms significantly impact on the magnitude of the deterrence effects of leniency programmes.

tual deviation occurs. Effectively, price wars are a mechanism that incentivises colluding firms to abstain from opportunistic behavior.

As already mentioned, a difficulty that firms face with tacit collusion is the coordination to a commonly acceptable equilibrium, given that there is a plethora of potential candidates.[53] The feasible set of equilibria consists of a set that ranges from the competitive to the monopoly equilibrium. With repeated interaction firms can enforce any of these equilibria, provided that they sufficiently discount future profits. Many theorems have been developed in the literature, commonly known as '*Folk theorems*', that allow for the justification of the selection of specific collusive equilibria in infinitely repeated games.[54] In effect, Folk theorems provide an appropriate tool for the analysis of the sustainability mechanism of collusion.

The second strand of the economic literature on collusion focuses on the identification of factors that enables firms to coordinate and sustain their cooperation without necessarily resorting to an explicit agreement. Although the assumptions of the papers in the relevant literature differ, the underlying mechanism that drives their results is the same. The mechanism operates through i) the effect on firms' ability to coordinate to a commonly acceptable collusive equilibrium and ii) the costs and benefits that emanate from cartel defection. Among the factors that are considered to facilitate the sustainability of collusion are the following: small number of firms[55]; high market concentration rate[56]; high frequency of interaction[57]; multi-market interaction[58]; market transparency[59]; and information exchange[60]. On the other hand, factors that facilitate surpassing the problems of a cartel's sustainability include: barriers to entry[61]; cost asymmetries[62]; buyer power[63] and; asymmetric information[64]. The effects of restrictions in productive capacity of

---

[53]Tirole called this the '*embarrassment of riches*'. See Tirole [1988], p. 244.

[54]For instance, see Friedman [1971], Aumann and Shapley [1976], Rubinstein [1979] and Fudenberg and Maskin [1986].

[55]Shapiro [1989].

[56]Bain [1956].

[57]Snyder [1996].

[58]Bernheim and Whinston [1990], Spagnolo [1999] and Scott [2008].

[59]Porter [1983], Green and Porter [1984] and Abreu [1986].

[60]Porter [1983], Farrell [1987], Kandori and Matsushima [1998], Compte [1998], Athey and Bagwell [2001] and Kühn [2001].

[61]Bain [1956], Modigliani [1958], Labini [1969], Tirole [1988], Neven [1989] and Wilson [1992] – for empirical research see Levenstein [1995], Clay and Troesken [2002] and Levenstein and Suslow [2006].

[62]Bain [1948], Rothschild [1999] and Miklós-Thal [2009].

[63]Stigler [1964] and Snyder [1996].

[64]Roberts [1985], Cramton and Palfrey [1990], Kihlstrom and Vives [1992], Athey and Bagwell [2001], Athey et al. [2004], Compte [1998], Kandori and Matsushima [1998] and Verboven [1998].

firms[65], asymmetries in the size and distribution of productive capacity of firms[66], product differentiation[67] and business cycles[68], are less clear-cut.

1.2.4 Asymmetry of information between antitrust authorities and colluding firms

The possibility of firms' engagement in illegal collusive activities brings to the surface an acute information problem. This emanates from the existence of asymmetric information between an antitrust authority and the firms that participate in a cartel. Rational firms, fully conscious of the illegality of collusive agreements, adopt a '*code of silence*' to conceal their unlawful conduct, and remain undetected by antitrust authorities. Consequently, the greatest challenge for an antitrust authority is to gather or elicit incriminating evidence to expose and eventually convict illegal cartels and their members. The hunt for evidence may not always be fruitful as collusive firms may manipulate or even destroy the evidence. Additionally, market investigations and appeal processes are costly for an antitrust authority, as they require the commitment of resources. More importantly, even though an antitrust authority detects such evidence it will nevertheless be hard to convict the cartel because of the high standard of evidentiary proof required by courts.

Enforcement policy instruments The portfolio of standard policy instruments an antitrust authority has to combat illegal cartels comprise of i) market investigations and ii) fines for convicted cartels. These two instruments significantly impact on cartel deterrence, as they jointly determine the expected fine for a cartel. An increase, either in the magnitude of the fine or the probability of detection, suffices to alter the relationship between the expected costs and benefits emanating from the participation to a collusive agreement, as is reflected by the participation constraint. Moreover, by changing the balance of the relationship between deviation and compliance profits, as reflected by the incentive compatibility constraints, such changes may instigate the collapse of existing cartels.

The efficacy of such an enforcement policy hinges on the size of the maximum feasible fine[69] and the resources available to the antitrust authority. The latter

---

[65]Brock and Scheinkman [1985] and Davidson and Deneckere [1990].

[66]Compte et al. [2002].

[67]Deneckere [1983], Chang [1991], Ross [1992], Raith [1996] and Symeonidis [2002].

[68]Rotemberg and Saloner [1986], Haltiwanger and Harrington [1991], Staiger and Wolak [1992] and Bagwell and Staiger [1997].

[69]Becker was the first to explicitly argue that fines are bounded by the offender's wealth. See Becker [1968], pp. 183-185 and pp. 191-193. In a similar vein, antitrust fines are also argued to be bounded. The argument is twofold. Firstly, high fines may jeopardize a firm's financial vigor, and even its market survival, contrary to the ultimate objectives of antitrust policy. Thus,

constrains the number of market investigations and, hence, the probability of detection. Ideally, the probability of detection would be equal to 1, assuming that all cartels retard competition and produce malign welfare effects.[70] However, given the scarce resources an antitrust authority has at its disposal, the actual probability of detection is significantly less than this level.[71] Therefore, the cap on the fine and the limited resources of the antitrust authority condition the efficiency of the enforcement policy.

The boundaries to the effectiveness of the enforcement policy may be relaxed with the advent of an information revelation mechanism, such as a leniency programme. Essentially, leniency programmes are mechanisms that incentivise collusive firms and/or individuals to come forward and denounce the cartel. A properly designed leniency programme can break the walls of secrecy that surround the cartel and its members from inside. The concession may range from a reduction in the fine to the grant of a positive reward/bounty to the whistle-blower(s).

A leniency programme has many constituent elements. Specifically, it stipulates, among other things, the number of possible beneficiaries of leniency, the stage of the investigation procedure at which leniency is offered (before or after the investigation), whether a ringleader or a repeated offender is qualified to apply for leniency, the size of the fine reduction and the impact of the quantity and quality of information, and the sequence with which such information is provided. The transparency and confidentiality are two other influential elements for the effectiveness of leniency programmes. All these elements are crucial for the design of leniency programmes achieving the optimal level of cartel deterrence.[72]

---

antitrust fines are bounded to avoid bankruptcy. Secondly, the principle of proportionality applies. According to this principle the severity of the fine must be correlated to the magnitude of the harm caused by the offender. In other words, the punishment should fit the crime. Taking also into account the possibility of legal errors, which implies that fines are not socially costless, the application of this principle also restricts the level of the fine. For a recent discussion on the optimal antitrust fines, see Buccirossi and Spagnolo [2007] and Houba et al. [2011].

[70]We deliberately neglect the possibility of welfare enhancing effects as this is not the focus of this survey paper.

[71]Bryant and Eckard [1991] estimate the probability of detection to be at most in the range of 13-17% for the US using data from 1961-1988. A similar study by Combe et al. [2008] estimate the probability of detection at EU level to be in the range of 12.9-13.3% using data from 1969-2007. A more recent study by Ormosi [2011] estimates that cartel detection rate in the EU has stayed below 20% for most of the analyzed period (1985-2005), and it frequently dropped under 10%.

[72]Hammond distinguishes three essential cornerstones of an effective Leniency Program: 1) severe sanctions for those firms participating in cartels, 2) increased fear of detection by antitrust authorities for non-reporting firms and, 3) transparent leniency programme so that firms can predict with a high degree of certainty how they will be treated if they apply for leniency, and what the consequences will be if they do not. See, Scott Hammond, Director of Criminal Enforcement, Antitrust Division, US Department of Justice, ICN Workshop on Leniency Programmes,

A leniency programme can have both deterrence and desistance effects. It may enable, respectively, the deterrence of cartels that would otherwise have formed, and the detection of existing cartels that would otherwise have gone undetected. As a result of the introduction of a leniency programme, the efficacy of antitrust enforcement may improve. To fully comprehend the underlying working mechanisms of such programs, we first need to identify the costs and benefits emanating from a leniency application. The determinative benefit for a cartel member that applies for leniency is its privileged treatment with regard to the level of the fine it will pay in case of conviction by the antitrust authority. This benefit is a function of the avoided fine, the probability of successful prosecution of the cartel, once detected, and the probability with which the whistle-blower actually receives leniency. On the other hand, the cost for a firm that actively collaborates with an antitrust authority, on the basis of a leniency programme, is represented by the collusive supra-normal profits that the firm forgoes due to the collapse of the cartel, and the likely retaliation measures by its ex partners.

The next subsection briefly cites the main effects of leniency programmes, both on cartel stability and cartel sustainability. A more systematic exposition of these issues is reserved for Section 2.4.

1.2.5 Deterrence effects of leniency programmes

To begin with, a leniency programme may have direct deterrence effects.[73] Theoretically, it is possible that a leniency programme achieves the first best of complete deterrence of cartels without the antitrust authority incurring any cost for active investigations.[74] Spagnolo [2004] proves that if only the first self-reporting firm is eligible for leniency, and if this firm is rewarded with the fines imposed on all other cartel members, then the efficacy of the enforcement policy would reach its maximum at zero cost.[75] A similar finding is uncovered by Buccirossi and Spagnolo [2006a]. Moreover, Motchenkova [2004] demonstrates that when fines are proportional to the accumulated collusive profits and the application procedure for obtaining leniency is confidential (firms cannot instantaneously react to the

---

22-23/11/2004, Sydney, Australia.

[73]The deterrence effects of Leniency Programs are explored in more details in Section 2.4.

[74]Notably, this theoretical result contradicts the results of a recent experimental study by Bigoni et al. [2012]. These authors show that deterrence is enhanced when positive rewards are not available to self-reporting (experimental) subjects.

[75]Houba et al. [2011] confirm this result in a more general setting without imposing any restrictions on functional forms and the mode of competition between firms. The main novelty of this paper is that it relaxes the assumption regarding the protection of a deviator form the punishment by other cartel members.

actions of their rivals), then complete deterrence is feasible, provided that the fine is sufficiently high.[76]

The deterrence power of the enforcement policy may also amplified by a leniency programme that offers bounties to individual self-reporters. As shown by Aubert et al. [2006], when such a leniency programme is available, firms must pay higher compensations to their employees in order to preserve their silence. Markedly, the higher the compensation, the higher the cost of collusion. Therefore, cartel deterrence may be fostered by the introduction of a leniency programme that is available to individuals.

A leniency programme may also impact on cartel deterrence indirectly. The collapse of cartels' secrecy walls from inside and the ensuing desistance of cartels may relieve prosecution costs and accelerate the judicial procedures. As demonstrated by Motta and Polo [2003], the released resources of the antitrust authority could be used to enhance cartel detection rates, and thereby leniency programmes can indirectly increase cartel deterrence.

1.2.6   Desistance effects of leniency programmes

A leniency programme may also trigger the collapse of existing cartels by undermining their (internal) stability.[77] In particular, as shown by Spagnolo [2004] and Harrington [2008], a leniency programme may increase the payoff of cheating in the market by protecting the defector from antitrust fines. This will be the case if a firm simultaneously defects in the market and applies for leniency. This impacts on firms' incentive compatibility constraints, and, thus, cartels' sustainability. Even if a firm is unwilling to blow the whistle, there is a genuine danger that some other firm will. This potential undermines the fragile trust among cartel members. Effectively, a leniency programme thrusts firms into a Prisoners' dilemma through a 'race to the court' which ultimately squeezes the expected profits from collusion. Thus, by undermining the already feeble trust among firms, a leniency programme may increase the *riskiness of collusion*. However, this effect transpires only when the whistle-blowers are offered advantageously discriminate treatment with regard to the rate of fine reduction.[78] A similar result exists for programs offering leniency to individuals. In effect, with the introduction of such programs, firms are not only in a race with their rivals but also in a race with their culpable employees.

---

[76]However, contrary to Spagnolo [2004], the optimal rate of law enforcement is positive. In other words, to achieve the complete deterrence of cartels, the antitrust authority must credibly commit some of its resources to market investigations.

[77]The desistance effects of leniency programmes are thoroughly explored in Section 2.4.

[78]An extreme example is when only the first self-reporting firm is eligible to receive lenient treatment.

Leniency programmes may also vitiate the punishment mechanisms that firms use to curb opportunism and secure the conformity to the agreement. As shown by Spagnolo [2004], a leniency programme may increase the future fines and reduce the expected profits from continuing the participation to the cartel.[79]

Despite their benefits, the economic literature highlights that only properly and carefully designed leniency programmes can fruitfully induce self-reporting. Indeed, the relevant literature underscores that poorly designed leniency programmes may result in counterproductive side effects. For instance, wrongly designed leniency programmes may exacerbate collusive behaviour or provoke cartel formation. As Motta and Polo [2003] argue, leniency programmes essentially decrease the expected cost of cartels, and thereby can, contrary to their principal objection, cultivate cartel formation. Moreover, Ellis and Wislon [2003] and Motchenkova and Leliefeld [2010] maintain that leniency programmes may provide collusive firms with a disciplining device to thwart defections from the illegal agreement. Similarly, Buccirossi and Spagnolo [2006a] demonstrate that leniency programmes may alleviate the inherent hold-up problems of some types of previously infeasible collusive arrangements. Aubert et al. [2006] also show that leniency programmes may prevent benign cooperation between firms (i.e., legitimate joint ventures) or detrimentally affect the internal structure of colluding firms.

The next section presents a systematic review of the literature on the optimal enforcement of (generic) law.

1.3   The Economics of the Optimal Enforcement of Law

The economic literature on the optimal law enforcement begins with the seminal paper by Becker [1968].[80] Becker's pioneering ideas have been the locomotive for the development of rigorous economic analysis of crime and punishment. The investigation of criminal behavior was, until then, dominated by sociologists and psychologists who argued that crime reflects social environment problems (e.g., problematic families, social inequality, class warfare, subculture and social oppression), and that criminals were helpless '*victims*'. Crime was also linked to biological characteristics (e.g. natural-born criminals, heredity, organic disorders) or to psychopathological characteristics (e.g., mental illness, distress, dysfunction). However, none of these *theories* managed to adequately interpret the causes of

---

[79]This will be the case if i) the collusive agreement is sustained by two phase stick and carrot punishment strategies à la Abreu [1986, 1988] and ii) a recidivist is severely punished.

[80]For a systematic and meticulous survey of this literature see Polinsky and Shavell [2000] and Garoupa [1997].

criminal behavior, nor had any such theory focused upon the premise that all individuals are potential criminals.

Becker's work is built on the assumption that individuals are rational expected utility maximizers. As such, their decision to engage in criminal activities is based on the comparison between the expected net returns from committing a crime, and the benefit derived from legitimate activities (the opportunity cost of crime). A public agency which implements and enforces law (enforcement agency) aims to minimize the social cost caused by crimes. The social cost of crimes involves both the harm caused to society by the illegal act and the enforcement costs. The latter includes all expenditures the agency makes for police and the courts. Prominently, these expenditures determine the probability of detection and conviction.[81] The law enforcement agency can choose the type of punishment (e.g. imprisonment or fines) as well as the severity of punishments.

In this framework, the optimal level of enforcement is attained when only those illegal acts for which the net gain that accrues to a violator is lower than the harm caused to society are prevented (inefficient crime).[82] Those criminal activities for which the individual benefit to the violator exceeds the loss caused to society should not be deterred (efficient crime). Thus, according to Becker's analysis only the socially inefficient crime must be deterred.[83]

Becker concludes that the least costly policy to enforce the law is to set the fine at its maximum feasible level[84] and to use the probability of apprehension as a complement to the fine. The idea is that the expected fine must be at a level that optimally deters crime.[85] This conclusion is crucially grounded on two assumptions. The first is that an increase either in the fine or in the probability of apprehension increases the expected cost of criminal behavior. The second is that fines are costless transfers. Although the former assumption seems to be realistic, the latter is controversial because of the non-binding nature of the enforcement agency's decisions and the costs associated with administrative control (e.g. possibility of erroneous decisions).

In particular, Becker's assumption that fines are simple transfers from criminals to taxpayers (and thus the cost of collecting a fine is unrelated to its size) does not hold when violators: i) engage in socially costly detection avoidance activities

---

[81] As we will see in Section 2.4, Motta and Polo [2003] discriminate the probability of detection from the probability of conviction. The rationale is that detection does not always guarantee conviction.

[82] This rule resembles to the contemplations of Beccaria [1986]. According to Beccaria, the main objective of law enforcement is to reduce the frequency of inefficient outlawed courses of action.

[83] This result is known in the literature as the 'Becker's rule'.

[84] The maximum feasible fine is restricted by individuals' wealth so as to prevent bankruptcy.

[85] In Becker's model the probability of detection serves as a 'price' incentive that induces the quantity of offences 'supplied' by individuals.

(Malik [1990]), ii) have an aversion to risk (Polinsky and Shavell [1979]) and, iii) can commit crimes of different gradation of seriousness (Stigler [1970]). In those instances, the optimal fine may not be the maximum feasible fine.

Finally, Becker argues that fines are preferred to other types of punishment, such as imprisonment. Unlike fines, imprisonment is socially costly, since a prisoner suffers a utility loss. In addition, with imprisonment the government ultimately ends up spending[86], rather than gathering resources, in order to maintain correctional institutions and facilities. Besides, Becker argues that prisoners are unable to productively work during their time in prison; hence the diminished labor force represents an additional cost to society. Nevertheless, he maintains that if fines are not feasible, then the enforcement agency ought to resort to imprisonment to achieve the optimal level of crime deterrence.

Malik [1990] contests, and indeed invalidates, Becker's main thesis, namely, that the optimal fine must be set at its maximum feasible level. He argues that if individuals can potentially engage in activities that reduce the probability of detection (avoidance activities), then it will not be optimal to impose the maximum feasible fine. The underlying rationale is that a higher fine induces individuals to invest more resources into socially unproductive avoidance activities.[87] This implies that fines are not costless transfers, as assumed by Becker [1968]. In the light of this, the design of the optimal law enforcement policy, should serve, in contrast to Becker [1968], two conflicting goals; optimizing deterrence and minimizing avoidance activities.

In this context, Malik shows that the optimal fine is such that the marginal benefit of a higher fine in reducing the enforcement cost of deterrence is equal to the marginal cost of raising wasteful avoidance efforts. The optimal level of the fine is, in principle, below the maximum feasible fine. He also argues that

---

[86]Resources used for establishing and maintaining jails and other correctional facilities.

[87]A limitation in Malik's analysis is that it disregards the possibility of cross-effects between crime and avoidance. This is basically due to the assumption that individuals make a binary decision, namely, to commit an illegal act of a specific degree or not. Any changes in the magnitude of the punishment (fine or probability of detection) may affect an individual's decision to commit a crime but not the severity of the crime. Thus, a change in the level of the punishment gives rise to two direct effects: the deterrence effect and the avoidance effect. The deterrence effect emerges because an increase in the punishment increases the marginal cost of crime, and, thus, reduces crime. The avoidance effect emerges because an increase in the punishment implies a higher benefit from avoidance, and, thus, increases investment in avoidance. However, there is another indirect effect between crime and avoidance. A greater investment in avoidance reduces the marginal expected punishment, thereby inducing more crime. Similarly, a lower level of crime reduces the expected punishment, thereby decreasing the marginal benefit of avoidance. The complementarity between crime and avoidance is explored by Nussim and Tabbach [2009]. The authors show that increasing the punishment may result in more crime, contrary to Malik [1990].

the social costs associated with avoidance activities provide a rationale to screen individuals and impose fines that are type dependent (i.e., fines that depend on the private benefits accrued to individuals by their criminal activity). In effect, screening allows the enforcement agency to eliminate the avoidance activities of those individuals whose private gain exceed the harm caused to society by their illegal act, without deterring them from engaging in the illegal activity. On the contrary, without screening, the optimal enforcement policy may deter the occurrence of socially efficient criminal activities for which the private benefit to the violator exceeds the cost caused to society. Hence, the possibility of screening a violator's private benefit may enable the law enforcement agency to reduce the avoidance costs without sacrificing efficient deterrence.

Polinsky and Shavell [1979] discuss the importance of individuals' attitudes towards risk. The authors argue that when individuals are risk averse, the optimal fine may be well below the maximal fine.[88] This is because when an individual is risk averse, a higher fine imposes substantial risk-bearing costs which are not translated into extra revenue for the enforcement agency (or the state). Since violators dislike risk, the expected fine falls by an amount reflecting the risk premium (the disutility caused by imposing greater risk on risk-averse individuals). Importantly, the real social cost of fines increases for risk-averse criminals as the fine increases.

Becker's remark that a higher fine implies a lower expected net utility to a potential offender is valid for single-act framework where each individual makes a binary decision, namely, whether to commit an illegal act or not. In a more realistic multi-act framework, where illegal activities can take on different graduations, higher fines may induce wrongdoers to switch to more harmful acts. The argument is that if the fine is unrelated to the severity of the crime, the marginal deterrence for more serious crimes could be very small or even negative. Therefore, a violator who is threatened with a uniform fine, regardless of the severity of the violation he commits, would have more incentives to substitute that crime for a more serious one, from which he obtains a higher benefit.[89] The marginal deterrence of less serious crimes is also distorted from higher fines. For instance, over-enforcement could emerge with the risk of judicial errors (e.g. prosecution of innocent individuals).[90] Taking into consideration the distortion to marginal deterrence, Stigler

---

[88]In the extreme case where the cost of investigation is sufficiently small, the optimal probability of detection converges to one. This is because the decrease in utility due to the risk of a fine is higher than the potential benefit from a lower probability of detection. In this case, the optimal level of the fine is equal to the benefit of wrongdoers from the law violation.

[89]For instance, if robbery is punished as severely as murder, the robber might as well kill his victim to eliminate a witness.

[90]The costs of defense of innocent individuals, whether borne by themselves or the government,

[1970] shows it is optimal to impose only moderate fines on less serious violations to maintain sufficient marginal incentives to deter more serious offenses. The condition for optimal marginal deterrence advocated by Stigler states that the marginal benefit of an illegal activity should be equal to the marginal expected fine to deter such an activity.

1.4 Enforcement of Antitrust Law

Building on the insights of Becker [1968] and Stigler [1970] that penalties should be sufficient to induce offenders to internalize the full social cost of their illegal actions, Landes [1983] lays the foundations for the economic analysis of antitrust enforcement. In particular, Landes develops the economic theory of the optimal antitrust penalties.

In a simple setting where detection of violations and enforcement of the law are perfect and costless, the optimally deterring penalty equals the net harm to persons other than the violator. In more realistic settings wherein the probabilities of detection and conviction are less than perfect, and enforcement costs are non-trivial, the optimal penalty equals the net harm (which includes the enforcement costs per case) multiplied by the inverse of the probability of a fine being effectively imposed (i.e., the product between the probability of detection and the probability of conviction).

Notably, this rule allows for possible benefits that accrue form the offenses to third parties. More precisely, the net harm rule makes the offender internalize all costs and benefits of the violation. Accordingly, this rule implies that the offender commits only the '*efficient violations*', i.e., violations whose total benefits exceed the total costs.

Applied to the case of cartels, the benefits consist of the additional collusive profits, plus any cost saving and quality advancement as a result of the coordinated practice, net of any cost of enforcement. On the other hand, the harm consists of the consumer surplus transferred to firms in the form of overcharges plus the utility of the foregone consumption due to the higher price and the induced misallocation of scarce social resources, i.e., the deadweight loss.[91]

The theory of optimal penalties developed by Landes [1983] does not explore the effects of antitrust enforcement on firms' behavior (e.g. pricing decisions). The first systematic effort to study the impact of antitrust enforcement on cartels' pricing

---

are part of the social costs of enforcement.

[91]The net harm rule fits the Chicago School view that the primary goal of antitrust laws is to maximize the total economic welfare, i.e., the sum of the economic welfare of both consumers and producers. Clearly, this rule gives equal weight to the welfare of antitrust violators and that of their victims.

23

decisions is by Block et al. [1981]. The authors consider a static model of collusion where the antitrust policy instruments are endogenous. In particular, a higher price cost-margin raises more suspicions by customers or antitrust authorities, thus implying a higher probability of detection and conviction.[92] In addition, the penalty (fine plus damages) is also influenced by firms' pricing decisions since it is a multiple of collusive profits.[93] In this context, firms want to increase prices in order to achieve higher profits but avoid detection and conviction by an antitrust authority.

The analysis of the paper shows that the optimal cartel price is below the monopoly price and is decreasing in the penalty multiple and the intensity of antitrust enforcement.

The authors conduct an empirical test of the implications of the model in the market of white pan bread. The empirical results reveal that raising the Department of Justice's (DOJ) enforcement capacity or indicting a DOJ price-fixing complaint had the deterrent effect of contracting markups in the industry.

An implicit assumption made by Block et al. [1981] is that firms' price-cost margins are publicly known. This requires that the antitrust authority (or, in general, the enforcement agency) observes the collusive firms' cost functions. Besanko and Spulber [1989] relax this assumption and explore the optimal monitoring policy of an antitrust authority when firms' production costs are private information. In particular, the authors assume that cartels can be of two types: high marginal cost and low marginal cost.[94]

In this context, neither the antitrust authority nor the buyers (who lodge complaints to the antitrust authority) observe the cartel. Instead, they draw inferences from the observed price (and expected price-cost margins) and decide whether or not to pursue a case. What makes the problem interesting is that with incomplete information it is difficult to distinguish situations with high price due to high marginal costs from situations with high price due to collusion.

A crucial assumption of the paper is that the antitrust authority credibly commits to a set of instruments, i.e., fines and probability of investigation, that make the expected fine contingent on some observed signal, as for instance the market output. Given that the litigation procedure is costly, the antitrust authority may not wish to carry out its threat to prosecute cartels. Anticipating that they are not going to be prosecuted, firms would not change their behavior. Thus, the an-

---

[92] The authors find empirical support of this evidence in their study of the bread industry.

[93] This mainly reflects the US legal system wherein a price fixing conspiracy is subject to both criminal and civil sanctions. Civil sanctions such as treble damages are a multiple of markups.

[94] Pénard and Souam [1996] generalize this approach to the case of where the cost of production take a continuum of values. The results of the analysis are qualitatively consistent with the original paper.

titrust policy would not have any deterrence effects unless the antitrust authority credibly commits to prosecute cartels.

Besanko and Spulber [1989] show that, under full information, the fully competitive equilibrium can be costlessly achieved. Similar to Becker's logic, the optimal fine is the maximum feasible and the probability of detection is such that the expected fine equals the supra-normal profits accrued to a cartel. However, with asymmetric information the antitrust authority will always tolerate some collusion by low-cost firms (i.e., adopt a *laissez-faire* policy for slightly harmful price agreements) to economize on the costs of enforcement. Since the enforcement costs are decreasing in firms' production levels, the authority can allow '*little*' collusion to go unpunished. This would enable the antitrust authority to deter more serious cartels at a lower expected enforcement cost. Prominently, this result holds regardless of the level of enforcement costs. Finally, the authors show that the cartel's equilibrium price is decreasing in the fines, a result which is consistent with Block et al. [1981].

A common characteristic of the papers by Block et al. [1981] and Besanko and Spulber [1989] is that they explore the impact of antitrust enforcement on firms' pricing decisions in static settings. Several papers consider dynamic settings where the parameter time influences firms' pricing behaviour. The first attempt to integrate the probability of detection and fines into a repeated game of collusion is put forward by Cyrenne [1999]. The author considers an antitrust enforcement policy which focuses on significant price changes in an industry as a sorting mechanism for the allocation of resources devoted to monitoring collusion. In particular, the enforcement agency monitors the market price, and if it significantly falls for a number of periods and then reverts back to a higher price, it intervenes and investigates all firms for price fixing. To determine the effect of this enforcement strategy, the author considers the Green and Porter [1984] model where firms use a (finite reversion) trigger strategy as a punishment mechanism to sustain collusion. According to this strategy, firms initiate a price war if the price is below an agreed threshold – the trigger price - for a finite number of periods and then revert to the initially agreed price.[95]

In this context, prices can change even if antitrust enforcement is not included in the analysis, contrary to Block et al. [1981] where prices adjust only in responds to changes in the enforcement policy parameters (fine and probability of detection).

The author considers two versions of the Green and Porter [1984] model, i) quantity competition with homogeneous goods and ii) price competition with differentiated goods.

---

[95] However, in Cyrenne [1999] firms observe the market price and can monitor rivals' sales. Unlike Green and Porter [1984], the market price is determined by industry output and an additive stochastic component.

While the enforcement policy typically reduces the expected gains from collusion, it has no effect on the frequency of collusion unless fines are sufficiently large.[96] Importantly, undesirable effects may emerge if firms anticipate that the fines will not be sufficiently high.

In the first scenario, where firms produce homogeneous goods and compete in quantities, the enforcement policy may have a negative impact on the length of the punishment period. A defection from the agreement (increase in output) increases the probability of a price war, and, thus detection is more likely. This implies a higher punishment for a defection from the collusive agreement, since firms pay a fine in the case of detection, which allows firms to reduce the costs of their own punishment mechanism by limiting the length of the reversion period.

In the second scenario, where firms produce differentiated goods and compete in prices, the analysis shows that, for given trigger quantity and length of the reversion period, the enforcement policy allows firms to charge a higher collusive price than would be sustainable otherwise.

Another crucial assumption of the paper by Block et al. [1981] is that the probability of detection and the penalty depend on the price levels. In a dynamic setting, this assumption leads to a counterfactual result, namely, that there is a radical price increase when the cartel forms and then the price declines. Intuitively, a higher price in the current period increases current profits, but lowers the future payoff by both increasing the probability of detection and the penalty. Since the probability of detection and the penalty are both increasing in price, colluding firms steadily lower their prices over time to make detection less likely.

The impact of antitrust enforcement on a cartel's pricing dynamics is explored by Harrington [2005]. The author assumes that the probability of detection depends on price changes, rather than on price levels. In particular, large price movements raise more concerns to buyers and antitrust authorities about collusion, implying a higher probability of detection.

In this context, the cartel gradually raises its price to the steady-state level which is less than the monopoly price. The steady-state price is set to equate the incremental profit with the expected present value of the marginal rise in damages due to a higher price.[97] Notably, the steady-state price is decreasing in the damage multiple and the probability of detection while it is independent of the level of fixed

---

[96] When fines are not sufficiently high, the choice of collusive output or price is unaffected since the antitrust policy does not influence the probability that an unexpected price fall is realized.

[97] Harrington [2005] assumes that the punishment in case of detection comprises of a (fixed) fine and a multiple of the damages instigated by the cartel (historical and current). The latter depend on the cartel's price path contrary to Block et al. [1981] where damages depend only on the price level. This difference is due to the static nature of the model by Block et al. [1981]. A similar assumption is made by Harrington [2004].

fines. The latter implies that if fines are the only penalty, so that the penalty is independent of the price path, the steady-state price equals the monopoly price. In this case, antitrust enforcement does not influence firms' pricing decisions.[98] Finally, the author shows that when the damage multiple increases, the steady state price is reduced. This is because an increase in the damage multiple increases the cost of detection for the cartel, and, thus, firms lower their prices to decrease the probability of detection.

An implicit assumption in Harrington [2005] is that firms' incentive compatibility constraints are not binding. In another paper, Harrington [2004] investigates the role of antitrust policy in destabilizing the internal stability of cartels. The analysis shows that two qualitatively different cartel price paths may develop depending on the parameter values. One is qualitatively the same as in Harrington [2005] - the cartel gradually increases its price converging in the long-run to a steady-state level. The second type of price path has the cartel gradually increase its price, but then the price declines down to the steady-state. Even though reducing the collusive price shrink profits, and cannot make detection less likely (since the probability of detection depends on price changes), a price decline is essential to discourage defections from the collusive agreement. Moreover, the author shows that the risk of detection and penalties may deter a firm from defecting, anticipating that a price war may spark suspicions about collusion. Thus, antitrust enforcement can loosen firms' incentive compatibility constraints and have preserve effects on cartels' pricing.

Motchenkova and Kort [2006] explore the deterrence power of two distinct penalty structures for antitrust law violations: fixed and proportional. The latter penalty scheme takes into account all accumulated illegal gains to the violating firm(s). An implicit assumption of the study is that the antitrust authority can perfectly observe the accumulated rents from the antitrust law violation. The authors focus their attention to illegal price-fixing agreements. Using a dynamic optimal control model wherein a cartel maximizes its discounted expected payoff in the presence of antitrust enforcement, the paper shows that the proportional penalty scheme is more effective in achieving complete deterrence. Although complete deterrence is feasible with fixed fines, the level of the fixed penalty that attains complete deterrence is too high and leads to bankruptcy of the fines participating in the cartel. The proportional penalty scheme is more appropriate than the fixed penalty since it ensures complete deterrence of cartels in the long-run, even when the penalty is moderate. Moreover, the analysis reveals that cartel

---

[98]This result is known in the literature as the 'neutrality' of antitrust enforcement. A similar result has been shown by Block et al. [1981] when the probability of detection is insensitive to price.

deterrence is achieved sooner with a combination of high probability of detection and more severe penalty. Finally, when the penalty is already high, then a further increase in the severity of the punishment is less efficient than the policy that increases the probability of detection.

1.5   Self-Reporting Schemes

The literature on self-reporting schemes for generic crimes is divided into two ramifications. One studies the effects of self-reporting schemes offered to single individuals who commit a single illegal act. The other one, studies the effects of self-reporting schemes offered to a group of individuals who jointly commit a crime (organized crime). We survey the two subdivisions of the literature below.

1.5.1   Crimes committed by individuals

Kaplow and Shavell [1994] and Malik [1993] are the first to investigate the effects of self-reporting schemes on the optimal design of law enforcement.

Kaplow and Shavell [1994] show that if self-reporting individuals pay lower fines, then the social costs of law enforcement diminish. . The crux of their argument is that if the fine for those who self-report is infinitesimally smaller than the expected fine without self-reporting, while the fine for those detected is the maximum feasible, then all (risk-neutral) violators will self-report. Although deterrence is unaffected, given that the fine for self-reporting is slightly lower than the one without self-reporting, the enforcement costs are lower. This is because only those individuals who have not reported must be monitored. The authors maintain that the optimum enforcement policy is such that all individuals who commit an illegal act self-report. Nevertheless, the *first best* (complete and costless deterrence of illegal acts) is infeasible as the enforcement agency has to credibly commit itself to costly monitoring. Otherwise, no individual will have any incentives to self-report.[99]

Kaplow and Shavell [1994] develop various extensions to their basic model. Specifically, they show that if there is an administrative cost associated with fines, then self-reporting is beneficial only if that cost is sufficiently higher than the monitoring cost. The reason is that in a self-reporting regime administrative costs are borne with certainty, while in a regime without self-reporting these costs are borne only with the probability of apprehension. Furthermore, the authors show that if

---

[99]Spagnolo [2004] is the first to show that the *first best*, that is complete and costless deterrence, is feasible. However, the framework of his study is different from that of Kaplow and Shavell [1994]. The latter authors examine individual violations, while Spagnolo [2004] studies violations by a group of individuals.

the monitoring technology is imperfect, that is, if there is a positive probability of *Type I* or *Type II errors*[100] transpiring, then the cost savings from self-reporting are greater than those estimated in the basic model. The underlying reasoning is that in the presence of imperfect monitoring technology, the frequency of violations is, *ceteris paribus*, higher (as with some probability complying with the law is pointless).[101]

Another beneficial effect that derives from a self-reporting scheme is the provision of insurance to risk-averse individuals. Thus, in a self-reporting regime, risk cost is reduced since (risk-averse) individuals who self-report pay a certain amount rather than uncertain (expected) fines.

In another extension to their basic model, Kaplow and Shavell [1994] investigate the effects of imprisonment as a punishment for those individuals who do not self-report. Although imprisonment is costly (as suggested by Becker [1968]), Kaplow and Shavell [1994] show that the benefits stemming from a self-reporting regime are extensive. This is because all violators self-report at the equilibrium, and thus the imposition of imprisonment obviates. However, this result holds as long as the enforcement agency credibly commits to put in jail those violators who do not self-report. This conclusion is relevant whenever imprisonment is an optimal punishment when a self-reporting scheme is not available. Nevertheless, as Becker [1968] argues, even if imprisonment is not an optimal punishment, the threat of imprisonment, as an ex-post sanction for those who fail to self-report their illegal acts, always augments the advantages of self-reporting schemes. The intuition is that a given level of deterrence (i.e., a given level of *ex-ante* sanction) can be attained at a lower cost with a lower probability of detection. Only the violators that falsely report face a greater ex-post sanction. Importantly, since the ex-post sanctions are never actually imposed, the social costs of imprisonment are evaded.

In a similar vein, Malik [1993] explores the effects of self-reporting when fines are subject to administrative control. This implies that fines, contrary to Becker [1968], are not costless transfers.[102] In particular, the author analyzes the effects of

---

[100]Type I errors occur when the enforcement agency fails to punish/convict a guilty party (also known as '*false positives*'). Type II errors occur when the enforcement agency punish/convicts an innocent party (also known as '*false negatives*').

[101]For a discussion of the effects of Type I and Type II errors on the enforcement of competition law see Schinkel and Tuinstra [2006]. The authors conclude that an increase in the enforcement errors exacerbate antitrust law violations. The reasoning is twofold. Firstly, the probability of detection drops, leading to lower expected fines for firms. Secondly, the risk of being fined increases (because of the probability of an unjust punishment), inducing firms to collude as a '*precautionary*' measure.

[102]The analogy between environmental pollution and cartels is limited and vague, if it exists at all. Nevertheless, Malik's [1993] assumption that fines are not costless seems to be more realistic for the analysis of antitrust policy. This is because antitrust authorities' decisions are also subject

self-reporting in a context of regulation of stochastic environmental pollution where the privately known effort level of a firm affects the probability of an environmental standard violation. Within this context, the goal of the regulatory authority is to design a policy that elicits a truthful self-reporting of firms' pollution levels. The portfolio of policy tools of the regulatory authority comprises of: i) random audits in order to verify each firm's report, ii) penalties for generating pollution and iii) fines for submitting dishonest reports.

The main result of the study is that self-reporting does not necessarily reduce regulatory costs. Therefore, a policy that does not require firms to self-report may be preferred to one that does. This is because with a self-reporting scheme available, firms need to be audited less often and punished more often than without self-reporting. Whether self-reporting reduces or not the regulatory and social costs depends on the relative magnitude of the costs pertinent to auditing and fining. Malik concludes that self-reporting is more likely to reduce (increase) the regulatory and social costs when a) the audit costs are high (low), b) the maximum feasible fine is low (high), c) the accuracy of the monitoring technology is low (high) and, d) the desired effort level by the firm is high (low). Finally, he argues that an improvement in the accuracy of the regulator's monitoring technology is socially less valuable when a self-reporting scheme is available to polluting firms.

Innes [1999] identifies additional benefits that emanate from self-reporting schemes, namely, introducing and allowing the possibility of the early prevention of damages. Motivated by environmental law, the author develops a model wherein firms engage in activities that may cause socially costly accidents. In this context, each firm exerts some accident prevention effort to reduce the accident rate. Accidents do not cause fixed and immutable harm, but rather harm that can be reduced ex-post if firms engage in costly post-accident remediation activities. Innes assumes that a firm is eligible for leniency provided that it engages in remediation activities that (partially) remedy the damage caused by the accident.[103]

The author shows that by increasing the frequency of remediation rate, a self-reporting scheme enhances the ex-post benefits of remediation. While non-reporting firms engage in costly remediation activities only when they are caught by the enforcement agency, self-reporting firms always engage in efficient remediation (before or after the accident occurs). Another result of the paper is that

---

to costly judicial review.

[103]It is worth noting that there is no such a requirement for obtaining leniency for antitrust law violations. A colluding firm is eligible for obtaining leniency as long as it terminates its participation to the cartel, at the latest at the time it lodges a leniency application, and actively cooperates with the antitrust authority. Moreover, leniency programmes in antitrust do not require ex-post remediation actions by self-reporting firms. However, firms can be held liable to private damage claims.

a self-reporting scheme allows the enforcement agency to costlessly impose stiffer fines to non-reporters. Thus, a lower effort is required by the enforcement agency to achieve a target level of deterrence. Innes' analysis shows that in a regime without self-reporting the optimal fine for an accident may be set below the maximum feasible level, contrary to Becker's argument. If remediation is possible, then an increase in the probability of detection not only increases the frequency of accident detection, but it also mitigates the harm caused by the accident through a socially valuable remediation. This implies that the probability of detection and the fine are no longer substitute policy instruments.

By comparing the two regimes, with and without self-reporting, Innes [1999] infers that the optimal fine in the setting where self-reporting is feasible is lower than the one without self-reporting. The rationale is that self-reporting firms must be incentivise to undertake precautionary measures (effort) to prevent the accident.

One common prediction of the papers discussed above is that at the equilibrium all violators self-report. Nonetheless, this theoretical result is inconsistent with empirical observations. The discrepancy between theory predictions and empirical facts is the primary motivation for a new study by Innes [2000]. By segregating all other advantages of self-reporting schemes (as indicated by Kaplow and Shavell [1994], Malik [1993] and Innes [1999]), Innes [2000] suggests that one possible explanation for the fact that law enforcement agencies sometimes detect non-reporters is the existence of heterogeneity in violators' likelihood of apprehension. The study shows that at the optimum enforcement policy only those violators with high probability of apprehension, who would be over-penalized under a regime without self-reporting, self-report. Those individuals pay a fine equal to the harm they cause by their illegal acts. However, Innes shows that those violators who do not self-report continue to be under-penalized. Thus, he concludes that self-reporting can provide a screening mechanism that enables the enforcement agency to tailor the fines based on individuals' heterogeneity. The introduction of a self-reporting scheme may thus foster the efficacy of the enforcement policy.

Building on Malik [1993], Innes [2001] explores the impact of self-reporting schemes on detection avoidance activities. In this study, Innes identifies two advantages of self-reporting, those being the savings achieved both in avoidance and, enforcement costs. The main result of the analysis is that with self-reporting there is no need to engage in socially inefficient investments in avoidance activities. By setting the fine for self-reporting equal to the average fine that the violators would otherwise face, including the optimal avoidance cost, the enforcement agency can induce self-reporting without dampening individuals' incentives to avoid engaging in illegal acts. The additional benefit from self-reporting is that those individu-

als who self-report do not engage in avoidance activities and, therefore, the costs related to these activities are economized. Moreover, by deterring avoidance, self-reporting allows the enforcement agency to thwart violations with a lower enforcement effort. This is achieved by raising the non-reporting fines to the maximum feasible level. Therefore, Innes [2001] concludes that at the optimal policy, the same benefits identified by Becker [1968] without the additional costs identified by Malik [1990] could be attained.

The introduction of *ex-ante* asymmetric information, as studied by Innes [2001], was criticized by Feess and Heesen [2002]. These authors suggest that, although *ex-ante* asymmetric information allows explaining partial self-reporting, it does not account for the main drawback of self-reporting schemes, which is that they, *ceteris paribus*, sharpen the incentives to violate the law.

The authors consider a scenario wherein each violator receives a private signal about the individual probability of apprehension, after the violation is committed. Technically, this implies that the decisions related to the violation and self-reporting are made on different information sets, contrary to Innes [2001] who assumes that individuals' probabilities of apprehension are ex-ante heterogeneous. The distinguishing feature of ex-post asymmetric information is that it creates an *option value* for individuals which bolsters their incentives to violate the law. Nevertheless, Feess and Heesen [2002] show that the advantages of self-reporting dominate this undesirable effect.

More specifically, the authors show that when there is ex-post asymmetric information, the opportunity to self-report may provide those individuals that receive an unfavorable signal (i.e., those with high probability of detection) with an option value. If the enforcement agency's effort remains the same, as in the setting without self-reporting, then the option to self-report tends to increase the frequency of law violations. However, self-reporting violators need not be investigated, and, therefore, the enforcement agency may achieve a target level of deterrence at a lower cost. Thus, by taking into consideration this feature of self-reporting, the enforcement agency may coin a second best optimal enforcement policy. As a result, social welfare increases, even with the restriction that the violation rate must not exceed the one without a self-reporting scheme.[104]

1.5.2   Crimes committed by groups of individuals

The observation by Feess and Heesen [2002] that the introduction of self-reporting schemes sharpens, *ceteris paribus*, the ex-ante incentives of individuals to vio-

---

[104]This *second best* solution entails 1) a higher monitoring effort because of value of the option to self-report and 2) a fine for self-reporting that is slightly below the expected fine for the violator with the highest detection probability.

late the law has stimulated the research on self-reporting schemes for groups of violators. A distinguishing feature of this stream of the literature is that the enforcement agency can exploit the strategic interdependence between violators in order to increase the expected fine for each violator. This is despite of the fact that the probability of apprehension diminishes. This contrasts to the single violator scenario where the expected fine with self-reporting cannot exceed the expected fine without self-reporting, since otherwise no individual would self-report.

One paper that investigates the deterrence effects of self-reporting schemes to groups of individuals that jointly act to commit an illegal act is by Feess and Walzl [2004]. Their analysis suggests that the adoption of self-reporting schemes improves the efficacy of the enforcement policy, in spite of the fact that violators behave cooperatively at the self-reporting stage. In particular, the authors show that if violators behave non-cooperatively at the self-reporting stage, the enforcement agency may induce them into a Prisoners' Dilemma situation. The agency can achieve this by setting the maximum feasible fine if all violators self-report and by granting full amnesty (i.e., the reduced fine is equal to zero) only to a single self-reporter. Although all group members can be better-off if they credibly commit that none of them will self-report, each of them has an incentive to unilaterally run to the court first and benefit from amnesty.[105] While not all crime is deterred, all violators self-report with the least enforcement cost given that the optimal probability of apprehension is (almost) zero. If, on the other hand, violators behave cooperatively at the self-reporting stage to minimize their joint expected fines[106], then, as in Kaplow and Shavell [1994], the optimal fine for each violator is equal to the expected fine that individuals are liable to pay in the setting without self-reporting. In this case, even if all violators self-report, the probability of apprehension is weakly higher than the one with non-cooperative behavior. This implies that the enforcement cost is higher with cooperative behavior.

In addition, Feess and Walzl [2004] explore the optimal enforcement policy when the enforcement agency has imperfect information about violators' actions at the self-reporting stage, i.e., it is ignorant whether violators act cooperatively or non-cooperatively. The authors show that if individuals' cooperation probability is exogenous, the optimal self-reporting scheme is independent of the cooperative behavior rate. The optimal policy calls for a fine against a single self-reporter equal to the expected fine without self-reporting. Moreover, if all violators self-report, then the optimal policy calls for the maximal feasible fine, so that to induce violators into a Prisoners' Dilemma situation. The threat of being driven to a Prisoners' dilemma situation may, however, motivate accomplices to search for credible commitment devices to provoke cooperation. Thus, when the probability

---

[105]This effect of self-reporting schemes is known in the literature as '*race to the courthouse*'.

[106]This is also the case in the paper by Motta and Polo [2003] discussed in the next section.

of cooperation is endogenously determined, the optimal policy offers fine reductions to all violators that non-cooperatively self-report. Given that the expected fine is increasing in the probability of conviction, and that higher fines induce a higher degree of cooperation, a non-cooperative self-reporter may pay a fine which is below the maximum feasible level. The rationale is to limit the cooperation rate of individuals at the self-reporting stage. Therefore, an endogenous rate of cooperation, as maintained by the authors, may explain the partial reduction to fines for all violators in case of simultaneous self-reporting.

A supplementary explanation of why all members of the group should be offered a fine discount is the existence of imperfect self-reporting technology. For instance, if self-reporting by a single violator does not necessarily imply the conviction of other group members. Thus, if violators retain evidence of different quality expost, i.e., after they commit the illegal act, then it will be optimal to induce full self-reporting by all group members and impose the maximum fines. This would force violators into a Prisoners' dilemma situation. The optimal fine in this case is lower than the maximum feasible fine. Furthermore, the fine is higher for the violator who has been more often convicted on the basis of his accomplices self-reporting. Finally, the authors show that the fine for each violator depends on how informative the evidence submitted by other group members is to the enforcement agency.

1.6   Leniency programmes in Antitrust

The literature reviewed in the previous section addresses the optimal level of enforcement in generic law. In this section, we survey the literature that explores the implications of self-reporting in antitrust law enforcement, where the centre of interest is firms rather than individuals. This ramification of the literature has proliferated over the last decade, after the pioneering papers of Motta and Polo [2003], and Spagnolo [2004]. It has to be emphasized, at the outset, that at the heart of this literature is the exploration of the impact of leniency programmes on cartel desistance and cartel deterrence.

Motta and Polo [2003] are the first to study the impact of leniency programmes in antitrust policy.[107] The main research question of their paper is whether firms already under investigation should be eligible for leniency. Their study shows that a leniency programme has two effects on law enforcement. On the one hand, it may lower proceedings costs and lead to a temporary cessation of cartels (anti-collusive effect). On the other hand, it may exacerbate cartel activity as the expected

---

[107]For an exellent survey of the literature (theoretical, empirical and experimental) on leniency and whistleblowers in antitrust, see Spagnolo [2008].

cost of collusion decreases (*ex-ante* pro-collusive effect). The latter effect was also identified by Feess and Heesen [2002].

The main assumptions of the paper by Motta and Polo [2003] are the following: i) the antitrust authority has an exogenous fixed budget constraint that can be used either to detect or to convict cartel members; ii) the same lenient treatment is offered to all self-reporting firms regardless of the reporting order[108]; iii) a deviating firm can neither be convicted for taking part in a cartel nor can report its former partners; and iv) firms sustain collusive agreements by adopting grim trigger strategies. Assumptions ii and iii are less intuitive and are in odds with the leniency programmes implemented in various jurisdictions. In fact, antitrust authorities discriminate self-reporters depending on the self-reporting sequence, as well as on the quality of information that self-reporting firms reveal to the antitrust authority. Theoretically, the assumption of '*no fine differential*' among self-reporting firms relaxes firms' incentive compatibility constraints, and thus enhances cartel sustainability. In addition, an antitrust authority can in practice convict a deviating firm for taking part in a collapsed cartel. More importantly, a deviating firm has strong incentives to self-report its illegal conduct to the antitrust authority so as to pay a lower fine. Essentially, assumptions ii-iv of the paper, drain firms' incentives to spontaneously and non-cooperatively run to the antitrust authority to self-report. As we will see later, this is the main motivation of the paper by Spagnolo [2004].

The two major results of the paper are the following: i) leniency programmes are not unambiguously optimal and ii) it may be optimal to grant leniency to firms already under investigation if the antitrust authority has not yet obtained actionable evidence. With respect to the first result, the authors show that, in principle, it may not be optimal to introduce a leniency programme if the antitrust authority's budget is sufficiently high, so that cartels could be deterred with a sufficiently high probability of detection without leniency. If, however, the antitrust authority has limited resources, a leniency programme may be an optimal ancillary policy instrument from a second best perspective. In this case, the introduction of a leniency programme may generate a Prisoners' dilemma situation at the conviction stage, which induces firms to (collectively) self-report. Taking into consideration the collective application for leniency and the fact that a leniency programme reduces the expected costs of the cartel, the authors demonstrate that the optimal enforcement policy should provide full amnesty (as in Feess and Walzl [2004]) to all self-reporting firms so as to increase the probability of conviction. The second result of the paper is rationalized on the basis that the probability of conviction radically increases once the antitrust authority prosecutes the cartel. Thus, the

---

[108]For a convenient way to model asymmetric treatments to self-reporting firms, which depends on the sequence of self-reporting, see Houba et al. [2011].

prospect of reduced fines becomes more appealing for a firm under investigation. Therefore, leniency programmes that are open to firms after the commencement of an investigation by the antitrust authority might play a pivotal role in the prosecution of cartels.

It is worth noting that the paper by Motta and Polo [2003] is a close mirror of Kaplow and Shavell [1994]. Both papers highlight two benefits of leniency programmes (or self-reporting schemes). On the one hand, help economize enforcement resources and, on the other hand, provide insurance to risk-averse individuals or firms (a leniency programme essentially provides a riskless alternative to a probabilistic conviction).

The paper by Spagnolo [2004] can be viewed as a complement to Motta and Polo [2003]. However, the two approaches crucially differ in their focus and results. The study by Spagnolo [2004] focuses on whether a leniency programme can induce spontaneous self-reporting by firms, contrary to Motta and Polo [2003] who focus their analysis on whether firms already under investigation should be eligible for leniency. Moreover, although Motta and Polo [2003] highlight the indirect effect of leniency programmes on deterrence by facilitating post-detection prosecution, Spagnolo [2004] emphasizes the direct effect on deterrence caused by the undermined trust among cartel members.

Another difference between the two papers is that in Motta and Polo [2003] the cost of enforcement is exogenous, while in Spagnolo [2004], it is a choice variable. Moreover, in Motta and Polo [2003] detection leads to conviction only with some probability, while in Spagnolo [2004] detection and conviction are identified with a single probability, as in Becker [1968]. A fundamental difference is that Motta and Polo [2003] do not address risk dominance issues in their model (they assume in contrary that the equilibrium selection criterion is Pareto dominance), and as a result a leniency programme cannot induce firms to spontaneously self-report. Conversely, Spagnolo [2004] introduces risk dominance considerations, which are captured in the spirit of Harsanyi and Selten [1988].[109] The idea is that a leniency programme increases the risk of unilateral self-reporting, and thus undermines the trust among collusive firms. An additional difference is that in Motta and Polo [2003] firms sustain collusive agreements with grim trigger strategies, while in Spagnolo [2004] collusive agreements are sustained via optimal penal codes à la

---

[109]Technically, an equilibrium is *risk dominant* if the Nash Product (product of surpluses) of that equilibrium is greater than the Nash Product of every other potential equilibrium. Such equilibrium is called risk dominant because players try to mitigate the risks associated with their choices by minimizing the risks for every player. The concept of '*risk dominance*' was introduced by Harsanyi and Selten [1988]. Risk dominance should be contrasted to Pareto dominance where players select the Nash equilibrium that Pareto dominates all other Nash equilibria. For a meticulous discussion on equilibrium selection criteria, see Fudenberg and Tirole [1991a].

Abreu [1986, 1988]. Contrary to Motta and Polo [2003], Spagnolo [2004] considers the possibility that a defecting firm cannot be convicted for having taken part to a cartel, nor is it able to report information on its former partners. Finally, in Spagnolo [2004] leniency programmes take into consideration recidivism, contrary to Motta and Polo [2003].

A first observation by Spagnolo [2004] is that when a leniency programme is not available, the optimal enforcement policy requires the antitrust authority to commit not to target the firms that unilaterally defect from the collusive agreement. Prominently, this policy must be publicly known. Intuitively, a firm is more prone to defect from the collusive agreement if it knows that it will not be fined for its past illegal conduct. Spagnolo calls this effect 'protection from punishment'[110]. In principle, this effect mitigates the costs associated with a deviation from the collusive agreement, and thereby makes collusion harder to sustain.[111]

The main result of the study by Spagnolo [2004] is that the first best, of complete and costless deterrence, is feasible.[112] To attain the first best, the antitrust authority must impose the maximum feasible fine on all cartel members except the first to self-report (known in the literature as the 'first informant rule').[113] At the optimal policy, the first self-reporting firm is rewarded with the fines imposed on all other cartel members. This finding is in stark contrast to Becker [1968], who argues that the optimal probability of detection and the investigation costs that generate such a probability are a dead-weight loss that keeps society away from the first best. It is also distinguished from the paper by Motta and Polo [2003], according to which the leniency programme can, in the best case scenario, result in a second best solution. This is attributed to the fact that some resources need

---

[110]Harrington [2008] rename this effect to 'deviator effect'.

[111]Effectively, this assumption of protecting a defecting firm from the risk of punishment for participating in a collusive agreement in the past allows to overcome the 'irrelevance result' in Motta and Polo [2003].

[112]This result is based on a rather strong assumption, namely, that the antitrust authority can offer positive rewards to self-reporting firms. This has been criticized in the literature as politically infeasible. It is worth noting that the UK, South Korea, and Hungary have in place rewards programs for reporting information on cartels. In 2002 the (South) Korean Fair Trade Commission introduced a cartel informant reward scheme. This program was modified in 2002 and 2005 to increase the amount of maximum reward from 20 million won (approximately 18,789 USD) to 1 billion won (approximately 1 billion USD). A similar informant reward program was adopted by the Office of Fair Trading in UK. The program provides a reward rising to 100,000 GBP. In 2010 Hungary also introduced an informant rewards program. The program provides a reward of at least 1% of the fine levied against the cartel's members with an upper bound of 50 million forints (approximately 238,000 USD). Remarkably, parallel application to the leniency program and the information reward program is prohibited.

[113]In Becker [1968] the first best of complete and costless deterrence could not be attained even with infinite fines. This is because when no resources are invested in the enforcement of law, the probability of detection is zero. Thus, even infinite fines have no impact on crime deterrence.

to be invested in detection activities in order for the enforcement policy to have teeth.

Spagnolo [2004] shows that when the leniency programme is sufficiently generous, then two conflicting effects on cartel deterrence may emerge. On the one hand, the value of the collusive agreement increases since the leniency programme may be exploited; firms can collude and self-report in every period. Thus, a very generous leniency programme can rein cartel deterrence. This effect is, however, mitigated when the antitrust authority restricts eligibility for leniency only to the first self-reporting firm. This is because if more than one firm is eligible for leniency, fewer firms will have to pay the fine, and thereby a lower amount will be available to reward self-reporting firms. On the other hand, the value of deviation from the illegal agreement increases. A firm that defects from the collusive agreement is protected both from the antitrust fines ('*protection from fines effect*'[114]) and from the (temporary) punishment by the other cartel members ('*protection from punishment effect*'[115]). Thus, a very generous leniency programme can enhance deterrence. It is important to underline that the *protection from punishment effect* exists when the collusive agreement is sustained with two-phase Stick and Carrot punishment strategies à la Abreu [1986] and the punishment for repeat offenders is stricter. The intuition is the following: with two-phase stick-and-carrot punishment strategies, firms suffer a cost when participating to the '*stick*' phase, and are willing to do so because it allows them to return back to the collusive ('*carrot*') phase the following period. Furthermore, if repeat offenders are punished more severely than first time offenders, then a report today by one firm raises future fines. This diminishes the expected value of further collusion. As a result, the amount that firms are willing to bear in punishing defections in the first place is limited (i.e., the carrot tightens the incentive compatibility constraint for the punishment phase). Effectively, this drains the strongest credible '*stick*' that can be used to discourage defections in the first place.

In this context, Spagnolo [2004] proves that if positive rewards are feasible, the optimal policy is to impose the maximum possible fine to all firms but the first to self-report.[116] Restricting eligibility (only to the first self-reporting firm) makes the leniency programme less exploitable as it triggers a '*race to the courtroom*' caused by the '*fear to arrive second*'. Thus, the author infers that a combination of sufficiently high fines and high powered leniency programmes (courageous leniency programmes) makes law enforcement redundant. Nevertheless, when the maximum fines are sufficiently small to achieve the first best, it may be optimal

---

[114]In Harrington [2008] this effect is called the '*cartel amnesty effect*'.

[115]Harrington [2008] calls this the '*deviator effect*'.

[116]High fines are now crucial not only because they reduce the expected value of the illegal cartel, as in Becker [1968], but because they allow the enforcement agency (e.g., the antitrust authority) to offer higher rewards to the first self-reporting firm.

to couple rewards with active investigations, given that investigations and rewards are substitute enforcement instruments.

Moreover, Spagnolo [2004] shows that when positive rewards are infeasible, there is another effect of leniency programmes on cartel stability and deterrence. In addition to the protection from fines and protection from punishment effects, a leniency programme influences the riskiness of the cartel agreement by undermining the trust among its members. Collusive agreements are strictly riskier when the eligibility to the leniency programme is restricted only to the first self-reporting firm. A leniency programme that does not restrict eligibility only to the first self-reporting firm makes the illegal agreement less risky as firms can exploit such programs and enjoy fine discounts. Therefore, contrary to Motta and Polo [2003], Spagnolo [2004] concludes that it is always optimal to have a leniency programme (even if moderate).

An additional reason why colluding firms spontaneously apply for leniency, before an investigation is launched in the market, is offered by Ellis and Wislon [2003]. Their study explores the impact of a leniency programme on firms' behavior under Bertrand price competition with differentiated products. Similar to Motta and Polo [2003], the authors assume that the collusive equilibrium is sustained by grim trigger strategies. The novelty of the paper is that fines are not fixed, but proportional to revenues.[117] This implies that a change in marginal revenue can impact on firms' behavior in the market, and, ultimately, on their incentives to collude.

In this context, Ellis and Wislon [2003] show that firms can exploit the leniency programme by defecting in the market and denouncing the cartel. Thus, a leniency programme may destabilize cartels. In particular, a defecting firm can gain market advantage by raising its rivals' costs through fines and compliance measures. The authors highlight that firms' incentives to defect and apply for leniency are sharper when the fine is higher and the market less concentrated.

However, forward looking firms anticipate this opportunistic behavior and, thus, strategically use the leniency programme as a mechanism to thwart defections from their illegal agreement. If a firm defects from the agreement but does not self-report, then the other firms may punish the defector by self-reporting the cartel to the antitrust authority. In this case, a leniency programme can strengthen the stability of the cartel.

Contrary to Motta and Polo [2003] and Spagnolo [2004] who assume constant (time independent) probabilities of detection and conviction, Hinloopen [2003] studies the effects of leniency programmes when the probability of detection is

---

[117]This is a realistic assumption. Importantly, it is consistent with the fine policies in the US and EU.

time dependent. In particular, the author assumes that the probability of detection increases with a cartel's record. This may, for example, correspond to a scenario where the efficiency of an antitrust authority's prosecution activities constantly improves over time. A novelty of the paper by Hinloopen [2003] is that the antitrust authority is assumed to continue its investigation for a finite number of periods after the collapse of the cartel due to a market defection by one of its members. This allows the author to examine the effects of the duration of the limitation period that comes with antitrust law violation.[118]

By distinguishing two manners whereby a cartel may collapse, either by a deviation in the market or a deviation in the self-reporting stage, the author concludes that the effectiveness of leniency programmes increase with the generosity of the fine reduction. The reasoning is that a larger fine reduction for self-reporting reinforces firms' incentives to deviate from the collusive agreement by self-reporting, as opposed to deviating in the market. Furthermore, the rate of reduction of fines does not impinge on firms' incentives to deviate in the market, while it affects their incentives to self-report, as the cost of self-reporting diminishes. Thus, an additional reduction in the fine for self-reporting increases the probability of leniency application. A necessary condition for this result to hold is the existence of a sufficiently transparent leniency programme so that cartel members can accurately estimate the benefits accruing from self-reporting. However, this does not mean that the antitrust authority should publicly disclose information about the firms that apply for leniency. Such practice may allow firms to simultaneously apply for leniency, and thus exploit the leniency programme by paying lower fines.[119] Moreover, it may increase the risk for private claims, the costs of which is potentially greater that the fine imposed by the antitrust authority.[120]

Hinloopen [2003] argues that if the antitrust authority consistently improves the effectiveness of its detection activities so that the probability of detection in future periods increases, then the expected fine, and hence the cost of collusion rises. This implies sharper incentives to deviate from the cartel and self-report to the antitrust authority. The author also argues that an increase in the fine has two conflicting effects on firms' incentives. On the one hand, a higher fine implies a higher cost for collusion. This strengthens firms' incentives to defect, either in the market or by self-reporting to the antitrust authority. On the other hand, a higher fine implies a higher absolute value of the fine, net of the reduction for self-reporting,

---

[118]Acording to Article 25 of the EC Council Regulation 1/2003, the limitation period is (a) 3 years in the case of infringements of provisions concerning requests for information or the conduct of inspections; (b) 4 years in the case of all other infringements.

[119]It is precisely for this reason that Motchenkova [2004] argues in favor of confidential leniency programmes.

[120]On the other hand, the cost for private damages increases the total fine for a cartel member, and therefore enhances cartel deterrence.

which mitigates firms' incentives to self-report. The net product of these two opposite forces depends on the fine reduction relative to the sequence of per-period probabilities. If the percentage of the fine reduction is relatively high, compared to the increase in future probabilities of detection, then an increase in the level of the fine is more likely to improve the effectiveness of the leniency programme. Finally, Hinloopen [2003] argues that an extension to the limitation period makes self-reporting more likely to be preferred to deviation from the market, given that with self-reporting the cartel would be convicted immediately. In the case of deviation from the market, there is a possibility of being convicted in the future for past violations of the law. The length of the limitation period does not have any impact on the decision to self-report the cartel. Thus, the effectiveness of the leniency would not be affected by the duration of the limitation period.

It has to be remarked that, unlike Motta and Polo [2003], the study by Hinloopen [2003] cannot predict self-reporting when an investigation is underway. This is because the probability of cartel detection once an investigation is launched in the market is intact, contrary to Motta and Polo [2003].

Most studies in the literature on leniency programmes analyze the effects of such programmes under fixed fines, that are independent of the accumulated illegal collusive gains. However, fines may also be proportional to the accumulated illegal collusive profits. This is also an assumption that is closer to realism and the actual fining policies adopted by antitrust authorities in various jurisdictions. A paper that acknowledges and incorporates the distinction between fixed and proportional fines is by Motchenkova [2004].

Motchenkova [2004] explores the effects of leniency programmes under two distinct fine regimes, fixed and proportional, and two distinct application procedures, confidential and non-confidential. With a non-confidential procedure, firms can simultaneously apply for leniency.[121] Although this is not a very realistic assumption, it allows to highlight the role of confidentiality of leniency applications. Based on the approach of Reiganum-Fudenberg-Tirole[122], Motchenkova [2004] analyzes whether the treatment of collusive firms by the antitrust authority should depend on the timing of their applications for leniency. The main conclusion of the paper is that the driving force for the effectiveness of leniency programmes in deterring cartels is the strength of firms' incentives to stop colluding and apply for leniency before preempted by some other firm (*preemption mechanism*).[123] A

---

[121]For instance, if a firm applies for leniency, then this becomes publicly known, and, therefore, the other firms can instantaneously react and lodge a leniency application too.

[122]See Reinganum [1981] and Fudenberg and Tirole [1985]. The authors study the decision of whether and when to build a new plant or adopt a new innovation when the market can accommodate only one such addition.

[123]In effect, this mechanism induces firms in a race to the antitrust authority in order to be the

well designed leniency programme may reduce the duration of cartel agreements; however this result is ambiguous. Only strict leniency programmes unambiguously erode cartel stability, and thus enhance the efficacy of antitrust law enforcement. Similar to Spagnolo [2004], Motchenkova [2004] also shows that programs that do not restrict the eligibility for leniency only to the first self-reporting firm may restrain deterrence because of the reduced expected fines.

Motchenkova [2004] demonstrates that when fines are proportional to the accumulated collusive gains and the application procedure for obtaining leniency is confidential, then complete cartel deterrence can be achieved, provided that the fine is sufficiently high. At any rate, the introduction of a leniency programme (with these specifications) does not facilitate cartel formation or increase cartel duration. When the application procedure is not confidential, the leniency programme may reduce the duration of cartel agreements, but not deter cartel formation. This result is again ambiguous. When the fines and the probability of detection are sufficiently low, the introduction of a leniency programme may, on the contrary, foster collusion. Motchenkova [2004] also shows that under fixed fines and confidential leniency application procedures, the introduction of a leniency programme does not affect the effectiveness of the enforcement policy at all. In this situation, cartel duration depends on the size of the fine, while cartel deterrence is not influenced by the leniency programme. Only a sufficiently high fine can guarantee the success of the leniency programme. A policy recommendation that derives from the paper is that leniency programmes must be confidential, which seems to be the procedure adopted in most jurisdictions.

A typical assumption in the literature (including the studies by Motta and Polo [2003], Hinloopen [2003], Spagnolo [2004], Motchenkova [2004] and Harrington [2008]), is that the incriminating evidence retained by firms, and which can be used to apply for leniency, is symmetric. Although this symmetry assumption simplifies the analysis, it is not very realistic. For instance, firms may retain different pieces of evidence. The amount of evidence revealed to the antitrust authority critically affects the magnitude of the fine reduction and the probability of successful conviction of the cartel. One of the few papers that explores the relevance of the amount of reported evidence in the determination of the fine reductions and the number of firms that should be eligible to receive leniency is by Feess and Markus [2005].[124]

More formally, Feess and Markus [2005] develop a (static) model with two self-reporting stages, one before the investigation and one after the investigation, to explore the different amount of revealed evidence that is required to obtain leniency. A key result of the paper is that the strategic interaction between cartel

---

first to self-report and benefit from a lenient treatment.

[124] A more recent paper that sheds some light on similar research questions is by Silbye [2010].

members is different at the two stages. Self-reporting before detection leads collusive firms to a race to the courthouse, whereas after detection, it leads them to a Prisoners' Dilemma situation.

Feess and Markus [2005] conclude that the optimal fine hinges on the degree the evidence provided to the antitrust authority. The team specific characteristics, which affect the probability of cartel detection, also play a pivotal role in the determination of the optimal fine. Specifically, the authors demonstrate that granting full amnesty at the pre-detection stage is never optimal for the low evidence provider. Nonetheless, depending on the distribution of team specific characteristics, it may be optimal for the high evidence provider. The authors also show that the optimal fines are increasing in the amount of additional information delivered by other members of the cartel. A crucial assumption of the paper is that each firm is perfectly informed about the evidence retained by others. Furthermore, the evidence retained by an individual does not influence the level of the fine, a result driven by the Prisoners' Dilemma structure at the conviction stage. However, high evidence providers should pay a lower fine if they self-report at the pre-detection stage. Another significant result of the paper is that it is optimal to offer leniency to the low evidence provider if the high evidence provider self-reports. This provokes self-reporting by another cartel member, and thus save investigation costs (as in Motta and Polo [2003]). Finally, the paper concludes that the fine for the firm providing less evidence should be the same in both self-reporting stages (pre-detection and conviction), while the fine for the high evidence provider should be higher in the pre-detection stage. The rationale is that the antitrust authority has more to economize if self-reporting occurs at an earlier stage of the antitrust proceedings, while the cartel has more to lose given that the cartel is dissolved sooner. The authors remark that the theoretical predictions of the paper are consistent with the EU leniency programme, which may provide leniency to more than one firm. However, the predictions are in dissonance with the US corporate leniency programme - which provides leniency only to the first self-reporting firm.

A critical assumption of the papers discussed above is that the hard evidence that is generated by cartel members lasts only for one period and thereafter evaporates.[125] This is not always a realistic assumption. Evidently, evidence can be durable, and as such it can play the role of a '*hostage*' to discourage defections from the collusive agreement. One paper that factors in this assumption is by Buccirossi and Spagnolo [2006a]. The authors study the effects of leniency programmes in occasional sequential bilateral illegal agreements (or other corruptive activities) wherein reputation plays no role. Crucially, these types of transactions are exposed to serious '*governance*' or '*hold up*' problems because of the lack of

---

[125]In the paper by Motta and Polo [2003] evidence is produced only if no defection occurs. Thus, a defecting firm cannot apply for leniency.

an enforcement mechanism (i.e., a contract) that can constrain the development of opportunistic behavior by the party delivering last in the transaction.

The authors demonstrate that if the enforcement agency grants a fine reduction, or at best amnesty, without paying a positive reward to the first firm that self-reports, the commitment to cooperate with the enforcement agency may facilitate the punishment of a defecting party at a relatively low cost. As a result, a poorly designed leniency programme, for which the fine paid by the self-reporting party is lower than the expected fine without self-reporting, may provide an effective governance mechanism or credible threat to any party that attempts to deviate from the agreed strategy. The transacting parties can collect hard evidence to discourage opportunistic behaviour. In case of a defection, the mislead party can report the evidence to the enforcement agency. This counter-productive effect can be mitigated by offering a sufficiently high reward (above the expected gain from the illegal transaction) only to the first self-reporting party. To this effect, the paper by Buccirossi and Spagnolo [2006a] is in line with Spagnolo [2004]. Both papers conclude that collusion can be completely deterred if the enforcement agency sufficiently rewards the first self-reporting firm.

The main shortcoming of the papers assuming a fixed probability of detection (and conviction) over time is that they produce the counterfactual result that a cartel will never use the leniency programme. Nevertheless, this theoretical prediction contradicts the empirical evidence that many cartels established after the introduction of leniency programmes and eventually applied for, and obtained, fine reductions afterwards. Harrington [2008] argues that a logical explanation for this empirical observation is that the probabilities of successful prosecution stochastically change over time.

The study by Harrington distinguishes three effects of leniency programmes on firms' incentives to self-report. These effects are referred as the '*deviator effect*', the '*cartel amnesty effect*' and the '*race to the courthouse effect*'. The deviator effect operates through the payoff of defecting in the market. It captures the reduction of fines when a defecting firm applies for leniency. Clearly, this effect makes collusion more difficult to sustain as it increases the payoff from cheating. The cartel amnesty effect operates through the expected payoff of collusion. It captures the reduction in the size of the fine when a firm self-reports. This effect makes collusion easier to sustain when a firm obtains leniency. Similar to the cartel amnesty effect, the race to the courthouse effect operates through the payoff of collusion. This effect transpires when only the first self-reporting firm is eligible to receive a lenient treatment. The race to the courthouse effect makes collusion harder to sustain. Harrington [2008] allows for these three effects to simultaneously interact, thus permitting a richer analysis of the design of optimal leniency programmes.

The paper shows that the overall effect of leniency programmes on firms' incentives hinges on the size of the fine reduction. With an extremely lenient program that offers leniency only to the first self-reporting firm only the deviator effect and the cartel amnesty effect would be at play. In general, the deviator effect dominates the cartel amnesty effect. This implies that a more generous fine reduction to self-reporting firms makes the cartel agreement more unstable. The underlying reasoning is that the benefit for a firm that applies for leniency and self-reports today is greater than the expected benefit from using the leniency programme in the future. If those were the only two effects in operation, then collusion would be made more difficult to sustain under a more generous leniency programme. In this case, a policy of maximal leniency would be optimal. If, however, the fine reduction is sufficiently mild, only the cartel amnesty effect and the race to the courthouse effect would be at operation. A deviating firm cannot apply for leniency because the fine reduction is too small, while a cartel would be detected with certainty if a firm self-reports to the antitrust authority. In other words, a firm prefers to pay a lower expected fine than a higher certain fine. In this case, the fine reduction affects only the future expected collusive payoff. If the *cartel amnesty effect* were the only force at work, then a more lenient program would raise the collusive payoff, and thus facilitate collusion. A more lenient policy would then revive cartel stability contrary to the objective of the antitrust policy. The picture is less clearcut once the race to the courthouse effect and the cartel amnesty effect are both in force. In this case, a more generous fine reduction may increase or decrease the stability of the cartel depending on which of the two effects dominates. Cartel stability erodes when the race to the courthouse effect dominates.

Harrington [2008] proves that under some plausible conditions[126], it is best to offer full amnesty only to the first self-reporting firm; otherwise, partial leniency is optimal. Intuitively, when more than one firm is eligible to receive lenient treatment, the expected fine with self-reporting decreases (i.e., the cartel amnesty effect strengthens). On the other hand, the deviation profits remain intact (i.e., the deviator effect is intact). It is important to note that this result is in line with Spagnolo [2004], Motchenkova [2004] and Buccirossi and Spagnolo [2006a]. The author also shows that it is not always optimal to offer unrestricted amnesty to the first self-reporting firm. Amnesty should be offered only when the probability of conviction without self-reporting is sufficiently small and a firm delivers evidence that substantially reinforces the antitrust authority's case. In particular, the author shows that it is always optimal ex-ante (not ex-post) to offer leniency only if the probability that the antitrust authority wins the case, without self-reporting, is less than 50%.

---

[126]Technically, the Cumulative Distribution Function (CDF) of the probability of detection must be weakly concave.

Finally, it is worth noting that, contrary to Spagnolo [2004], the paper by Harrington [2008] have no implications for spontaneous self-reports.

The papers discussed above consider symmetric firms. The effects of a leniency programme on cartels' stability when firms are asymmetric is explored by Motchenkova and Leliefeld [2010]. In particular, the authors consider two firms which for historical reasons have asymmetric market shares (one has a large market share while the other has a low market share). The asymmetry in the size of firms implies different collusive profits for cartel members, and, thus, asymmetric punishment threats in case of defection from the collusive agreement. In effect, the larger firm has less to fear from a retaliation than the smaller firm. In this context, the authors explore the impact of a leniency programme on a cartel's stability, by means of its effect on the severity of the threat of retaliation.

In a setting where the small firm decides whether to self-report or not and subsequently the large firm decides whether to retaliate or not, the authors show that the leniency programme may foster cartel stability. The underlying reasoning lies in the ability of the large firm to use the leniency programme as a mean to enhance its trust to the small firm not to report the cartel to the antitrust authority. In other words, the leniency programme facilitates the large firm to '*coerce*' the small firm not to report with the threat of a severe punishment. Thus, firms may exploit the leniency programme to enhance cartel stability. It is worth noting that this adverse effect on cartels stability crops up even if the leniency programme restricts complete amnesty only to the first self-reporting firm. The authors show that increasing the fine or the probability of detection helps to alleviate this adverse effect of leniency programmes.

A policy recommendation that derives from the paper is that self-reporting firms should be protected from possible retaliation measures by other cartel members. For instance, the antitrust authority may closely monitor the market after a firm self-reports so that to discourage retaliatory measures, which may manifest with abuse of dominant positions (e.g. predatory pricing). Moreover, the antitrust authority may focus its investigations in markets with intermediate degrees of asymmetry wherein the adverse effects of leniency programmes are more likely to evolve. Finally, the antitrust policy should put greater emphasis on aggravating factors, such as coercion, when calculating the fines for illegal cartels.

A common limitation of the papers discussed above is that the collusive price and profits are fixed. In addition, the pricing behaviour of collusive firms does not influence the probability of cartel detection. These two assumptions imply that a leniency programme does not influence the price charged by a defecting firm, nor does it affect its profits.

The impact of a leniency programme on the cartel price path and cartel stability

is explored by Chen and Harrington [2007]. The authors extend Harrington [2004, 2005], where the probability of detection and fines are endogenous to colluding firms' prices[127], by considering a leniency programme as an additional tool for the enforcement of antitrust law. In this context, the cartel chooses a price path that ensures compliance by all firms (i.e., it is incentive compatible) without raising suspicions to buyers or the antitrust authority about the cartel.

A distinctive assumption of the paper is that detection can occur in the period after the collapse of the cartel (i.e., the period after which a firm defects from the illegal agreement). This assumption implies that the expected penalties can be higher when partial leniency is offered, compared to the policy of no leniency. As a result, a defecting firm can lodge a leniency application even if the ensuing payoff is lower than without a leniency programme to preempt others to self-report in the following period.

The authors show that a leniency programme providing full amnesty only to the first firm to self-report the cartel makes collusion more difficult to sustain. Moreover, their analysis shows that offering partial leniency can facilitate collusion (compared to offering no leniency), nonetheless, this result is ambiguous. When the probability of detection is weak, a defecting firm would prefer not to self-report to the antitrust authority. However, anticipating that the other firms will self-report in the next period and considering that only the first self-reporting firm receives a generous treatment, the firm chooses to defect in the market and simultaneously self-report. This strategy reduces the payoff to cheating which allows the cartel to sustain a higher price path. On the other hand, as the fine discount rate increases, firms have more incentives to defect in the market and simultaneously self-report to the antitrust authority. This tightens the incentive compatibility constraint which induces the cartel to price lower. Hence, even if a leniency programme is unsuccessful in deterring cartel formation, it may still achieve to impel the cartel to price lower in order to preserve the stability of the cartel.

The papers by Motta and Polo [2003], Hinloopen [2003], Spagnolo [2004], Motchenkova [2004], Feess and Markus [2005], Buccirossi and Spagnolo [2006a], Harrington [2008], Motchenkova and Leliefeld [2010] and Chen and Harrington [2007] investigate the effects of leniency programmes on firms' incentives to self-report. However, as discussed in Section 2.2, an antitrust authority has also other investigative tools to detect cartels and achieve compliance with antitrust law. One tool is individual leniency programmes. Such programmes may provide a new source of information for antitrust authorities. Indeed, if individuals are provided with sufficient incentives, then the firms that breach antitrust law would not only

---

[127]In particular, the probability of detection is increasing in price changes while the magnitude of the penalty, in case of conviction, is increasing in prices (current period and historical prices). See Section 1.4, pp. 26.

be in a race for leniency with their competitors but also with their own employees.

In addition, a typical assumption of the literature on leniency programmes is that collusion always generates hard evidence.[128] The hard evidence may last only for one period (as in Motta and Polo [2003]), and then evaporate, or it may be indelible (as in Ellis and Wislon [2003] and Buccirossi and Spagnolo [2006a]). The role of evidence retention has not been given sufficient attention in the literature. A notable exception is Aubert et al. [2006] who consider several explanations, in a rather informal way, for the puzzling fact of evidence retention. Evidence is crucial both for an antitrust authority and colluding firms. To credibly build a case against cartels, an antitrust authority needs to find cogent incriminating evidence. For instance, documents proving that firms have agreed to coordinate their conduct (e.g., fix prices). Thus, an essential assumption of the literature is that firms axiomatically keep hard evidence.[129] Moreover, to be eligible for leniency a firm must provide hard evidence to the antitrust authority, which is a realistic assumption to be made. However, the interesting question is not whether a cartel produces hard evidence or not but, given that hard evidence is produced, whether firms consciously decide to retain or destroy the incriminating evidence. This suggests that it may be indispensable to make the decision to retain or destroy the hard evidence endogenous, so that to analyze and comprehend the dynamics of the potential trade-offs involved. Considering these remarks, the conclusions derived by Aubert et al. [2006] are the following.

Firstly, courageous leniency programmes offering positive rewards to individuals have a greater deterrence effect than moderate leniency programmes offering at best amnesty (as in Spagnolo [2004]). Secondly, rewards to individuals can be more effective than rewards to firms. Thirdly, sufficiently low rewards may instigate adverse effects on the benign cooperation among firms, as well as on firms' organizational structure. Finally, the probability of retaining hard incriminating evidence by firms, or individuals, increases with the magnitude of the reward.

In particular, the authors contend that positive rewards may exacerbate the agency problems within firms by encouraging individuals to self-report. Given that firms must pay higher bonuses to their employees, in order to '*bribe their silence*', collusion becomes less lucrative, and thus less alluring. The higher the reward,

---

[128]An exception is Motta and Polo [2003] who assume that evidence is produced, provided that defection occurs in the market. For instance, the evidence comprises of *scoresheets* tracking the compliance of colluding firms to the illegal agreement. Essentially, this assumption implies that a colluding firm cannot defect in the market and simultaneously self-report to the antitrust authority.

[129]Certainly, a more realistic assumption is that colluding firms can be convicted based on indirect evidence, e.g. circumstantial evidence, however, with a lower probability than the one associated with hard evidence. For example, this may be justified on the fact that it is easier and less costly to prosecute and successfully convict a cartel based on hard evidence rather than on circumstantial evidence.

the higher the compensation a firm must pay to its employees to discourage them from running to the antitrust authority. In addition, a high reward may provide sufficient compensation to the employees for their anticipated reduction in future earnings, in case of dismissal. It is for this reason that positive rewards have a stronger deterrence effect than reduced fines.

Nonetheless, a poorly designed individual leniency programme may give rise to unfavorable effects, such as, the chilling of valuable cooperation between firms or the restriction of beneficial information flows within the firm. For instance, the exchange of information regarding the future prospects of market demand or costs which, in effect, may allow firms to adopt better informed decisions.[130] Moreover, they can cause rigidities in the job market or the internal structure of firms. For instance, colluding firms may increase the duration of their informed employees' contracts, rather than hiring new, and probably more productive, employees. This will be the case if the bribe that has to be paid to new employees is significantly higher than the productivity differential between new and current employees. Moreover, positive rewards to individuals may also induce firms to inefficiently adopt '*innocent attitudes*' in order to avoid raising suspicions to their employees. In particular, firms may strive to prevent their employees from '*sensing trouble*' so as to discourage them from investing their working time into non-productive activities (e.g. sneaking about and acquiring convicting evidence).[131] In addition, employees may strategically exploit the leniency programme by self-reporting even if they do not keep valid evidence. Rewards may also result in deterrence of valuable inter-firm or intra-firm cooperation. This could be the case if prosecution is costly for firms and the antitrust authority erroneously consider a legal conduct as illegal (a type I error). An antitrust authority may, however, mitigate these inefficiencies by introducing fines for information fabrication or coloring.

The authors also investigate plausible explanations for the puzzling fact that firms or individuals keep hard evidence that can be discovered by an antitrust authority. However, the explanations offered cannot justify why firms keep hard evidence in the absence of leniency programmes, or why firms that keep evidence, nevertheless fail to report it to the antitrust authority.

Aubert et al. [2006] contend that firms or individuals may decide to keep hard evidence, running the risk of a higher probability of detection and conviction, to prevent deviations from the collusive agreement, or to mitigate the agency problems within the firm. In particular, the authors argue that firms may want

---

[130]See, for instance, Athey and Bagwell [2001].

[131]A more recent study by Avramovich [2011] shows that the hunt for bounties by individuals may generate intra-firm productive inefficiencies. For instance, employees' attention may be destructed from production towards activities related to gathering evidence on firms' misconduct. The author shows that by rewarding whistle-blowers, the antitrust authority may enhance deterrence at the expense of compromising productive efficiency.

to keep hard evidence so as to decrease the expected fine they have to pay in case of detection by the antitrust authority. The retained evidence can also be used by firms as a disciplining device to threaten to denounce their partners in case of a deviation from the collusive agreement, as in Ellis and Wislon [2003] and Buccirossi and Spagnolo [2006a]. Another plausible explanation is that individuals retain evidence to take advantage of the agency problems within firms. For instance, an individual may keep evidence pertaining to the negotiations to form a cartel in order to increase his bargaining power *vis-à-vis* the manager or the shareholders of the firm. An individual may also keep hard evidence that proves his conformity to the orders of his superiors in order to evade conviction by the antitrust authority, or to threaten to report it in case of deprivation of his bonus, or promotion, or even his employment to the firm.

1.7   Empirical Literature on Leniency Programmes

From the papers discussed above, it is clear that the theoretical literature on leniency programmes is inconclusive with regard to the effectiveness of such incentive schemes in deterring cartels. Some dynamics work in favour of cartel deterrence after the introduction of a leniency programme while others work against it. For example, reducing the fines for self-reporting firms makes collusion less costly ex-ante which fosters collusion.[132] Moreover, when the probability of cartel detection exogenously increases, as a result of a policy change, the option to self-report becomes more valuable to firms.[133] On the other hand, a cartel which is otherwise stable may become unstable after the introduction of a leniency programme. In particular, a leniency programme may reduce the expected fine in case of conviction[134], or increase rivals' costs through revenue based fines and other compliance measures[135], or make the collusive agreement more risky in the sense of Harsanyi and Selten [1988]. Finally, the deterrence effect of leniency programmes hinges on i) the punishment strategies that firms adopt to align their incentives and conform with the illegal agreement[136], ii) the confidentiality of the leniency application procedure[137], iii) the transparency of the leniency programme[138], iv) the fine policy (fixed or proportional to the illegal gains)[139], v) whether the probabilities of

---

[132]Motta and Polo [2003].
[133]Chen and Harrington [2007] and Harrington [2008].
[134]Spagnolo [2004] and Harrington [2008].
[135]Ellis and Wislon [2003].
[136]Ellis and Wislon [2003] and Spagnolo [2004].
[137]Motchenkova [2004].
[138]Hinloopen [2003].
[139]Motchenkova [2004].

detection are endogenous[140] and, vi) the heterogeneity of colluding firms[141].

Because of the inherently clandestine nature of cartels, the empirical analysis of the efficacy of leniency programmes in deterring cartels is difficult. The major challenge involved is the lack of information regarding undetected cartels. The sample selection problem due to observing only the detected cartels may lead to biased estimates and misleading conclusions with regard to the impact of leniency programmes on cartel deterrence.[142]

There are several empirical studies that attempt to measure the impact of leniency programmes on cartel deterrence. The success of such programmes can be measured by their impact on the duration and profitability of cartels.

With regard to the duration of cartels, it is worth noting that the theoretical literature on collusion (see Section 2.2) does not directly consider cartel durations. It rather focuses on the conditions under which collusion can be sustained after a change in market characteristics or a policy innovation (e.g. introduction of a leniency programme). An exception to the literature is Chang and Harrington [2010]. The authors develop a theoretical framework for the evaluation of the impact of a critical policy change, which affects the magnitude of the fines on the duration of cartels (e.g. introduction of a leniency programme). The model predicts that if the policy change is effective, then the average duration of detected cartels should rise in the short-run and fall (below the initial levels) in the long-run. Intuitively, after a policy change (e.g. introduction of a leniency programme) the short-run benefit from defecting exceeds the long-run gains from colluding. Thus, the policy change shifts the 'marginal' cartel (indifferent between competing and colluding) from the population of sustainable and longer-lived cartels, to the population of unstable and shorter-lived cartels. As a result, the ensuing cartel detection comes from a population of longer-lasting cartels. This implies that the average duration of cartels increases in the short-run. In the long-run, the duration decreases (below pre-leniency levels) due to the enhanced overall deterrence. It is important to note that no information regarding undetected cartels is necessary to assess the impact of a change in antitrust policy. The implications of this model are empirically tested by Brenner [2009] and Zhou [2012].

---

[140]Chen and Harrington [2007].

[141]Motchenkova and Laan [2011].

[142]For instance, an increasing number of detected cartels or detection of cartels with shorter duration, after a change in the antitrust policy (e.g. introduction of a leniency programme) may be attributed to an increase in the probability of detection (e.g. increase in the antitrust authority's budget which allows more market investigations) or an increase in the propensity to collude (Ellis and Wislon [2003], Buccirossi and Spagnolo [2006a] and Aubert et al. [2006]). Failing to take these concerns into account may lead to misleading conclusions regarding the effectiveness of a change in antitrust policy (Posner [1970]).

An alternative model, which infers the deterrence effects of leniency programmes from the rate of cartel detection, is suggested by Miller [2009]. Similar to Chang and Harrington [2010], the model developed by Miller [2009] predicts that the number of cartel discoveries increases immediately after the introduction of a leniency programme and subsequently decreases below the initial levels.

With regard to the impact on the profitability of cartels (price-cost margins), the situation is more clear-cut. An effective leniency programme should enhance competition and thus squeeze firms profit margins. Therefore, the hypothesis to be tested empirically is whether the price-cost margins contract after the introduction of a leniency programme.[143] This proposition is empirically tested by Klein [2010] and Cloutier [2011].

Below we survey the most influential contributions to the empirical literature on leniency programmes.

Brenner [2009] empirically investigates the efficacy of the first EU Leniency programme of 1996 for the period 1990-2003. In particular, the author addresses two questions. The first is whether the information revealed under leniency applications i) induces higher fines per-case, compared to traditional prosecution methods, and ii) accelerates the prosecution procedure (as advocated by Motta and Polo [2003]). The second is whether the leniency programme destabilizes existing cartels and deters future cartel activity (as shown by Spagnolo [2004] and Ellis and Wislon [2003]).

The econometric analysis shows that the leniency programme induces a higher level of per-firm fines, controlling for a number of factors that influence the level of the fines. To the extent that the level of the fines is influenced by the strength of the case against the cartel, this confirms that additional and value-added information is revealed to the European Commission which would not have access to without self-reporting. Moreover, the analysis of the data reveals that after the introduction of the leniency programme in 1996, the average duration of a cartel investigation decreased by almost 1.5 years, which hints at the economized investigation and prosecution costs.

The change in cartels' stability is measured by cartels' duration prior and after the introduction of the leniency programme. The author shows that although there is a sharp increase in the number of prosecuted and convicted cartels after adopting the leniency programme, this cannot be interpreted as ample evidence

---

[143]However, to get reliable and unbiased results the empirical analysis should take into account the potential problems due to endogeneity and omitted variables. A reduction in price-cost margins may be due to other policies that concurrently change with the leniency policies. See Klein [2010] and Cloutier [2011].

of the deterrence effects of leniency programmes. The reasoning is twofold. On the one hand, the particular temporal pattern of detections is not consistent with sufficient conditions on the deterrence effects of leniency programmes as established by Miller [2009]. On the other hand, the empirical results of the paper do not corroborate Chang and Harrington [2010] who show that an effective leniency programme increases the average duration of detected cartels in the short-run.

Zhou [2012] criticizes Brenner [2009] because the econometric specification of his paper does not allow the impact of a leniency programme on cartel durations to change over time. As a result, the short-run and long-run effects of leniency programmes, pointed out by Miller [2009] and Chang and Harrington [2010], are confounded, leading to biased estimates.

The paper by Zhou [2012] adapts a dynamic model of cartel formation developed by Chang and Harrington [2010] which predicts that the impact of a more efficacious policy on the duration of discovered cartels is time-dependent. In particular, following an increase in the detection capabilities of an antitrust authority the marginal cartel (indifferent between colluding and competing) immediately collapse and the ensuing cartel detection comes from a population of longer-lasting cartels. Because of such a sample selection effect, the average duration of discovered cartels increases in the short-run; in the long run, the duration decreases due to the enhanced overall deterrence.

The adapted model is then used to evaluate the efficacy of the new EC leniency programme of 2002 for the period 1990-2010. In particular, the author examines whether cartel durations increase immediately following the introduction of the new leniency programme in 2002[144], and whether it falls below the pre-leniency levels, immediately after the change in the leniency policy.

The econometric results of the paper are consistent with the theoretical predictions of Chang and Harrington [2010] that following the introduction of a more efficacious leniency programme, the average duration of discovered cartels increases in the short-run and falls in the long-run below pre-leniency levels.

Miller [2009] develops a dynamic model of cartel behavior that provides predictions and moment conditions regarding the temporal distribution of the number of convicted cartels. The theoretical model predicts that the number of detected cartels increases immediately after the adoption of the leniency programme (because of a higher rate of detection) and decreases in the long-run (because of a lower rate of cartel formation). These effects, which depend on the time horizon, are subsequently used to empirically identify the impact of the amended US cor-

---

[144]The change in the leniency policy in 2002 essentially provides an exogenous shock which identifies the impact of the leniency programme on the duration of detected cartels.

porate leniency programme of 1993, on detection and deterrence capabilities.[145] In particular, the author applies the model to the set of convicted cartels by the DOJ over the period 1985-2005.

The econometric results are consistent with the theoretical predictions suggesting that the number of cartels detected increases immediately following leniency introduction and then falls below the initial levels. Thus, the author concludes that the amended US corporate leniency programme of 1993 enhances both deterrence and detection of cartels.

As Miller [2009] remarks, the results must be interpreted with caution. The predictions of the theoretical model which the author tests empirically is based on the assumption that the probability of detection is equal for all cartels. This remains valid as long as the detected cartels are representative (for the total pool of cartels – detected and undetected) in some fashion. However, a more realistic assumption is that the probability of detection may depend on the level of the cartel overcharges.[146] Moreover, the empirical results of the paper may be biased. This is because the identification of the impact of the leniency programme on detection and deterrence is based on a single time series with only one exogenous policy change.

A common characteristic of the papers discussed above is that they indirectly measure the effectiveness of leniency programmes via the change in the probability of detection or the duration of detected cartels. Below we discuss two papers, by Klein [2010] and Cloutier [2011], that follow a different identification approach. In particular, the effectiveness of leniency programmes is directly measured by their impact on the intensity of competition at the industry level. Since the ultimate goal of leniency programmes is to enhance the desistance and deterrence of cartels, which are harmful for society, a successful leniency policy should have a negative impact on the profitability of firms.

The paper by Klein [2010] studies the effectiveness of leniency policies over a 20 years period from a population of 23 OECD countries. Rather than focusing on cartel detection, the author measures the effect of leniency policies on price-cost margins at industry levels. In particular, the author considers whether a country's specific leniency policy has an impact on the price cost margin of its industries. To avoid potential endogeneity problems and measurement problems from omitted variables, since a change in price-cost margins can be due to other factors than the introduction or amendment of a leniency policy, the author employs the

---

[145]The major innovation of the amended corporate leniency programme of 1993 is that the first self-reporting firm is automatically guaranteed full amnesty, provided that an investigation is not underway. Moreover, amnesty was still available even if an investigation was underway.

[146]See Block et al. [1981], Harrington [2004, 2005] and Houba et al. [2010].

instrumental variables approach. The author uses two instruments: 1) the implementation of leniency programmes in other OECD countries and 2) indicators for the political environment.[147]

The results of the empirical analysis suggest that the national leniency programmes, as well as the EU leniency programmes of 1996 and 2002, have a positive impact on industries' competition intensity, with an approximate decrease in the price-cost margins of 3-5%. Moreover, the instrumental variable estimation reveals that these results do not suffer from significant endogeneity and omitted variable bias. Thus, the author concludes that leniency programmes are efficient in detecting and deterring cartels. Robustness checks indicate that a leniency programme takes on average 1 year after its implementation to become effective, and that its effectiveness proliferates over time. Finally, the legal system wherein the underlying leniency programme is implemented seems to have an important impact on its effectiveness.[148]

Cloutier [2011] criticizes the instrumental variables approach implemented by Klein [2010] arguing that this method may lead to biased estimates because of endogeneity. In particular, the author argues that the EU is a considerably integrated region where many large firms operate at an EU-wide or supranational level; thus they are affected by the leniency policies of other countries. Therefore, other countries leniency policies could be directly related to home country's cartels and their profitability. A similar problem exists with the second instrument used, the policy position of political parties of a country, which is influencing several other policies and not only the leniency policy of a country.

To overcome the potential endogeneity problems, Cloutier [2011]uses the industry concentration as a proxy of the inherent competitiveness of an industry. To the extend that highly concentrated industries are more prone to collusion, Cloutier [2011] suggests that leniency programmes should potentially be more effective in highly concentrated industries.

Thus, he assumes that the leniency policy should be potentially more effective in highly concentrated industries since those industries are more prone to collusion. The implicit assumption is that changes in other policies, but the leniency policy influence all industries regardless of their concentration levels. The author uses the difference-in-differences approach, where the low concentration industries is the control group and the high concentration industries is the treatment group.

---

[147]A good instrument must be correlated with a country's leniency policy, but uncorrelated with any other policy or shock that affects the profitability of the firms operating in the country.

[148]On average, profits are lower in countries that embrace the English and Scandinavian legal systems, and, especially, in countries with the German legal system.

The author uses annual firm-level data of all publicly traded firms in the US over the period 1991 to 1997. The econometric results show that the amendment of the US corporate leniency programme in 1993 had no significant effect on price-cost margins in the short-run while a persistent effect is present after 1-2 years. The results are consistent with Klein [2010].

To conclude this section, the main findings of the existing empirical literature on the effectiveness of leniency programmes are synopsized below:

i)   leniency programmes seem to strengthen the case of the antitrust authority against the cartel - reflected in the higher per-firm fines when a leniency programme is available (Brenner [2009]);

ii)   leniency programmes accelerate the judicial procedure, thus economizing investigation and prosecution costs (Brenner [2009]);

iii)   more transparent leniency programmes (the amendment of the US corporate leniency programme in 1993) seem to have a positive impact on deterrence (Miller [2009]);

iv)   offering full amnesty to the first self-reporting firm enhances the deterrence effect of the leniency programme (compare Brenner [2009] and Miller [2009]);

v)   leniency programmes enhance both desistance and deterrence of cartels (Miller [2009], Zhou [2012] and Klein [2010]).

1.8   Conclusions

Cartels are a symptom of free markets. Because of their inimical effects to consumers, and the economy as a whole, cartels are considered as per-se illegal. Antitrust authorities encounter two key challenges in the war against cartels. The first is the budget stringency which restricts the frequency and success of their investigations. The second is the asymmetry of information which is due to the inherently secret nature of cartels: firms know whether they participate in a cartel while antitrust authorities have only an indication about the existence of the cartel.

A powerful policy instrument to combat illegal cartels is leniency programmes. Leniency programmes are information revelation mechanisms incentivising the instability of cartels. The rationale for using such incentive schemes is twofold: facilitate the detection and/or prosecution of existing cartels and deter the formation of new cartels.

The economic literature on leniency programmes has thrived in the last decade. This survey paper reviews a selective subset of the most influential contributions to the relevant literature. To comprehend the operational mechanism of leniency programmes, and their effects on firms' incentives to collude, this paper reviews

the literature on collusion in Industrial Economics. Moreover, it surveys the literature on generic law enforcement, with a focus on self-reporting schemes, both for individuals and groups of individuals.

The main conclusions derived from the literature on leniency programmes in antitrust, in the light of the research articles reviewed in this paper, are the following:

1. Post-investigation leniency may provide colluding firms with incentives to reveal evidence of significant added value to the antitrust authority. As a result, the judicial procedure accelerates and the investigation costs diminish. These beneficial effects of leniency programmes are more pronounced when an antitrust authority's investigations are not fruitful in detecting incriminating evidence of good quality.

2. The deterrence effects of a leniency programme may be indirectly amplified to the extent that the economized resources for the antitrust authority are used to conduct additional market investigations.

3. The punishment strategy (grim trigger v. stick and carrot) adopted by colluding firms in order to enforce their illegal agreement influence the deterrence effects of leniency programmes.

4. Restricting eligibility to the first self-reporting firm (*first informant rule*) curbs colluding firms' ability to exploit the leniency programme - firms can collude and report in each period, however, all but one has to pay the full fine.

5. Offering a positive reward to the first self-reporting firm, financed by the fines imposed on all other cartel members, provides high-powered incentives to firms in order to self-report (*courageous* leniency programmes). In this case, the leniency programme generates a Prisoners' dilemma situation where all firms run to the antitrust authority to confess their illegal act and benefit from leniency.

6. Extending eligibility to receive leniency to more than one firm when colluding firms retain asymmetric evidence facilitates the prosecutorial procedure of cartels.

7. A more transparent leniency programme allows firms to accurately estimate the benefits accruing from a leniency application.

8. More severe punishment for repeat offenders enhance the efficacy of a leniency programme.

9. A confidential application procedure restricts firms' ability and scope to exploit the leniency programme.

10. A leniency programme for individuals is more effective than a corporate leniency programme.

11. Reduced fines (because of a lenient treatment) make collusion less costly ex-ante (pro-collusive effect).

12. Wrongly designed leniency programmes may provide firms with a credible mechanism to curb the opportunism which is inherent to illegal cartel agreements. Thus, contrary to the policy objectives, a poor leniency programme may exacerbate cartel formation and/or make existing cartels more robust.

13. A poorly designed leniency programme for individuals may have adverse effects on intra-firm hiring strategies and inter-firm benign cooperation.

Chapter 2: Keep to sustain or keep to exploit? Why firms keep hard evidence

2.1 Introduction

In the last decades, several cartels have been dismantled either because some of their members have blown the whistle to antitrust authorities[149] (henceforth "AAs") or because AAs' own investigations have exposed incriminating hard evidence.[150] Moreover, numerous cartels have been prosecuted based on indirect evidence, as the AAs investigations failed to disclose hard evidence.[151]

Intuitively the cartels whose members keep hard evidence have a higher probability of detection and, hence, conviction by an AA. For instance, if an AA conducts a *dawn raid*[152] in a market whereby it detects cogent and irrefutable incriminating hard evidence (e.g., meetings notes/memos/minutes, emails, videos, voice recordings, "*scoresheets*" tracking a cartel's members compliance etc.), it can directly and speedily convict the cartel. On the other hand, if an AA's investigations are not fruitful in tracing hard evidence, then presumably more effort is required to substantiate the case against the cartel, all else equal.

A question that logically arises is why collusive firms do not destroy the hard evidence of their illegal communication to curtail the risk of detection by an AA. A reasonable speculation is that firms keep hard evidence to exchange it with a fine discount on the basis of a leniency programme (henceforth "LP"). However, there are several instances in case law where AAs investigations made clear that firms keep hard self-incriminating evidence even when LPs were not part of the antitrust enforcement policy. This suggests that the existence of a LP may not be the exclusive motive rationalizing a firm's decision to keep hard evidence. Nevertheless, the introduction of a LP arguably influences firms' decisions to keep hard evidence.

---

[149]For instance, see the following cases: Case COMP/E-1/36 604 - *Citric acid*, 5/12/2001; Case COMP/E-1/37.152 – *Plasterboard*, 27/11/2002; Case COMP/E-1/38.069 - *Copper plumbing tubes*, 3/9/2004; Case COMP/C.38.281/B.2 – *Raw tobacco Italy*, 20/10/2005; Case COMP/F/38.899 – *Gas insulated switchgear*, 24/1/2007; Case COMP/E-1/38.823 - *PO/Elevators and escalators*, 21/2/2007; Case COMP/38.628 - *Nitrile butadiene rubber*, 23/1/2008; Case COMP/38511 – *DRAMs*, 19/5/2010.

[150]For instance, see the following cases: Case IV/31.865 - *PVC*, 27/7/1994; COMP/C-38.279 - *French beef*, 2/4/2003; COMP/38.432 - *Professional videotapes*, 20/11/2007; COMP/39165 - *Flat glass*, 28/11/2007; COMP/38.543 - *International removal services*, 11/3/2008; COMP/39125 - *Car glass*, 12/11/2008.

[151]For instance, see the following cases: Cases 89/85, 105/85, 114/85, 116-117/85, 125-129/85 - *A. Ahlström Oy v. Commission*, 27//9/1988; Case IV/34.621, 35.059/F-3 - *Irish Sugar*, 24/5/1997; Case IV/33.133 - *Solvay ICI*, 19/12/1990; COMP/E-1/36.756 - *Sodium gluconate II*, 29/9/2004; COMP/B-2/37.666 - *Dutch beer market*, 18/4/2007.

[152]A '*dawn raid*' is an unannounced inspection at suspected firms' premises. The name is derived from the time of the day when the investigations are normally held.

Pertinent to this is also the fact that those firms that keep hard evidence do not necessarily lodge a leniency application. The relevant case law illustrates several cases where the AAs' investigations were fruitful in detecting and exposing hard evidence in firms hands.

The above discussion brings into the surface several interesting questions in relation to the apparently puzzling behavior of collusive firms. The most critical questions pertain to i) the factors that influence firms' decisions to keep or destroy the hard incriminating evidence, ii) the impact of a LP on the said decisions and, iii) the underlying incentives of firms to keep, nevertheless not report the hard evidence to the AA. Shedding additional light with regard to these questions is the primal objective of this study.

The majority of the papers in the literature on LPs assume that firms axiomatically keep the hard evidence that is generated by the cartel.[153] Nonetheless, there are a few studies that endogenize firms' decisions to keep or destroy the hard evidence. These studies offer plausible explanations for the apparently puzzling questions. For instance, Aubert et al. [2006] maintains that hard evidence can serve as a device to discipline firms to the collusive agreement. In other words, hard evidence can mitigate firms' incentives to defect from the cartel. In the case where one firm defects in the market (e.g. undercut the collusive price), the other firms can run to the AA and self-report. As a result, the defecting firm will pay the full fine with certainty, implying that the benefit from a defection contract. Moreover, a firms that keeps hard evidence and does not self-report can credibly convince its partners that it complies with the collusive agreement. This is because a defecting firm would have strong incentives to self-report to the AA at the same time it defects in the market. Thus, a firm may keep hard evidence in order to dilute any shades regarding its compliance to the agreement, in particular, in situations of uncertainty and imperfect information. Another reasonable explanation offered by Aubert et al. [2006] is that firms keep hard evidence to pay a reduced fine in the case of detection by an AA. In particular, firms may exploit the LP by colluding and systematically self-reporting to the AA. A limitation of the paper by Aubert et al. [2006] is that it cannot justify a firm's decision to keep hard evidence when a LP is not an ingredient of the policy enforcement tools. Furthermore, it cannot provide an adequate answer why firms keep without self-reporting to the AA if the market conduct of individual cartel members can be perfectly monitored (e.g. when the market is sufficiently transparent). A relevant shortcoming of that paper is that it cannot illustrate the impact of a LP on firms' decisions to keep or

---

[153]The existing literature provides answers, among others, as to the optimal level of fine discount that should be offered to self-reporting firms, the number of firms that should be eligible for obtaining leniency and the stage at which leniency should be offered (before or after the commencement of AA's investigation in the market). See Motta and Polo [2003], Spagnolo [2004], Motchenkova [2004] and Harrington [2008]. For an excellent literature review see Spagnolo [2008].

destroy hard evidence and whether to self-report or not to the AA.

Another paper that explores firms' incentives to destroy (partially or fully) or even to create additional evidence is by Silbye [2010]. This paper highlights the trade-off associated with firms' decisions to keep or destroy the hard evidence. On the one hand, more evidence qualifies a self-reporting firm with a higher fine discount (the fine discount is positively related to the quantity of evidence submitted). More evidence, on the other hand, implies a higher probability of detection (and conviction) by the AA, as it is easier for the AA to unravel the cartel if more evidence exists. Interestingly, Silbye [2010] concludes that the option to destroy the hard evidence does not affect the design of the LP.

To tackle some of the open questions in the relevant literature, we develop a model wherein firms' decisions to keep or destroy the hard evidence is endogenous, as in Aubert et al. [2006] and Silbye [2010]. A key assumption of the model is that a cartel's sustainability crucially depends on hard evidence.[154] Specifically, we assume that if at least one cartel member destroys the hard evidence, the cartel may, with some probability, collapse, and thus firms may forego the collusive profits. The underlying reasoning for this assumption is that hard evidence allows more effective implementation and monitoring of the terms of the collusive agreement. We call this effect the *cartel collapse effect* of hard evidence (henceforth the *"CCE"*). This effect may rationalize a firm's decision to keep hard evidence even in the absence of a LP.

The decision to keep hard evidence does not, however, come without a cost for the cartel. Similar to Silbye [2010], we assume that the probability of cartel detection is higher when firms keep hard evidence.[155] If the AA's investigation in the market exposes hard evidence, then the prosecution of the cartel is facilitated. Thus, the AA does not need to undertake additional effort to validate or further substantiate its case against the cartel. We call this effect the *cartel detection*

---

[154] It is worth noting that in Aubert et al. [2006] a cartel's sustainability is not directly influenced by the existence of hard evidence. A cartel, in their model, may collapse if a firm's drastic innovation is successful, as the *successful* firm would prefer to compete forever rather than to collude. Within this context, hard evidence may operate as a mechanism to thwart firms to defect from the cartel. Essentially, by threatening to denounce the cartel, in case of defection from the collusive agreement, the deviant firm's incentive compatibility constraint tightens, as it has to pay the full fine to the AA. Conversely, our model assumes that the hard evidence directly affects a cartel's sustainability by facilitating colluding firms to closely monitor and react to each other's behaviour.

[155] Aubert et al. [2006] also make this assumption. However, contrary to Aubert et al. [2006], we assume that a cartel can be detected even in the absence of hard evidence, based only on indirect evidence. In our view, this is a more realistic assumption. As already remarked, in the real world there are cases where AAs in various jurisdictions have convicted firms for collusion based exclusively on indirect evidence. This assumption would allow us to explore the trade-offs between keeping and destroying the hard evidence, when a LP is not available.

effect of hard evidence (henceforth the "*CDE*"). Intuitively, this effect amplifies the expected cost of the cartel by increasing the expected fine.

Both effects, the CCE and the CDE, cause a contraction to the net expected collusive profits. On the one hand, when firms destroy the hard evidence the cartel's coordination becomes less effective (the cartel collapse with a positive probability) and the gross expected collusive profits decrease. On the other hand, when firms keep the hard evidence, the probability of cartel detection is higher, thus the expected cost of the cartel increases. The strength of the two effects will reveal the conditions under which colluding firms choose to keep or destroy the hard evidence, even though the portfolio of AA's enforcement tools does not include a LP.

Similar to Aubert et al. [2006], the introduction of a LP offers to collusive firms an opportunity to exploit the LP.[156] In particular, firms may keep the hard evidence and subsequently exchange it with a reduced fine within the framework of a LP. We call this effect the *cartel amnesty effect* of hard evidence (henceforth the "*CAE*"). This effect causes a reduction to the expected cost associated with a firm's decision to keep hard evidence. Therefore, a firm may choose to keep hard evidence not only to evade the costly breakdown of the cartel but also to exploit the LP. The interplay of the three effects, CCE, CDE and CAE, allows us to explore the impact of a LP on firms' decisions to keep or destroy hard evidence. Moreover, the introduction of a LP allows the investigation of the conditions under which collusive firms keep hard evidence without disclosing it to the AA.

The main conclusions of our study is that firms are more likely to keep hard evidence when: *i*) the cartel's sustainability is more sensitive to hard evidence, *ii*) the probability of cartel detection is less depended on hard evidence and, *iii*) the collusive profits-fine ratio is high. Moreover, we show that the introduction of a LP sharpens firms' incentives to keep hard evidence. Finally, we demonstrate that firms may keep hard evidence without reporting it to the AA if the cartel's sustainability is very sensitive to hard evidence and the (aggregate) probability of cartel detection is sufficiently low.

The paper is organized as follows. Section 2.2 outlines the model. Section 2.3 presents the baseline model without LP. The extended model with LP is presented in Section 2.4. Section 2.6 concludes.

---

[156]Motta and Polo [2003] are the first to show that firms may strategically exploit the leniency programme in order to pay lower fines. This potential has been confirmed in a lab experiment conducted by Hinloopen and Soetevent [2008]. The experiment confirms that firms collude and report systematically when the LP is sufficiently generous. For a more general discussion on LPs exploitation see Wils [2008], p. 137.

2.2 Outline of the model

Similar to Motta and Polo [2003], we consider two firms that play an infinitely repeated game in the presence of an AA which enforces antitrust law. This is a significant difference to Aubert et al. [2006] who study a dynamic non-repeated game. In effect, the authors assume that the hard evidence is indelible so that firms can use it in any future period, as long as the cartel has not been detected. On the contrary, we assume that there is full information decay after the end of each period. The repeated structure could be justified on the ground that the collusive agreement has limited time frame due to changing market conditions. For instance, the product becomes obsolete and thus firms shift to other markets which guarantee more lucrative opportunities for trade.

To simplify the analysis, we assume markets where cartel formation is always profitable for firms. We call a market with this feature a *pro-collusive* market.[157] This study does not aim to analyze the deterrence and/or the desistance effects of LPs. The main goal is to concentrate on the impact of a LP on firms' incentives to keep or destroy hard evidence, and their incentives to report or not the retained evidence to the AA. Therefore, the assumption for a pro-collusive market allows to focus on firms' decisions after establishing a cartel, rather than on their decisions to form a cartel a such.

We assume that collusion generates and leaves symmetric hard incriminating evidence to the two colluding firms. In addition, we assume that the existence of a cartel cannot be observed by the AA unless the latter conducts an investigation in the market. The AA's investigation is successful only with some probability. This probability crucially depends on whether hard evidence exists. Destroying the hard evidence does not necessarily imply that firms will escape unscathed by the AA's investigations. The cartel can be detected and successfully prosecuted based on indirect evidence; evidence of facts and circumstances supporting an inference of a cartel.[158] The probability of detection if both firms destroy the hard evidence is denoted by p. We call this probability the *base probability of detection*.[159]

If the two collusive firms decide to keep the hard evidence, the probability of detection increases to $p + \Delta$. The latter is called the *aggregate probability of detection*. Parameter Δ represents the incremental probability of cartel detection when

---

[157]Technically the existence of a pro-collusive market is guaranteed by Assumptions 2 and 4, i.e., $\delta \geq \frac{1}{2}$ and $\frac{\pi}{F} \geq 1$, respectively, of our paper.

[158]For example, this evidence could include market data on prices or a sudden change in firms conduct not justified otherwise by the market conditions.

[159]This probability is constant and independent of firms' pricing behaviour (as in the seminal papers by Motta and Polo [2003] and Spagnolo [2004]). A more realistic assumption would be that this probability is an increasing function of the collusive price; a higher price (or price cost margin) may attract more attention by the law enforcement agency, and thus make detection more likely (as in Harrington [2004, 2005] and Houba et al. [2010]).

both firms choose to keep rather than to destroy the hard evidence. The parameter has a two-fold interpretation. It may reflect 1) the quality of hard evidence or 2) the AA's efficacy in prosecuting cartels when its investigations expose hard evidence. The better the quality of hard evidence is or the more effective the AA is in assessing the hard evidence detected, the higher the incremental probability of detection would be. We assume that Δ is exogenous.[160]

If only one firm chooses to keep hard evidence, the incremental probability of detection is $\frac{\Delta}{2}$.[161] The underlying reasoning is that the AA may, in that case, need to further investigate the market in order to corroborate the quality of the reported evidence. In the extreme case where the AA's competency in detecting cartels is not influenced by the existence of hard evidence or the quality of the hard evidence is unreliable, then $\Delta = 0$.

To sum up, the probability of detection when both firms destroy the hard evidence is p, when both firms keep the hard evidence is $p + \Delta \in [0, 1]$, and when only one firm keeps the hard evidence is $p + \frac{\Delta}{2}$.[162]

In case of detection by the AA, each cartel member must pay a fine $f \in [0, F]$, where F is the maximum feasible fine.[163] A typical result in the literature is that the optimal level of law enforcement is attained by setting the fine at its maximal level.[164] Thus, we assume that $f = F$. Both p and F are enforcement policy instruments which are exogenously fixed. We define those instruments as the *antitrust enforcement policy*.[165] A critical assumption we make is that the antitrust policy is stationary and that it treats all markets alike.[166]

---

[160]Silbye [2010] assumes that colluding firms can choose how much evidence to keep, and thus probability Δ is, in effect, endogenous in his study.

[161]A similar assumption is made by Aubert et al. [2006]. However, these authors, contrary to this paper, assume that the base probability of detection is zero, that is, a cartel cannot be uncovered without hard evidence. As hinted before (see footnote 155), this normalization implies that firms would always keep the hard evidence when a LP is not available.

[162]The upper bound on p reflects our assumption that a cartel can be detected even if all of its members destroy the hard evidence albeit with a weakly lower probability to the one associated when hard evidence exists.

[163]To simplify the exposition, we assume that the fine is constant and independent of the nature and/or severity of the illegal cartel. In practice, however, the antitrust fines are, directly or indirectly, related to the accrued (net) illegal profits or the damages instigated to society as a result of the law violation. Nonetheless, the maximum feasible fine is capped so that the colluding firms' financial stability is not jeopardized.

[164]See the excellent survey by Polinsky and Shavell [2000].

[165]The probability of detection p is a determinative factor of the deterrence effect of the enforcement of antitrust policy. The limited resources available to AAs condition the level of this probability. For empirical estimation of the probability of detection see Bryant and Eckard [1991], Combe et al. [2008] and Ormosi [2011].

[166]In reality, a convicted cartel would stay under close scrutiny by the antitrust authority, making it hard to collude for at least some time. Nonetheless, it is not unrealistic to assume that firms could start colluding again at some future point (as it is assumed by Motta and Polo

When the enforcement policy encompass a LP, the collusive firms that keep hard evidence have an additional decision to make. In particular, they may choose either to exchange the hard evidence with lenient treatment or to keep the hard evidence without revealing it to the AA. Given that the hard evidence is symmetric only the first self-reporting firm is eligible for leniency (*first informant rule*). Moreover, we assume that the first self-reporting firms receives full amnesty. When both firms simultaneously apply for leniency, then each of them must pay an expected fine of $\frac{F}{2}$.

Hard evidence is of pivotal importance for the successful implementation of the cartel. On the one hand, it may allow firms to overcome the potential administration and implementation problems due to the complexity of their illegal agreement. More precisely, we assume that if at least one firm destroys the hard evidence, the management of the cartel is rendered more demanding and challenging, ultimately causing, with some probability, the collapse of the cartel. We denote this probability by $\beta \in [0, 1]$ (henceforth "*the cartel collapse probability*"). Parameter β measures the sensitivity of the cartel to hard evidence. A higher β implies that the cartel agreement is more complex (for instance, in terms of administration, allocation of duties and tasks) and requires the existence of detailed hard evidence to overcome the administration deficit and/or implementation problems it encounters. At the extreme case where $\beta = 0$, the cartel's sustainability (administration and implementation) is independent of hard evidence. Furthermore, if only one firm keeps the hard evidence, then again the cartel may collapse with the same probability β. Although the firm that keeps the hard evidence can administer the agreement, the one that destroys may fail, and thereby the cartel may, as before, collapse with probability β.[167]

We further assume that if the cartel is implemented, the colluding firms can perfectly monitor the conformity to the agreement. For instance, the market is sufficiently transparent so that firms can instantly react to any defection from the agreement.

Assumption 1. *Provided that the collusive agreement is implemented firms instantly react to any defections in the market when hard evidence exists.*

We concede that this is a very strong assumption to be made. However, it allows us to simplify the analysis and focus our attention on deviations from the individual terms of the agreement other than those referred to firms' market conduct.[168]

[2003]). To simplify the analysis, we assume that in the period following a conviction by the AA, firms can re-establish a cartel.

[167] A popular song written by Al Hoffman and Dick Manning and released in 1952 reminds us that it "*Takes Two to Tango*". The 'tango' metaphor remind us that for some thing in life, more than one party is involved in an inextricable manner.

[168] The study by Aubert et al. [2006] is more general and thus does not put the emphasis on

Given that we assume a pro-collusive market, when the enforcement policy does not include a LP, firms have only one decision to make: either to keep or destroy the hard evidence. When a LP is available, collusive firms have two sequential decisions to make: either to keep or destroy, and if they keep, either to report or not the hard evidence to the AA. We explore these strategy profiles and the corresponding payoffs in Sections 2.3 and 2.4.

To simplify the exposition we normalize non-collusive profits to zero. Per-firm collusive profits are denoted by $\pi \in (0, \pi^M]$. The upper bound denotes the per-firm monopoly profits. Moreover, we assume that firms have the same discount factor, denoted by δ, and adopt standard grim trigger strategies to sustain their illegal agreement. To simplify the analysis we assume that $\delta \geq \frac{1}{2}$.[169]

Assumption 2. $\delta \geq \frac{1}{2}$.

Another critical assumption of the paper is that a cartel that collapses due to administration problems avoids prosecution.

Assumption 3. *An unsuccessful cartel avoids prosecution.*

Essentially, unsuccessful cartels fail to coordinate their conduct and obtain collusive profits.[170]

2.3 Baseline Model: without Leniency Programme

This section analyzes the baseline model where the portfolio of antitrust enforcement tools does not include a LP. Given that we assume a pro-collusive market, the two firms have only one decision to make, namely, to keep or destroy the hard incriminating evidence generated by the cartel. The analysis of this scenario allows to rationalize firms' decisions to keep hard evidence even though a LP is not available. Moreover, it provides a benchmark to analyze the impact of a LP on the said incentives of firms. Before deriving the subgame perfect equilibrium (henceforth "SPE"), we formally present the game; timing, strategies and payoffs.

2.3.1 The timing of the stage game

In each period, the timing of the game is as follows:

---

this aspect associated with firms' decisions to keep hard evidence.

[169]The lowest bound to δ is the one we obtain in the standard Bertrand duopoly game where firms sustain collusion by adopting grim trigger strategies in a setting without antitrust enforcement. See Friedman [1971] and Tirole [1988].

[170]In practice, firms entering a discussion to establish a cartel, even if the discussions are fruitless, are liable for a breach of competition law.

- STAGE 0 (POLICY DESIGN): The AA commits and announces a certain enforcement policy (p and F).

- STAGE 1 (CARTEL FORMATION): Having observed the policy parameters and having full and complete information about parameters Δ and β, firms decide whether to communicate and form a cartel. Hard evidence is produced.

- STAGE 2 (KEEP OR DESTROY): Firms simultaneously decide whether to keep or destroy the hard evidence of their illegal agreement.

- STAGE 3 (POLICY IMPLEMENTATION): The AA launches an investigation into the market. The success of its investigation hinges on the existence of hard evidence by firms.

2.3.2 Firms' strategies and payoffs

For the purposes of our analysis, we consider only the two symmetric strategy profiles: (KEEP, KEEP) and (DESTROY, DESTROY).[171]

First, consider the strategy profile (KEEP, KEEP). According to this strategies both firms keep the hard evidence of their illegal communication. With probability $p + \Delta$ firms obtain the collusive profits π but have to pay a fine F. With the complement probability, that is, with $1 - p - \Delta$, firms evade detection and obtain the collusive profits π. If no deviation occurs, the game is repeated forever and the collusive agreement is stable. The expected discounted value of this strategy profile is equal to

$$V_K = \frac{\pi - (p + \Delta)F}{1 - \delta}$$

Notice that V_K is non-negative if $\frac{\pi}{F} \geq p + \Delta$.

If a deviation occurs, that is, if one firm destroys the hard evidence instead, the cartel will encounter administration problems. As a consequence, the cartel collapses with probability β. At the same time, given that only one firm keeps the hard evidence, the probability of cartel detection decreases to $p + \frac{\Delta}{2}$. Thus, the short-run gain associated with a deviation from the strategy prescription is the lower expected fine, given that the probability of detection decreases. This gain

---

[171]There is also an asymmetric strategy profile that firms may coordinate on. In particular, firms may agree that one of them will destroy, while the other will keep the hard evidence, i.e., coordinate on (DESTROY, KEEP). The expected discounted payoff associate with this profile is $V_{KD} = \frac{(1-\beta)}{1-\delta}[\pi - (p + \frac{\Delta}{2})F]$. However, this strategy profile is always Pareto dominated by the symmetric ones. Notice that $V_K = \frac{\pi-(p+\Delta)F}{1-\delta} > \frac{(1-\beta)}{1-\delta}[\Pi - (p + \frac{\Delta}{2})F] = V_{KD}$ and $V_D = \frac{(1-\beta)(\Pi-pF)}{1-\delta} > \frac{(1-\beta)}{1-\delta}[\Pi - (p + \frac{\Delta}{2})F] = V_{KD}$. Therefore, to simplify our analysis, we discard the possibility that firms coordinate on asymmetric strategies.

will be realized by both firms only if the cartel does not collapse, that is, with probability $1 - \beta$. With probability β the cartel collapses and both firms obtain zero profits. After a deviation from the strategy prescription, firms revert to a permanent punishment phase where they obtain zero profits forever. Therefore, the expected payoff for the deviant firm is equal to

$$V_K^d = (1 - \beta)[\pi - (p + \frac{\Delta}{2})F]$$

(the superscript stands for deviation).

By Assumption 1, and since the cartel is implemented with certainty, any defection in the market is instantly detected and punished. Thus, such a deviation is inferior to the one described above.

Consider the strategy profile (DESTROY, DESTROY). According to this, both firms destroy the hard evidence of their illegal communication. Thus, the cartel collapses with probability β. Moreover, the cartel is detected with probability p, since no hard evidence exists, in which case the firms must pay a fine F. If no deviation occurs, then the game is repeated forever and the collusive agreement is stable. The expected discounted payoff of this strategy profile is

$$V_D = \frac{(1 - \beta)(\pi - pF)}{1 - \delta}$$

Notice that V_D is non-negative if $\frac{\pi}{F} \geq p$.

If a firm instead keeps the hard evidence, then the probability of cartel detection increases to $p + \frac{\Delta}{2}$. The cartel remains fragile and collapse with probability β since one firm, the one that conforms to the strategy prescription, destroys the hard evidence. Given that a deviation occurs, firms revert to a permanent punishment phase wherein they obtain zero profits forever. Therefore, the expected payoff for the deviant firm is equal to

$$V_D^d = (1 - \beta)[\pi - (p + \frac{\Delta}{2})F]$$

By the same line of reasoning, as elucidated before, a simultaneous deviation in the market, provided that the cartel is implemented, is inferior to the one described above.

At the beginning of this section, we emphasize that we consider a pro-collusive market. To guarantee this we need to restrict V_K and V_D to be non-negative. For this we need to restrict the fine so that it is bounded above by collusive profits.[172]

---

[172]Notice that the condition that is required so that the strategy profile (KEEP, KEEP) results in a non-negative payoff is more stringent than the one required for (DESTROY, DESTROY).

Assumption 4. $\frac{\pi}{F} \geq 1$.

Effectively, this assumption implies that firms always find it profitable to collude irrespective of the strictness of antitrust policy. This assumption could also be justified on the basis that the fine level must not jeopardize firms' financial stability.

2.3.3   Equilibrium Analysis

The game we have to analyze is a repeated simultaneous move game with imperfect information. Thus, the appropriate solution concept is subgame perfect equilibrium (henceforth "SPE").[173] We only consider pure strategies. When multiple equilibria exist we apply the Pareto dominance criterion.[174] Given that in the scenario that we consider, firms communicate with the aim to reach a collusive agreement it is not unrealistic to assume that firms agree to orchestrate their actions by coordinating to a particular collusive strategy.

Solution of the baseline game   Suppose that the two firms agree to keep the hard evidence of their illegal communication, i.e., coordinate on (KEEP, KEEP). This strategy profile is a SPE if and only if (henceforth "iff") $V_K \geq V_K^d$. The latter inequality holds iff $p \leq p_1 \equiv \frac{\pi}{F} - \frac{(\delta+\beta-\beta\delta+1)\Delta}{2(\delta+\beta-\beta\delta)}$. Observe that $p_1 > 0$ iff $\Delta < \Delta_1 \equiv \frac{\pi}{F}\frac{2(\delta+\beta-\beta\delta)}{(\delta+\beta-\beta\delta+1)}$ and $p_1 < 1-\Delta$ iff $\Delta > \Delta_2 \equiv (\frac{\pi}{F}-1)\frac{2(\delta+\beta-\beta\delta)}{(1-\beta)(1-\delta)}$. Hence, if the additional probability of detection is sufficiently low ($\Delta < \Delta_2$), then (KEEP, KEEP) is a strict SPE. If this probability is sufficiently high ($\Delta > \Delta_1$), then (KEEP, KEEP) is not a SPE.

Lemma 1. *(*KEEP, KEEP*) is a SPE of the Keep-Destroy game if* $p \in [0, \min\{p_1, 1-\Delta\}]$.

Notice that p_1 increases in β and $\frac{\pi}{F}$, while it decreases in Δ. The intuition is the following. As the cartel becomes less sensitive to hard evidence then, *ceteris paribus*, the deviation gain from the strategy profile (KEEP, KEEP) rise. Moreover, the expected foregone collusive profits due to the collapse of the cartel increase, *ceteris paribus*, in $\frac{\pi}{F}$. Thus, an increase either in β or $\frac{\pi}{F}$ makes the *collapse effect* (CCE) sharper. In the case where $\frac{\pi}{F}$ is sufficiently high, and in particular if $\frac{\pi}{F} > \frac{(\beta+\delta-\beta\delta+1)\Delta}{2(\beta+\delta-\beta\delta)}$, then the strategy profile (KEEP, KEEP) is a strict NE. On

---

[173] A strategy is optimal in the sense of SPE if it maximizes players' payoff for every period and every history of the game.

[174] A strategy satisfies this criterion if there is no other strategy for which each player of the game has a strictly higher payoff. Such a strategy is also called payoff-dominant. The notion of Pareto-dominant equilibrium is well established in the literature. See [Fudenberg and Tirole, 1991a, pp. 20-22].

the contrary, an increase in the additional probability of detection, causes, *ceteris paribus*, the deviation profits to shrink since the expected fine increases. Therefore, the *detection effect* (CDE) intensifies with a higher Δ.

Suppose now that the two firms agree to destroy the hard evidence, that is, coordinate on (DESTROY, DESTROY). A unilateral deviation from this strategy profile is not profitable if $V_D \geq V_D^d$. The latter inequality holds iff $p \leq p_2 \equiv \frac{\pi}{F} + \frac{(1-\delta)\Delta}{2\delta}$. Assumption 4 implies that $p_2 \geq 1$, and hence (DESTROY, DESTROY) is a strict SPE. This is due to the fact that if a firm keeps hard evidence, rather than destroying it, the probability of detection increases by $\frac{\Delta}{2}$, while the stability of the collusive agreement remains intact. Since one of the two firms complies with the prescription of the agreed strategies, and accordingly destroys the hard evidence in every period, the cartel continues to collapse with probability β. Thus, such deviation does not enhance the sustainability of the cartel. On the contrary, it increases the expected fine, and thus the expected cost, of the cartel. Moreover, given that a deviation from the agreed strategies occurs, firms are deprived of future collusive profits, as they enter in an eternal punishment phase.

Lemma 2. *(*DESTROY, DESTROY*) is a strict SPE of the Keep-Destroy game.*

For the set of parameter values of p and Δ where the *Keep-Destroy* game has two SPE we apply the Pareto criterion. Thus, (KEEP, KEEP) Pareto dominates (DESTROY, DESTROY) iff $V_K > V_D$. The latter inequality holds iff $p < p_3 \equiv \frac{\pi}{F} - \frac{\Delta}{\beta}$.

Lemma 3. *(*KEEP, KEEP*) Pareto dominates (*DESTROY, DESTROY*) iff $p < p_3$, where $p_3 \equiv \frac{\pi}{F} - \frac{\Delta}{\beta}$.*

Considering that (i) $p_3 > 0$ iff $\Delta < \Delta_3 \equiv \frac{\beta\pi}{F}$ and (ii) $p_3 < 1 - \Delta$ iff $\Delta > \Delta_4 \equiv \frac{\beta}{1-\beta}(\frac{\pi}{F} - 1)$ and Lemmata 3, 5 and 3 we can state the main result of this section.

Proposition 1. *For given policy and other parameters (p, Δ, F, β, π) the Pareto dominant SPE of the baseline model without a LP is:*

1. *(*KEEP, KEEP*) iff:*

   (a) $\frac{\pi}{F} \in [1, \frac{1}{\beta}]$ *and*
       i. $\Delta \in [0, \Delta_4]$ *and* $p \in [0, 1 - \Delta]$ *or*
       ii. $\Delta \in [\Delta_4, \Delta_3]$ *and* $p \in [0, p_3)$ *or*
   (b) $\frac{\pi}{F} > \frac{1}{\beta}$, $\Delta \in [0, 1)$ *and* $p \in [0, 1 - \Delta]$

2. *(*DESTROY, DESTROY*) iff $\frac{\pi}{F} \in [1, \frac{1}{\beta}]$ and*

(a) $\Delta \in [\Delta_4, \Delta_3]$ *and* $p \in [p_3, 1 - \Delta]$ *or*

(b) $\Delta \in (\Delta_3, 1)$ *and* $p \in [0, 1 - \Delta]$

where $p_3 \equiv \frac{\pi}{F} - \frac{\Delta}{\beta}$, $\Delta_3 \equiv \beta \frac{\pi}{F}$ *and* $\Delta_4 \equiv \frac{\beta}{1-\beta}(\frac{\pi}{F} - 1)$.

Proof. See Appendix A. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Proposition 1 can be intuitively explained as follows. Firms keep hard evidence, rather than destroying it, if the base probability of detection is relatively low, and in particular lower than the critical value p_3. The latter balances the CCE, that is, the expected foregone collusive profits if firms destroy the hard evidence $(\beta \cdot \pi)$ and the CDE, that is, the incremental expected fine due to keeping hard evidence $(\Delta \cdot F)$.[175] The critical value p_3 is positively related to the collusive profits-fine ratio and the cartel collapse probability and negatively related to the additional probability of detection due to retaining hard evidence. Thus, an increase either in the collusive profits-fine ratio or in the cartel collapse probability makes it more likely that firms would refrain from destroying the hard evidence.

Consider the effects of an increase in the cartel collapse probability. If the firms destroy the hard evidence, then the cartel's expected profits shrink, since the cartel collapses with a higher probability. At the same time, if firms keep the hard evidence, the expected collusive profits are intact. In other words, the CCE is reinforced, while the CDE remains unchanged. An increase in collusive profits also strengthens the CCE, since firms forego higher future profits when the cartel collapses due to a defection from the agreed strategies. Moreover, the CDE dehydrates since the net gain from keeping hard evidence increases. A similar argument holds when the fine decreases. To sum up, firms are more inclined to keep hard evidence both with an increase in the cartel collapse probability and an increase in the collusive profits-fine ratio. These effects are illustrated in Figure 1 below.

When the cartel collapse probability is zero, that is, the sustainability of the cartel is insensitive to hard evidence (i.e., $\beta = 0$), firms always destroy the hard evidence. Retaining evidence does not bring about any benefit in that case. On the contrary, when the cartel collapse probability in the absence of hard evidence is sufficiently high (i.e., $\beta > \frac{1}{2}$), the foregone profits linked to firms' decisions to destroy the hard evidence are considerably high. In that case, the CCE dominates the CDE. Thus, firms keep hard evidence to sustain the cartel, although they run the risk of a higher probability of detection.

While an increase in the additional probability of detection does not influence the CCE, such an increase bolsters the CDE. Thus, firms are more likely to destroy

---

[175]Notice that p_3 can can be expressed as $p_3 = \frac{1}{\beta F}(\beta\pi - \Delta F)$. The first term in the parenthesis captures the CCE, while the second term captures the CDE.
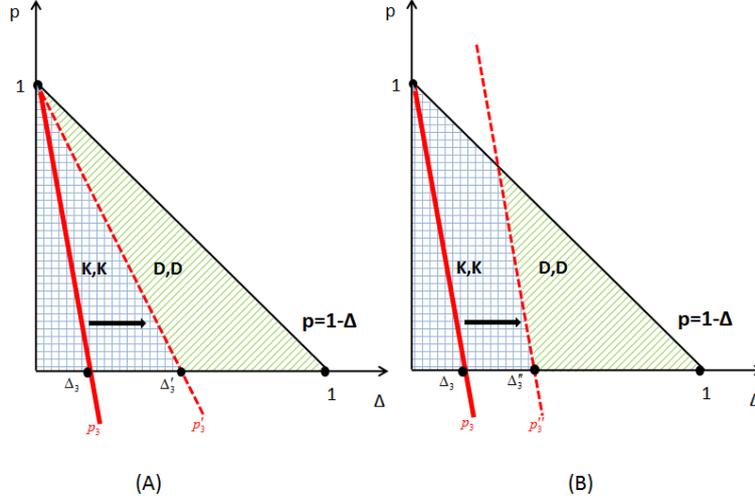
Figure 1: Without LP. Graph A: An increase in β causes a rotation of p_3 to the right to p_3'. Graph B: An increase in $\frac{\pi}{F}$ causes a parallel shift of p_3 to the right to p_3''. Both graphs are drawn with the initial assumption that $\frac{\pi}{F} = 1$.

the hard evidence. When this probability is sufficiently low ($\Delta < \Delta_5$), then the incremental fine associated with firms decisions to keep hard evidence is very low. Thus, the CDE dominates the CCE, making firms to keep the hard evidence in order to enhance the stability of the cartel. Conversely, when Δ is sufficiently high ($\Delta > \Delta_3$), then the incremental fine is very high, the CDE dominates the CCE, inducing firms to destroy the hard evidence (see Figure 1, Graph B).

Corollary 1. *Firms are more likely to keep the hard evidence with a higher β and a lower Δ. In the extreme case where $\beta = 0$, firms always destroy the hard evidence. If $\beta > \frac{1}{2}$ firms always keep the hard evidence.*

2.4 Extended model: with Leniency programme

In this section, we extend the baseline model so that the portfolio of policy instruments includes a LP. This extension allows to shed some light on two crucial decisions firms have to make after forming a cartel, namely, to keep or destroy the hard evidence and whether to report or not the evidence to the AA. Moreover, it enables to investigate the conditions under which firms keep hard evidence albeit not reveal it to the AA.

Within this richer framework, the cartel may be detected either by an AA's own investigations or after a cartel member self-reports to the AA, before an investigation is initiated in the market.

2.4.1 The timing of the stage game

In each period the timing of the game is as follows:

- STAGE 0 (POLICY DESIGN): The AA commits and announces a certain enforcement policy which includes a LP.

- STAGE 1 (CARTEL FORMATION): Having observed the enforcement policy and having full and complete information about parameters Δ and β, firms decide whether to communicate and form a cartel. Hard evidence is produced.

- STAGE 2 (KEEP OR DESTROY): Firms simultaneously decide whether to keep or destroy the hard evidence of their illegal agreement. If firms choose to keep the hard evidence then move to STAGE 3. If firms destroy, the hard evidence then move to STAGE 3 with probability $1-\beta$. With probability β the cartel collapse and firms obtain zero profits. The game restarts from STAGE 0.

- STAGE 3 (REVELATION OF HARD EVIDENCE): Firms simultaneously decide whether to report or not the hard evidence to the AA.

- STAGE 4 (POLICY IMPLEMENTATION): If at least one firm applies for leniency, then the cartel is detected. If no firm applies for leniency then the AA initiates an investigation into the market. The success of its investigation hinges on the existence of hard convicting evidence.

2.4.2 Firms' strategies and payoffs

For the purposes of our analysis, we consider only the three symmetric strategy profiles: (DESTROY, DESTROY), (KEEP AND NOT REPORT, KEEP AND NOT REPORT) and (KEEP AND REPORT, KEEP AND REPORT).[176]

To begin with, consider the strategy profile (DESTROY, DESTROY). The two colluding firms destroy the hard evidence. As shown in Section 2.3.2, the expected discounted payoff associated with this strategy profile is $V_D = \frac{(1-\beta)(\pi-pF)}{1-\delta}$. A

---

[176]There are also other possible (asymmetric) strategy profiles that firms may coordinate on: (DESTROY, KEEP AND REPORT), (DESTROY, KEEP AND NOT REPORT) and (KEEP AND REPORT, KEEP AND NOT REPORT). (DESTROY, KEEP AND NOT REPORT) is Pareto dominated by (DESTROY, DESTROY), as $V_D = \frac{(1-\beta)(\Pi-pF)}{1-\delta} > \frac{(1-\beta)[\Pi-(p+\frac{\Delta}{2})F]}{1-\delta} = V_{D,KNR}$. (DESTROY, KEEP AND REPORT) is Pareto dominated by (KEEP AND REPORT, KEEP AND REPORT), as $V_{KR} = \frac{\pi-\frac{F}{2}}{1-\delta} > \frac{(1-\beta)(2\pi-F)}{2(1-\delta)} = V_{D,KR}$. Note also that the strategy profile (Keep and Report, Keep and Not Report) yields the same payoff as the strategy profile (Keep and Report, Keep and Report). This is intuitive given that only one firm is eligible to receive amnesty.

deviant firm that keeps the hard evidence, rather than destroying it, has two options; either to report the evidence to the AA and obtain full amnesty or retain the evidence without self-reporting it to the AA. In the former case, the deviant firm receives an expected payoff equal to $V_D^{(d,r)} = (1-\beta)\pi$, while in the latter case its expected payoff is $V_D^{(d,nr)} = (1-\beta)[\pi - (p + \frac{\Delta}{2})F]$. Clearly, a deviant firm will never keep hard evidence without reporting it to the AA, as $V_{KNR}^{(d,r)} > V_{KNR}^{(d,nr)}$. An analogous argument to the baseline model holds with regards to defections in the market.

Remark 1. *A firm that deviates from (*DESTROY*,* DESTROY*) will never keep hard evidence without reporting it to the AA.*

Consider now the strategy profile (KEEP AND NOT REPORT, KEEP AND NOT REPORT). The two colluding firms keep the hard evidence, nonetheless, not report it to the AA.[177] Given that both firms keep hard evidence, the cartel is detected with probability $p + \Delta$, in which case firms must pay a fine F. The expected discounted payoff of this strategy profile is $V_{KNR} = \frac{\pi - (p+\Delta)F}{1-\delta}$. Two types of deviation from this strategy profile are possible. A firm may deviate at the *revelation* stage and report the hard evidence to the AA. In that case, the deviant firm receives full amnesty. Given that the cartel does not collapse, since the two firms continue to keep hard evidence, the deviation payoff is $V_{KNR}^{(d,r)} = \pi$. If instead a deviation occurs at the *keep-destroy* stage, whereby the deviant firm destroys the hard evidence, the probability of detection decreases to $p + \frac{\Delta}{2}$. However, the cartel may in that case collapse with probability β. Thus, the expected deviation payoff is $V_{KNR}^{(d,nr)} = (1-\beta)[\pi - (p + \frac{\Delta}{2})F]$. Clearly, a deviant firm will never destroy the hard evidence, given that $V_{KNR}^{(d,r)} > V_{KNR}^{(d,nr)}$.

A similar argument to the benchmark model without LP holds for a deviation in the market. Therefore, a simultaneous deviation in the *revelation* stage or the *keep-destroy* stage with a deviation in the market is inferior, in terms of profitability, to the deviations described above.

Remark 2. *A firm that deviates from (*KEEP AND NOT REPORT*,* KEEP AND NOT REPORT*) will never destroy the hard evidence.*

Lastly, consider the strategy profile (KEEP AND REPORT, KEEP AND REPORT). The two colluding firms agree to keep hard evidence and subsequently report it to the AA. In this case, the cartel is detected with certainty and firms

---

[177](KEEP, KEEP) from Section 2.3 is equivalent to (KEEP AND NOT REPORT, KEEP AND NOT REPORT). However, the deviations from this strategy profile are different. A firm that deviates from (KEEP AND NOT REPORT, KEEP AND NOT REPORT) can now report the hard evidence. This option is not available to firms when a LP is not a part of the portfolio of enforcement tools.

pay an expected fine $\frac{F}{2}$. The expected discounted payoff of this strategy profile is $V_{KR} = \frac{\pi - \frac{F}{2}}{1-\delta}$. Again, two types of deviations may occur from this strategy profile. A firm may deviate at the *keep-destroy* stage and instead of keeping the hard evidence destroy it. In this case, the cartel collapses with probability β. At the same time, the cartel is detected with certainty given that the compliant firm keeps and reports the hard evidence to the AA. The deviant firm's payoff is $V_{KR}^d = (1 - \beta)(\pi - F)$. Clearly, a deviation from this strategy does not bring about any benefit, as such deviation radically increases the probability of paying a fine and, at the same time, it deprives the deviator, as well as the compliant firm, all future collusive profits. This is because firms defect from the agreed strategy profile. The second type of deviation may occur at the *revelation* stage. Trivially, any deviation at this stage will be unprofitable as the deviant firm pays the full fine F with certainty, rather than an expected fine of $\frac{F}{2}$. Besides, firms are deprived of future collusive profits as they enter a permanent punishment phase. The same line of reasoning for the simultaneous deviation in the market holds as for the strategy profiles discussed above.

2.4.3 Equilibrium Analysis

The game we have to analyze in this section is a repeated sequential move game with imperfect information. Contrary to the game in Section 2.3, the firm that chooses to keep the hard evidence at the *Keep-Destroy* stage has an additional decision to make at the *revelation* stage: either to report the hard evidence to the AA or not.

As in Section 2.3, the appropriate solution concept is SPE. We only consider pure strategies. When multiple equilibria exist we apply the Pareto dominance selection criterion.

***Revelation subgame* - Report v. Not Report** This subgame is reached if firms have already chosen to keep the hard evidence at the *Keep-Destroy* subgame. Firms' decisions are either to report the hard evidence to the AA or keep the hard evidence without reporting it to the AA.

To begin with, suppose that the two firms agree to keep the hard evidence and subsequently report it to the AA, i.e., they coordinate on (KEEP AND REPORT, KEEP AND REPORT). The latter is a NE iff $V_{KR} \geq V_{KR}^d$. Given Assumption 4 this inequality always holds.

Lemma 4. *(*KEEP AND REPORT, KEEP AND REPORT*) is a strict NE of the Revelation subgame.*

Suppose now that firms agree to keep the hard evidence, nevertheless not report

it to the AA, i.e., coordinate on (Keep and Not Report, Keep and Not Report). The latter is a NE iff $V_{KNR} \geq V_{KNR}^d$, which holds iff $p + \Delta \leq \frac{\delta \pi}{F}$.

Lemma 5. *(*Keep and Not Report, Keep and Not Report*) is a NE of the Revelation subgame iff $p + \Delta \leq \frac{\delta \pi}{F}$.*

Notice that when $\frac{\delta \pi}{F} > 1$, (Keep and Not Report, Keep and Not Report) is a strict NE.

Given Lemmata 4 and 5, we apply the Pareto criterion. Thus, (Keep and Report, Keep and Report) Pareto dominates (Keep and Not Report, Keep and Not Report) iff $V_{KR} > V_{KNR}$. The latter inequality holds iff $p + \Delta > \frac{1}{2}$.

Lemma 6. *The Pareto dominant NE of the Revelation subgame is:*

1. *(*Keep and Not Report, Keep and Not Report*) iff $p + \Delta \leq \frac{1}{2}$*

2. *(*Keep and Report, Keep and Report*) iff $p + \Delta > \frac{1}{2}$*

Proof. See Appendix B. $\qquad \square$

In words, if the (aggregate) probability of detection when firms keep hard evidence is relatively high, i.e., $p + \Delta > \frac{1}{2}$, firms report the hard evidence. The underlying reasoning is that when the probability of detection with hard evidence is greater (lower) than $\frac{1}{2}$, the fine that firms expect to pay, in case of detection by the AA, is greater (lower) than the one associated with self-reporting (i.e., $(p + \Delta)F > \frac{1}{2}F$).

We now proceed backwards to the *Keep-Destroy* subgame. Depending on which strategy dominates at the *revelation* subgame, by Lemma 6 we distinguish two cases.

Keep-Destroy subgame for $p + \Delta > \frac{1}{2}$ If $p + \Delta > \frac{1}{2}$, then at the *revelation* stage the firms self-report to the AA, provided that they keep hard evidence. Thus, for those parameters' values firms' decisions is either to destroy the hard evidence or keep the hard evidence and then report it to the AA.

To begin with, suppose that the two firms agree to keep the hard evidence (and subsequently report it to the AA). This is a NE iff $V_{KR} \geq V_{KR}^d$. The latter inequality always holds given that $\beta \in [0, 1]$.

Lemma 7. *(*Keep and Report, Keep and Report*) is a strict NE of the Keep-Destroy subgame for $p + \Delta > \frac{1}{2}$.*

Suppose now that firms agree to destroy the hard evidence, i.e., coordinate on (Destroy, Destroy). This strategy profile is a NE iff $V_D \geq V_D^d$. The latter inequality holds iff $p \leq \frac{\delta \pi}{F}$.

Lemma 8. (Destroy, Destroy) *is a NE of the Keep-Destroy subgame if* $p \leq \frac{\delta\pi}{F}$.

From Lemma 8 and considering that $p \in [0, 1 - \Delta]$, we distinguish two cases: *i)* $\frac{\delta\pi}{F} \leq 1 - \Delta$ and *ii)* $\frac{\delta\pi}{F} > 1 - \Delta$. If the former condition holds, then for $p > \frac{\delta\pi}{F}$ (Keep and Report, Keep and Report) is the unique NE of the *Keep-Destroy* subgame. Otherwise, both (Destroy, Destroy) and (Keep and Report, Keep and Report) are NE of the *Keep-Destroy*.

Lemma 9. *For* $p + \Delta > \frac{1}{2}$ *the Keep-Destroy subgame has two NE iff:*

1. $\frac{\delta\pi}{F} \leq 1 - \Delta$ *and* $p \in [\frac{1}{2} - \Delta, \frac{\delta\pi}{F}]$ *or*

2. $\frac{\delta\pi}{F} > 1 - \Delta$ *and* $p \in [\max\{0, \frac{1}{2} - \Delta\}, 1 - \Delta]$

Given Lemma 9 we apply the Pareto criterion. Thus, (Keep and Report, Keep and Report) Pareto dominates (Destroy, Destroy) iff $V_{KR} > V_D$. The latter inequality holds iff $p > p_4 \equiv \frac{F - 2\beta\pi}{2(1-\beta)F}$.

Lemma 10. (Keep and Report, Keep and Report) *Pareto dominates* (Destroy, Destroy) *iff* $p > p_4 \equiv \frac{F - 2\beta\pi}{2(1-\beta)F}$.

From Lemma 10 and considering that *i)* $p_4 < 0$ iff $\frac{\pi}{F} > \frac{1}{2\beta}$, *ii)* $p_4 < \frac{1}{2} - \Delta$ iff $\Delta < \Delta_5 \equiv \frac{\beta}{1-\beta}\frac{\pi}{F} - \frac{\beta}{2(1-\beta)}$ and *iii)* $p_4 > 1 - \Delta$ iff $\Delta < \Delta_6 \equiv \frac{1-2\beta}{2(1-\beta)} + \frac{\beta}{1-\beta}\frac{\pi}{F}$, we establish the Pareto dominant SPE of the game when $p + \Delta > \frac{1}{2}$.

Lemma 11. *The Pareto dominant SPE of the model with LP for* $p + \Delta > \frac{1}{2}$ *is:*

1. (Keep and Report, Keep and Report) *iff:*

   (a) $\beta < \frac{1}{2}$ *and*

       i. $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta < \Delta_5$ *and* $p \in [\frac{1}{2} - \Delta, 1 - \Delta]$ *or*

       ii. $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta \in [\Delta_5, \Delta_6]$ *and* $p \in (p_4, 1 - \Delta]$ *or*

       iii. $\frac{\pi}{F} > \frac{1}{2\beta}$ *and* $p \in [\max\{0, \frac{1}{2} - \Delta\}, 1 - \Delta]$ *or*

   (b) $\beta > \frac{1}{2}$, $\frac{\pi}{F} \geq 1$ *and* $p \in [\max\{0, \frac{1}{2} - \Delta\}, 1 - \Delta]$

2. (Destroy, Destroy) *iff* $\beta < \frac{1}{2}$, $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$ *and*

   (a) $\Delta \in [\Delta_5, \Delta_6]$ *and* $p \in [\max\{0, \frac{1}{2} - \Delta\}, p_4]$ *or*

   (b) $\Delta > \Delta_6$ *and* $p \in [0, 1 - \Delta]$

where $p_4 = \frac{F-2\beta\pi}{2(1-\beta)F}$, $\Delta_5 = \frac{\beta}{1-\beta}\frac{\pi}{F} - \frac{\beta}{2(1-\beta)}$ *and* $\Delta_6 = \frac{1-2\beta}{2(1-\beta)} + \frac{\beta}{1-\beta}\frac{\pi}{F}$.

Proof. See Appendix B. $\qquad\square$

The intuition underlying Lemma 11 is the following. Firms keep the hard evidence and subsequently report it to the AA, rather than destroying it, if p is relatively high, and, in particular, greater than the critical value p_4. The latter balances the CCE ($\beta \cdot \pi$) and the CAE ($\frac{F}{2}$).[178] The critical value p_4 is negatively related to both, β and $\frac{\pi}{F}$. Thus, with an increase either in β or $\frac{\pi}{F}$ firms are less inclined to destroy the hard evidence. An increase in β makes the CCE sharper, as the foregone profits in case of the collapse of the cartel are greater, leaving the CDE intact. When β is sufficiently high ($\beta > \frac{1}{2}$), the CCE outweights the CAE. The foregone profits in the case where the cartel collapses are greater than the expected fine with self-reporting. Thus, firms keep and report the hard evidence to the AA. An increase in $\frac{\pi}{F}$ makes the CCE sharper, while it mitigates the CDE. Moreover, when $\frac{\pi}{F}$ is sufficiently high ($\frac{\pi}{F} > \frac{1}{2\beta}$), firms always keep the hard evidence.

Suppose that β is relatively low ($\beta < \frac{1}{2}$) and that p, the base probability of detection, is also relatively low. Furthermore, suppose that Δ is sufficiently high ($\Delta > \Delta_6$), but the aggregate probability of detection is relatively high ($p + \Delta > \frac{1}{2}$). Then, the expected cost associated with firms' decisions to destroy the hard evidence, that is, the sum of the fine the firms have to pay in case of detection by the AA on the basis of indirect evidence and the expected foregone profits in case of the collapse of the cartel, is lower than the fine associated with self-reporting to the AA. Thus, firms destroy the hard evidence. On the contrary, if Δ is sufficiently low ($\Delta < \Delta_5$), then the expected fine with self-reporting ($\frac{F}{2}$) is lower than the cost associated with firms' decision to destroy the hard evidence (expected fine and expected foregone profits). Thus, firms keep the hard evidence to exploit the LP and pay lower expected fines.

Corollary 2. *Firms are less likely to destroy the hard evidence with a higher β and $\frac{\pi}{F}$ and lower Δ.*

Keep-Destroy subgame for $p + \Delta \le \frac{1}{2}$ If $p + \Delta \le \frac{1}{2}$, then by Lemma 6 firms do not self-report at the *revelation* stage, provided that they indeed keep hard evidence. Thus, for those parameters' values, a firms' decision is between destroying and retaining the hard evidence without disclosing it to the AA.

To begin with, suppose that the two firms agree to destroy the hard evidence, i.e., coordinate on (DESTROY, DESTROY). This strategy profile is a NE if no profitable unilateral deviation exists, i.e., iff $V_D \ge V_D^d$. The latter inequality always holds.

---

[178]Notice that p_4 can be expressed as: $p_4 = \frac{1}{(1-\beta)F}(\frac{F}{2} - \beta\pi)$.

Lemma 12. (DESTROY, DESTROY) *is a strict NE of the keep-Destroy subgame for $p + \Delta \leq \frac{1}{2}$.*

Suppose now that firms agree to keep the hard evidence that, however, they do not report to the AA at the *revelation* stage, i.e., coordinate on (KEEP AND NOT REPORT, KEEP AND NOT REPORT). This strategy profile is a NE iff $V_{KNR} \geq V_{KNR}^d$. This inequality holds iff $p \leq p_5 \equiv \frac{\pi}{F} - \frac{(\delta + \beta - \beta\delta + 1)\Delta}{2(\delta + \beta - \beta\delta)}$.

Lemma 13. (KEEP AND NOT REPORT, KEEP AND NOT REPORT) *is a NE of the Keep-Destroy subgame iff $p \leq p_5$.*

Given Lemmata 12 and 13 for $p \leq p_5$ there are two NE. Taking into account that *i)* $p_5 > \frac{1}{2} - \Delta$ and *ii)* $p_5 > 0$ iff $\Delta < \Delta_7 \equiv \frac{\pi}{F} \frac{2(\beta + \delta - \beta\delta)}{\beta + \delta - \beta\delta + 1}$, we conclude that (KEEP AND NOT REPORT, KEEP AND NOT REPORT) Pareto dominates (DESTROY, DESTROY) if $V_{KNR} > V_D$. The latter inequality holds iff $p < p_3 \equiv \frac{\pi}{F} - \frac{\Delta}{\beta}$. Notice that this is the same condition as in the baseline model, however here it holds only for $p + \Delta \leq \frac{1}{2}$.

Lemma 14. KEEP AND NOT REPORT, KEEP AND NOT REPORT) *Pareto dominates* (DESTROY, DESTROY) *iff $p < p_3 \equiv \frac{\pi}{F} - \frac{\Delta}{\beta}$.*

Considering that: *i)* $p_3 > 0$ iff $\Delta < \Delta_3 \equiv \beta \frac{\pi}{F}$ and $p_3 < \frac{1}{2} - \Delta$ iff $\Delta \geq \Delta_5 \equiv \frac{\beta}{1-\beta}(\frac{\pi}{F} - \frac{1}{2})$ and Lemma 14, we establish the Pareto dominant SPE of the game for $p + \Delta \leq \frac{1}{2}$.

Lemma 15. *The Pareto dominant SPE of the model with LP for $p + \Delta \leq \frac{1}{2}$ is:*

1. *(KEEP AND NOT REPORT, KEEP AND NOT REPORT) iff:*

   (a) $\beta < \frac{1}{2}$ and

       i. $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta < \Delta_5$ and $p \in [0, \frac{1}{2} - \Delta]$ or

       ii. $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta \in [\Delta_5, \Delta_3]$ and $p \in [0, p_3]$ or

       iii. $\frac{\pi}{F} \in [\frac{1}{2\beta}, \frac{1}{\beta}]$, $\Delta < \frac{1}{2}$ and $p \in [0, \frac{1}{2} - \Delta]$ or

       iv. $\frac{\pi}{F} > \frac{1}{\beta}$, $\Delta < \frac{1}{2}$ and $p \in [0, \frac{1}{2} - \Delta]$ or

   (b) $\beta > \frac{1}{2}$, $\frac{\pi}{F} \geq 1$, $\Delta < \frac{1}{2}$ and $p \in [0, \frac{1}{2} - \Delta]$

2. *(DESTROY, DESTROY) iff $\beta < \frac{1}{2}$ and*

   (a) $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta \in [\Delta_5, \Delta_3]$ and $p \in [p_3, \frac{1}{2} - \Delta]$ or

   (b) $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $[\Delta_3, \frac{1}{2}]$ and $p \in [0, \frac{1}{2} - \Delta]$

where $p_3 = \frac{\pi}{F} - \frac{\Delta}{\beta}$, $\Delta_3 = \beta \frac{\pi}{F}$ *and* $\Delta_5 = \frac{\beta}{1-\beta}\frac{\pi}{F} - \frac{\beta}{2(1-\beta)}$.

Proof. See Appendix B. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

The reasoning behind of Lemma 15 is the following. Firms keep hard evidence without reporting it to the AA, rather than destroying, if p is relatively low, and in particular lower than the critical value p_3. The latter balances the CCE and the CDE. The critical value of p_3 is positively related to $\frac{\pi}{F}$ and β and negatively related to Δ. Thus, an increase either in $\frac{\pi}{F}$ or in β makes firms less prone to destroy the hard evidence. Equivalently, it makes it more likely that firms will keep, nevertheless not report the hard evidence. Such changes make the CCE more acute, while the CDE is alleviated. When β is sufficiently high ($\beta > \frac{1}{2}$), the CCE outweights the CDE so that the firms keep the hard evidence without reporting it to the AA, rather than destroying it.

The CCE is not influenced by Δ. Such a change, on the contrary, markedly affects the CDE. Specifically, an increase (decrease) Δ makes the CDE more (less) acute. Thus, an increase in Δ makes firms' decision to destroy the hard evidence more likely. When Δ is sufficiently low ($\Delta < \Delta_5$), then the CCE dominates the CDE (p_3 becomes greater than 1) and thus firms always keep the hard evidence. Conversely, when Δ is sufficiently high ($\Delta > \Delta_3$) then the CDE dominates the CCE (p_3 becomes negative), and, thus, firms always destroy the hard evidence.

Corollary 3. *Firms are less prone to destroy the hard evidence with a higher β or $\frac{\pi}{F}$. They are less likely to keep the hard evidence with a higher Δ.*

By Lemmata 11 and 15, we can state the main proposition of this section.

Proposition 2. *For given policy and other parameters (p, Δ, F, β, π) the Pareto dominant SPE of the extended game with LP is:*

1. (DESTROY, DESTROY) *iff* $\beta < \frac{1}{2}, \frac{\pi}{F} \in [1, \frac{1}{2\beta}]$ *and*

   (a) $\Delta \in [\Delta_5, \Delta_3]$ *and* $p \in [p_3, \frac{1}{2} - \Delta]$ *or*

   (b) $[\Delta_3, \frac{1}{2})$ *and* $p \in [0, \frac{1}{2} - \Delta]$ *or*

   (c) $\Delta \in [\Delta_5, \Delta_6]$ *and* $p \in [\max\{0, \frac{1}{2} - \Delta\}, p_4]$ *or*

   (d) $\Delta > \Delta_6$ *and* $p \in [0, 1 - \Delta]$

2. (KEEP AND NOT REPORT, KEEP AND NOT REPORT) *iff:*

   (a) $\beta < \frac{1}{2}$ *and*

       i. $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta < \Delta_5$ *and* $p \in [0, \frac{1}{2} - \Delta]$ *or*

$ii.$ $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta \in [\Delta_5, \Delta_3]$ *and* $p \in [0, p_3)$ *or*

$iii.$ $\frac{\pi}{F} > \frac{1}{2\beta}$, $\Delta < \frac{1}{2}$ *and* $p \in [0, \frac{1}{2} - \Delta]$ *or*

(b) $\beta > \frac{1}{2}$, $\frac{\pi}{F} \geq 1$, $\Delta < \frac{1}{2}$ *and* $p \in [0, \frac{1}{2} - \Delta]$

3. (KEEP AND REPORT, KEEP AND REPORT) *iff:*

(a) $\beta < \frac{1}{2}$ *and*

$i.$ $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta < \Delta_5$ *and* $p \in [\frac{1}{2} - \Delta, 1 - \Delta]$ *or*

$ii.$ $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta \in [\Delta_5, \Delta_6]$ *and* $p \in (p_4, 1 - \Delta]$ *or*

$iii.$ $\frac{\pi}{F} > \frac{1}{2\beta}$ *and* $p \in [\max\{0, \frac{1}{2} - \Delta\}, 1 - \Delta]$ *or*

(b) $\beta > \frac{1}{2}$, $\frac{\pi}{F} \geq 1$ *and* $p \in [\max\{0, \frac{1}{2} - \Delta\}, 1 - \Delta]$

where $p_3 = \frac{\pi}{F} - \frac{\Delta}{\beta}$, $p_4 = \frac{F - 2\beta\pi}{2(1-\beta)F}$, $\Delta_3 = \beta\frac{\pi}{F}$, $\Delta_5 = \frac{\beta}{1-\beta}\frac{\pi}{F} - \frac{\beta}{2(1-\beta)}$ *and* $\Delta_6 = \frac{1-2\beta}{2(1-\beta)} + \frac{\beta}{1-\beta}\frac{\pi}{F}$.

Proof. See Appendix B. $\square$

Depending on the parameters' values of the game (p, Δ, β, π, F), all three strategies could emerge as Pareto dominant SPE. If the base probability of detection is relatively low, and, in particular, lower than the critical value p_3, firms agree to keep, nevertheless not report the hard evidence to the AA. In that case, firms keep the hard evidence in order to enhance cartel's sustainability rather than to exploit the LP. Conversely, if the base probability is sufficiently high, and, in particular, greater than the critical value p_3, then firms keep the hard evidence and report to the AA. In that case, firms keep the hard evidence to exploit the LP rather than to enhance cartel's sustainability. For intermediate values of the base probability of detection, that is, for $p \in [p_3, p_4]$, the CDE and the CAE prevail over the CCE. Thus, firms destroy the hard evidence in order to minimize the risk of cartel detection.

Firms are more prone to keep hard evidence with a higher β or $\frac{\pi}{F}$. Their decisions on whether to self-report or not depend on the aggregate probability of detection. Firms are less prone to keep without reporting if p or Δ increase.

Corollary 4. *Firms are less likely to destroy the hard evidence when β or $\frac{\pi}{F}$ increase. They are more likely to destroy hard evidence when Δ decrease.*

We have to underline that when β is sufficiently high ($\beta > \frac{1}{2}$), firms never destroy the hard evidence. With a sufficiently high β the expected foregone profits are higher than the expected fine associated with firms' decisions to keep the hard evidence. Noteworthy, this is not the case when a LP is not included in the
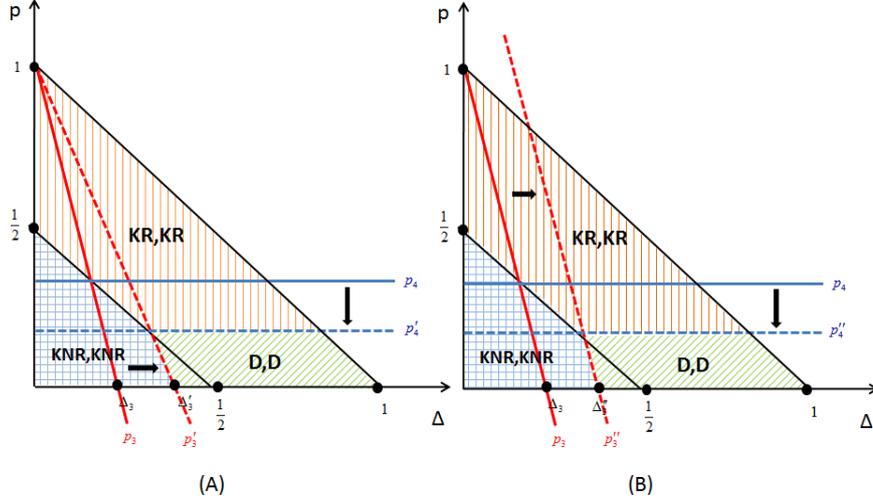
Figure 2: With LP. Graph A shows the effect of an increase in β. Graph B shows the effect of an increase in $\frac{\pi}{F}$. Both graphs are drawn with the initial assumption that $\frac{\pi}{F} = 1$.

portfolio of policy instruments. As we have seen in Section 2.3, firms destroy the hard evidence of their illegal cartel when $p > p_3$ regardless of the level of β.

Moreover, if Δ is sufficiently high ($\Delta > \frac{1}{2}$), firms always keep and exploit the LP by reporting the evidence to the AA. In this latter case, the expected fine with self-reporting is lower than the one without self-reporting. The same conclusion is reached if $\frac{\pi}{F}$ is sufficiently high ($\frac{\pi}{F} > \frac{1}{2\beta}$).

Corollary 5. *If $\beta < \frac{1}{2}$ and $\frac{\pi}{F} > \frac{1}{2\beta}$ or $\beta > \frac{1}{2}$ then firms always keep hard evidence. If in addition $\Delta < \frac{1}{2}$ then the Pareto dominant SPE is (*KEEP AND REPORT, KEEP AND REPORT*) if $p \in [\frac{1}{2} - \Delta, 1 - \Delta]$ and (*KEEP AND NOT REPORT, KEEP AND NOT REPORT*) if $p \in [0, \frac{1}{2} - \Delta]$. However, if $\Delta \geq \frac{1}{2}$, the unique Pareto dominant SPE is (*KEEP AND REPORT, KEEP AND REPORT*) $\forall p \in [0, 1 - \Delta]$.*

If $\beta = 0$ and $\Delta > \frac{1}{2}$ (implying that $p < \frac{1}{2}$) then firms always destroy the hard evidence.[179] In words, if a cartel's sustainability is independent of hard evidence and if the additional probability of detection is sufficiently high (i.e., $\Delta > \frac{1}{2}$), the unique SPE will be (DESTROY, DESTROY). Given that $\beta = 0$, the CCE evaporates. Moreover, for these parameters' values the CDE dominates the CAE, i.e., the expected fine with self-reporting is lower than the one without self-reporting. Given that the expected fine with self-reporting is higher than the cost

---
[179]If $\beta = 0$ then $\Delta_3 = 0$, $\Delta_5 = 0$ and $\Delta_6 = \frac{1}{2}$.

associated with firms' decision to destroy the hard evidence (foregone profits and expected fine), firms destroy the evidence. On the other hand, if the additional probability of detection is relatively low ($\Delta < \frac{1}{2}$), firms destroy the hard evidence as long as the base probability of detection is relatively low ($p < \frac{1}{2}$). In that case, the CDE dominates both the CAE and the CCE. The expected fine that the firms have to pay if convicted on the basis of indirect evidence is lower than the fine firms have to pay if they decide to keep the hard evidence regardless of reporting or not to the AA. On the contrary, if the base probability of detection is sufficiently high ($p > \frac{1}{2}$), then firms keep the hard evidence and exploit the LP. In the latter case the CAE dominates the CDE. It is worth noting that in the baseline model firms always destroy the hard evidence when $\beta = 0$. This is because when a LP is not available, the CAE does not materialize. In this case, retaining hard evidence would increase the risk of detection without improving the functioning of the cartel. For a graphical illustration see Figure 3 below (Graphs A and B).

Corollary 6. *If $\beta = 0$ then for any $p < \frac{1}{2}$ ($p > \frac{1}{2}$) firms always destroy (keep) the hard evidence.*

If $\beta = 1$, that is, if hard evidence is imperative for the survival of the cartel, then firms always keep the hard evidence. The decision on whether to apply for leniency depends on the aggregate probability of detection. If it is relatively high ($p + \Delta > \frac{1}{2}$), then firms always find it advantageous to exploit the LP as the expected fine with self-reporting is lower than that without self-reporting. If it is relatively low ($p + \Delta < \frac{1}{2}$), then firms keep the hard evidence, nevertheless not report to the AA. The expected fine without self-reporting is lower than the one associated with firms' decision to self-report. For a graphical illustration see Graph C in Figure 3.

When $\beta = 1$, firms always keep hard evidence. In addition, if $p + \Delta > \frac{1}{2}$, then they exploit the LP. Otherwise, they keep the hard evidence only to facilitate the administration and enhance the implementation of the cartel.

Before concluding this section, it may worth saying a word about the impact of the generosity of the LP on firms' decisions. We have assumed that the LP provides full amnesty to the first self-reporting firm. Instead of full amnesty, consider a LP which is characterized by a fine discount rate equal to ρ. Then, the expected fine with self-reporting is $\frac{(1-\rho)F}{2}$. Clearly, an increase in the fine discount rate has similar effects to a decrease in the fine for self-reporting firms. Therefore, a more generous LP has two effects on firms' decisions: 1) enhance their incentives to keep the hard evidence and, 2) enhance their incentives to report the retained evidence to the AA. Intuitively, a more generous LP decreases the expected fine with leniency (equivalently, it strengthens the CAE).
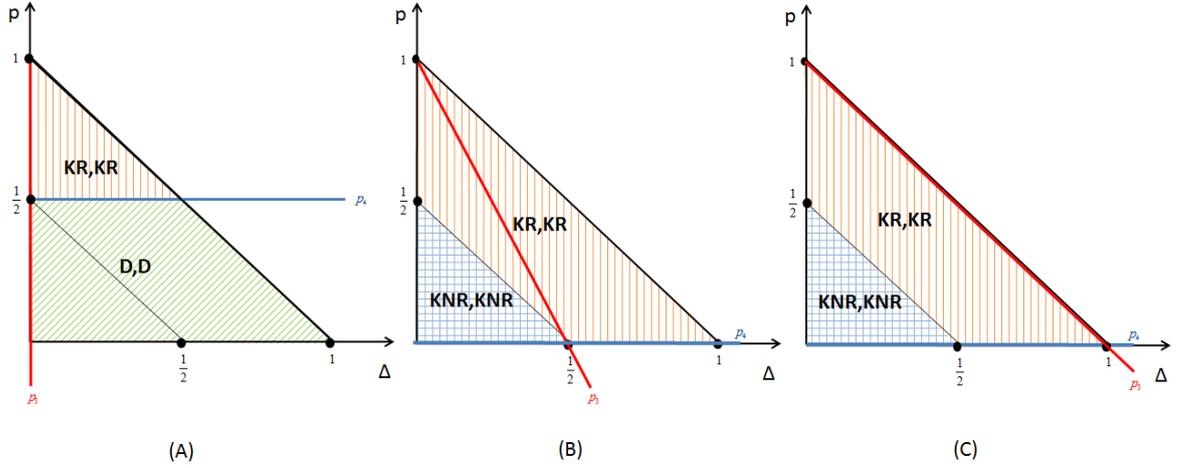
Figure 3: Some extreme cases. Graph A: $\frac{\pi}{F} = 1$ and $\beta = 0$. Graph B: $\frac{\pi}{F} = 1$ and $\beta = \frac{1}{2}$. Graph C: $\frac{\pi}{F} = 1$ and $\beta = 1$.

Corollary 7. *Firms are more prone to keep and report the hard evidence with a more generous LP.*

2.5 Comparison with the baseline model without a LP

To simplify the comparison of the two settings, with and without a LP, suppose that $\pi = F$ so that $\frac{\pi}{F} = 1$. A simple contrast of Propositions 1 and 2 reveal that the introduction of the LP, as part of the enforcement policy, makes firms more prone to keep the hard evidence. The reason is that the introduction of the LP provides an *option value* to those firms that keep hard evidence. Firms keep hard evidence not only to enhance the administration of their agreement, and sustain the cartel, but also to exploit the LP, when such an exploitation is advantageous for them.

Figure 4 below illustrates the effects of the introduction of a LP. Notice that for $p \in [\frac{1}{2} - \Delta, p_3]$ firms continue to keep the hard evidence, however, they exploit the LP by reporting to the AA (indicated by $NR \rightarrow R$). The principal objective for keeping hard evidence is to enhance the administration of their unlawful agreement and not to exploit the LP. However, firms also find it profitable to exploit the LP and pay a lower fine in case of conviction. For $p > max\{p_3, p_4\}$, firms' decisions at the keep-destroy stage are reversed. Now firms keep the hard evidence primarily to exploit the LP (indicated by $D \rightarrow KR$ in Figure 4 below). The expected fine linked to firms' decisions to keep and report the hard evidence is lower than the cost associated with firms' decisions to destroy the evidence (foregone profits

and expected fine). Therefore, by retaining the hard evidence, firms enhance the administration of their agreement, and thus sustain the cartel. At the same time, firms exploit the LP by self-reporting to the AA.

Corollary 8. *When a LP is available, then for $p \in [\frac{1}{2} - \Delta, p_3]$ firms' decisions to keep the hard evidence remain intact.*[180] *Firms keep the hard evidence primarily to enhance the implementation of the cartel and secondarily to exploit the LP. On the contrary, for $p > max\{p_3, p_4\}$ firms' decisions to destroy the hard evidence are reversed when a LP is available.*[181] *Firms keep the hard evidence primarily to exploit the LP.*
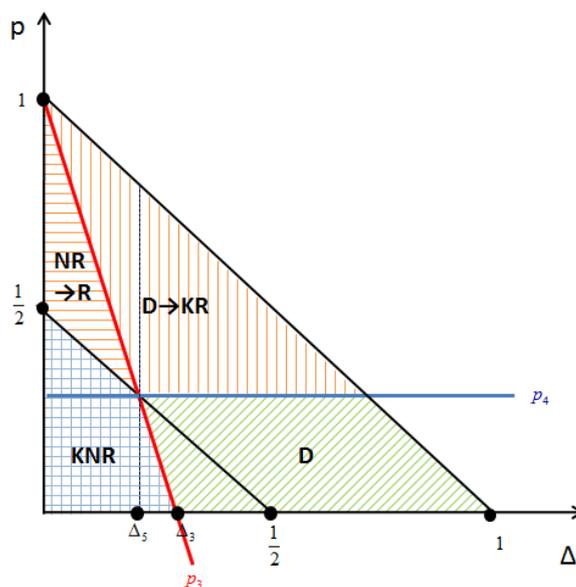
Figure 4: The impact of a LP on firms' decisions.

2.6 Conclusions

The motivation of this paper derives from the empirical observation that many collusive firms keep hard evidence, even when the enforcement policy does not encompass a LP. We provide plausible answers to three relevant questions: 1) why firms keep hard evidence when a LP is not available, 2) what is the impact of a LP on firms' decisions' to keep or destroy hard evidence and 3) why firms keep hard evidence, nevertheless not report to the AA.

---

[180]Notice that $p \in [\frac{1}{2} - \Delta, p_3]$ if $\Delta < \Delta_5$.
[181]Equivalently, if (i) $\Delta > \Delta_5$ and $p \in [p_3, 1 - \Delta]$ or (ii) $\Delta > p_5$.

In the absence of a LP, we have showed that firms' decisions to keep or destroy the hard incriminating evidence hinges upon i) the sensitivity of cartel's administration and implementation to hard evidence, ii) the increased risk of detection when firms keep the hard evidence and iii) the collusive profits-fine ratio. Reflecting on these theoretical results, this study puts forward three alternative, but not exclusionary, explanations for retaining hard evidence even when a LP is not part of the enforcement policy. First, cartel's administration and implementation is very sensitive to hard evidence, so that firms keep hard evidence to enhance the stability of their illicit agreement. Second, the efficiency of the AA in detecting and prosecuting cartels with hard evidence is relatively low, or the quality of hard evidence is poor, so that the expected incremental fine, if firms keep the hard evidence, is lower than the expected forgone profits associated with firms' decisions to destroy the hard evidence. And third, the level of the fine imposed on detected cartels is sufficiently lower than the accrued collusive profits, so that firms have more to lose if they destroy the hard evidence, and less to pay if detected by the AA.

We have also showed that the introduction of a LP reinforce firms' incentives to keep hard evidence. Essentially, the LP provides to those firms that keep hard evidence an option value, which, *ceteris paribus*, decreases the expected fine. Thus, firms may keep the hard evidence not only to enhance the stability of the cartel but also to exploit the LP and pay lower fines. The empirical fact that we are experiencing more cartels self-reporting may, therefore, be a *side-effect* of the introduction of LPs.

Finally, collusive firms may keep hard evidence without reporting to the AA. We have showed that this behavior is more likely when i) the probability in which the cartel collapse in the absence of hard evidence increases, ii) the collusive profits-fine ratio increases and iii) the additional probability of cartel detection decreases. Accordingly, the empirical fact that AAs' investigations expose hard evidence retained by cartelists can be justified on three grounds. First, that the cartel is very sensitive to hard evidence, so that firms keep hard evidence to enhance the survival rate of the cartel. Second, that the efficiency of the AA in detecting and prosecuting cartels with hard evidence is sufficiently low, so that the incremental and total expected fine is lower than the expected fine with self-reporting. And third, that the level of the fine imposed on detected cartels is sufficiently lower than the accrued collusive profits, so that firms have more to lose if they destroy the hard evidence and less to pay if detected by the AA. It is worth noting that all results in our paper are obtained without allowing for positive rewards contrary to Aubert et al. [2006].

An assumption of the paper is that the LP offers full amnesty only to the first self-reporting firm. This is a reasonable assumption to be made and it is

in line with other studies in the literature advocating for the maximum feasible fine discount for the first self-reporting firm. A question that arises is what is the impact of the level of the fine discount rate on firms' decisions to keep or destroy the hard evidence that is generated by the illegal cartel. It is not difficult to notice that a change in the fine discount rate has similar effects to a change in the level of the fine. Thus, with a higher (lower) fine discount rate, firms are more (less) inclined to keep and report the hard evidence to the AA. This observation can be useful for those designing a leniency policy.

A final remark. Despite the simple setting, our paper sheds some additional light on firms' decisions to keep or destroy the hard incriminating evidence as well as on their decisions to report or not the hard evidence to the AA. By abstracting from behavioral elements or cognitive biases our paper predicts a particular pattern of behavior by colluding firms. However, our understanding of the underlying reasons justifying firms' decisions to keep or destroy the hard evidence may be enriched if we diverge from the realm of rational choice theory by integrating bounded rationality in the analysis. For instance, collusive firms may be time inconsistent or overconfident, although they know they have to destroy the hard evidence when they do it is too late!. Such extension to the literature could offer alternative explanations with regard to the research questions of this paper, and indeed may come up with very different patterns of behavior. This research avenue remains open to future exploration.

Chapter 3: Strategic Antitrust Authorities and Leniency Programmes - A Signalling Game

3.1 Introduction

Cartels are the most egregious form of anticompetitive behaviour. Because of their inimical effects to consumers welfare, cartels are considered as per-se illegal. A central priority of antitrust authorities (henceforth "AAs"), in various jurisdictions, is dissolving existing cartels, and ultimately deterring the formation of new cartels.[182] This goal is challenging to achieve because of the inherently clandestine nature of cartels. Inevitably, two thorny economic problems emerge. The first transpires because of the limited enforcement resources available to the AA to combat illegal cartels. Effectively, an AA cannot monitor all markets.[183] The second is due to the asymmetry of information between an AA and a cartel, which can be two-sided. On the one hand, firms know whether they participate in a cartel while, on the other hand, an AA knows how compelling its evidence against the cartel is. Unlike previous literature, we explore the implications of the second source of information asymmetry - cartel members do not know the strength of the AA's case. The main question we aim to answer is whether an AA should disclose its evidence to the cartel or exploit its informational lead and induce the cartel to self-report, despite the weakness of its evidence. For this, we develop a stylized signalling game which resembles to the stylized signalling game introduced by Cho and Kreps [1987].[184] Crucial to the outcome of the strategic interaction between the AA and the cartel is the impact of the leniency program (henceforth "LP") whereby a cartel member that self-reports its unlawful conduct to the AA pays a reduced fine.[185] The AA must consider its information revelation policy alongside its LP design. To our knowledge this is the first paper in the literature to consider the impact of a LP in the framework of a signalling game.

The existing literature on LPs typically assumes that the information deficit

---

[182]This is reflected in the number of prosecuted cartels. For instance, for the periods 1995-1999, 2000-2004 and 2005-2009, the European Commission has issued 10, 30 and 33 cartel decisions, respectively.

[183]Thus, cartels are subject to probabilistic law enforcement.

[184]The simplest form of a signalling game considers two players which strategically interact. One player has some private information which influences the payoff of the uninformed player. Critically, the informed player can signal the privately known information to the uninformed player with its actions.

[185]According to the applied European Commission's Notice on Immunity from fines and reduction of fines in cartel cases (OJ C298, 8/12/2006, pp. 17-22, par. 24), "*In order to qualify (for a reduction in fines), an undertaking must provide the Commission with evidence of the alleged infringement which represents significant added value with respect to the evidence already in the Commission's possession*".

is only on the AA's side. However, it may well be the case that an AA retains some critical information, which is not accessible by cartels. For instance, after an investigation in a market, an AA may discover evidence whose quality is unknown to all cartel members. The reversal in the information asymmetry may have pivotal implications for the outcome of the strategic interaction between an AA and a cartel. Importantly, it may influence the efficacy of the LP. Notably, the literature has not adequately put the spotlight on these implications hitherto.

This remark establishes the foundation stone of our study. In particular, we explore the strategic interaction between a cartel and an AA whose strength of evidence against the cartel is private information. The main research question is whether the AA should outrightly reveal or not how compelling its evidence is to the cartel. Given the social harm instigated by cartels, an AA would ideally prefer to unravel the cartel as soon as possible. Thus, crucial in the game is the time wherein the AA initiates the formal proceeding against colluding firms. Typically, the formal proceedings commence with the issue of a Statement of Objections (henceforth "SO") by the AA. A SO outlines the evidence against the cartel and delineates a preliminary conclusion.[186] Realistically, the SO is issued on the basis of the AA's incriminating evidence. The better the quality of evidence, the higher the probability of conviction. Thus, an AA with good evidence would prefer to send a SO to the cartel sooner than an AA with poor evidence, which may prefer to investigate the market more. The ensuing strategic situation provides a fertile ground for modeling a signalling game. In this context, the AA can essentially use the timing of issuing a SO to signal (or reveal) how compelling its evidence is. If the AA's evidence is not good, then it can invest in costly effort to discover additional incriminating evidence or interpret its existing evidence. We assume that the AA behaves strategically without committing ex-ante to the intensity of its prosecutorial effort. In that way, we overcome the problem of dynamic time inconsistency.

The study provides the conditions under which an AA should refrain from revealing the strength of its evidence to the cartel. In particular, the analysis shows that, with the appropriate design of its leniency policy, the AA can successfully induce the cartel to self-report at an early stage of the prosecutorial process, despite its feeble evidence. All else equal, the more generous the fine reduction, the easier it is to dissolve cartels (i.e., more pronounced desistance effects). The study casts light on the role of a) the (prior) beliefs the cartel holds regarding

---

[186] Drafting and sending a Statement of Objections to the suspected cartel members is a formal step in an Antitrust Authority investigation. It is a typical procedure rule in EU (and EU member states) antitrust proceedings. Essentially, a SO sets out the AA's provisional findings against the cartel and outlines its preliminary conclusion. See, *Antitrust Manual of Procedures*, Internal DG Competition working documents on procedures for the application of Articles 101 and 102 TFEU, March 2012.

the AA's evidence, b) the quality or strength of the AA's evidence, c) the social harm instigated by the cartel, and d) the severity of the fine imposed on convicted cartels.

A policy implication that logically derives from the analysis is that AAs should try to manipulate cartelists' (prior) beliefs regarding the evidence they hold against them. For instance, after a dawn raid, an AA may overstate, in a prudent and careful manner, by an announcement (e.g. a press release) that it has gathered potentially significant incriminating evidence and/or that it intends to send a SO soon. This remains credible as long as it induces a leniency application. Moreover, an AA may want to build a reputation of winning the antitrust cases it prosecutes to the courts. For instance, by opening a few profound investigations, rather than many superficial investigations, an AA may establish itself as successful or effective in convicting cartels. As a result, firms may revise their prior beliefs regarding the strength of the AA's case against them. An additional implication that develops from the analysis is that AAs should prioritize their activities to more serious and harmful cartels. All the same, it is easier for an AA to fruitfully elicit self-reporting by cartels that cause a greater harm to society. This is because AAs divert more prosecutorial resources to serious cartel cases, implying higher probabilities of conviction. A necessary condition for the last implication is that AAs must have sufficient resources to finance their prosecutorial effort if the cartel does not self-report its unlawful conduct. A higher budget can alleviate the AA's constraints. More importantly, it may enable an AA to credibly threaten, without pre-committing its resources, to successfully prosecute those cartels that decline to self-report.[187]

3.1.1 Relevant Literature

The economic literature on LPs has developed over the last two decades after the pioneer contributions by Motta and Polo [2003] and Spagnolo [2004].[188] A general finding of the relevant literature is that by eroding a cartel's internal trust, LPs may

---

[187]This resembles to the credible nuclear deterrence. If you have the weapons you will not need to use them.

[188]LPs were enacted and implemented for the first time in the US in 1978 (reformed in 1993) and the EU in 1996 (revised in 2002 and in 2006). The general consensus is that LPs have been a story of success. In the U.S., the Corporate Leniency Policy has been the *"most effective generator of international cartel cases"*. See the speech by James Griffin (Deputy Assistant Attorney General, U.S. DOJ, Antitrust Division), *The modern leniency program after ten years - 'A summary overview of the antitrust division criminal enforcement program'*, DOS, 2003, US (available at http://www.justice.gov/atr/public/speeches/201477.htm). Likewise, in the EU *"The leniency instrument has proven to be a formidable tool for encouraging firms to cooperate with the competition authorities"*. See the speech by Philip Lowe (Director General, DG Competition, E.C.), *What's the future of cartel enforcement*, Understanding Global Cartel Enforcement – Brussels, 11 February 2003 (available at http://ec.europa.eu/competition/speeches/text/sp2003044en.pdf).

evoke substantial desistance and deterrence effects (e.g. Motta and Polo [2003], Spagnolo [2004], Ellis and Wislon [2003] and Harrington [2008]). The direction and magnitude of these effects depend, among others, on market concentration (Ellis and Wislon [2003]), the degree of evidence heterogeneity (Feess and Walzl [2004] and Silbye [2010]), the heterogeneity of firms' market shares (Motchenkova and Leliefeld [2010]), whether fines are fixed or proportional to the accumulated collusive profits (Motchenkova [2004]), whether the probability of cartel detection deterministically or stochastically changes over time (Hinloopen [2003] Harrington [2008]) and whether the antitrust policy instruments (penalties and probability of detection) are endogenous to firms' pricing decisions (Chen and Harrington [2007]). Nevertheless, LPs are not unambiguously optimal. Wrongly designed LPs may exacerbate collusive behaviour or provoke cartel formation (e.g. Ellis and Wislon [2003], Buccirossi and Spagnolo [2006a], Aubert et al. [2006] and Motchenkova and Leliefeld [2010]).[189]

Our paper is closely related to Sauvagnat [2010]. The author develops a model wherein the AA has private information regarding the probability of successful conviction if an investigation is launched by the AA. In this context, Sauvagnat [2010] shows that with the appropriate design of its LP, the AA can fool cartel members to self-report, even if it knows that its investigation will be futile in detecting incriminating evidence. A major difference of the two studies is that we consider a signalling game. In effect, the scenario we consider allows the AA to signal or reveal the strength of its evidence to the cartel. Importantly, our study highlights the vital role of the cartel's beliefs regarding the AA's evidence. Contrary to Sauvagnat [2010], in our model, the AA does not commit ex-ante to a certain investigation policy. Thus, our study is not susceptible to dynamic time inconsistency problems.[190] Moreover, in our study the rate of law enforcement is endogenously determined by the AA, via its decisions about the optimal prosecutorial effort level. Another difference is that we consider proportional fines (i.e., fines that depend on the collusive profits and the duration of the illegal cartel) contrary to Sauvagnat [2010] who considers fixed fines. Regardless of the differences, both studies conclude that the AA can induce the cartel to self-report even

The empirical findings must be interpreted with prudence because only the detected cartels are observed, which unavoidably leads to an intrinsic sample selection bias. For empirical studies that aim to draw inferences regarding the impact of LPs on cartels see Brenner [2009], Miller [2009], Chang and Harrington [2010], Klein [2010] and Zhou [2011].

[189]For a comprehensive survey of the literature on LPs see Spagnolo [2008].

[190]The ex-ante threat of audit alone suffices to induce the desirable compliance with antitrust law. Nonetheless, ex-post the AA may have deficient incentives to comply with its initial audit rate commitments, given that audits are costly. This could potentially result in dynamic time inconsistency problems.The role of commitment in antitrust enforcement has been recently discussed by Feizi [2010]. The author finds that a credible commitment by the AA regarding the frequency of its audit policy does not unambiguously enhance the efficacy of antitrust policy.

if it has unsound incriminating evidence. Finally, both studies confirm that the desistance effects of LPs are enhanced with more generous fine discounts, which is in line with previous studies in the literature.

The remainder of the paper is organized as follows. Section 3.2 describes the game and its main assumptions. Section 3.3 presents the baseline game wherein information is perfectly symmetric. The game with asymmetric information is presented in Section 3.4. Section 4.5 concludes.

3.2   Description of the game

We consider a scenario of strategic interaction between a cartel, comprising of two identical firms, and an AA which enforces antitrust law. The AA's objective is to minimize the expected social cost of cartels. Because of the secret nature of the cartels and the high standard of evidentiary proofs required by the courts, the AA cannot convict the cartel unless it discovers hard incriminating evidence.[191] To attain its goal, the AA commits to a certain leniency policy whereby it provides a fine discount only to the first firm to come forward and denounce the illegal cartel.[192] This is a realistic assumption to be made, as it reflects a key feature of LPs in real word (e.g. US and EU leniency programmes).[193] Rather than assuming that the AA commits to a fixed probability of conviction, we consider a strategic AA which chooses its prosecutorial effort (which determines the probability of conviction) on a case-by-case basis. Realistically, an AA decides how many of its limited resources to invest in each case after processing some preliminary evidence with regard to the suspected cartel. Depending on the quality of the evidence detected, the AA may prioritize the prosecution of more serious cartels.

An initial step to the AA's investigations is to conduct a market investigation, (e.g. a surprised inspection - known as 'dawn raid'[194]) at suspected firms' premises or elsewhere) whereby it collects all readily available evidence.[195] Evidence is disorganized and bulky, and crucially its review involves a cost for the AA. Thus, the

---

[191]An implicit assumption we made is that the cartel leaves some evidence, which the AA can discover, if it investigates the market.

[192]This resembles to the optimal design of leniency programmes by Harrington [2008].

[193]The EU leniency programme allows for more than one firms to receive lenient treatment, provided that they submit value added incriminating information to the EU Commission. An implicit assumption we made in our study is that colluding firms retain symmetric evidence with regard to their illegal agreement. This implies that self-reporting by more than one firms does not add anything to the AA's case against the cartel. On the contrary, it may make the LP more easily exloitable by colluding firms.

[194]The name is derived from the time of the day when the investigations are normally held.

[195]The European Commission and national competition authorities of the EU member states have wide powers to inspect. Among otherts, they can carry out unannounced inspections or dawn raids at suspected firms' premises. The investigations can be extended to other places

AA randomly surveys and assess a subset of the accumulated evidence. The preliminary scrutinized evidence (henceforth the *"prima facie"* evidence) is privately known to the AA. To simplify the analysis, we assume that the strength of the prima facie evidence can be either good or poor.[196] The AA is said to be *good* when the exposed prima facie evidence is of high quality. On the other hand, the AA is said to be *bad* when the prima facie evidence is of poor quality. The better the quality of evidence, the higher the probability of conviction. We are going to elaborate this later. What is important to underscore at this point is that the cartel is uncertain about the quality of the prima facie evidence that the AA holds. However, it has some prior beliefs regarding the strength of the AA's evidence.

On the basis of the strength of its prima facie evidence, the AA decides in which period to send a SO to the cartel. We assume that evidence is perishable and, therefore, the AA has two options, either send the SO in the first period (SO at t=0) or in the second period (SO at t=1).[197] Ideally, the AA, wishing to minimize the expected social harm instigated by the cartel, prefers to expose the cartel at an earlier stage of the prosecutorial proceedings.

After receiving the SO, a colluding firm has two options. Either to collaborate with the AA and self-report the cartel, in which case it pays a reduced fine, or refute the SO and do not self-report. In the latter case, that is, if the cartel declines to self-report, the AA decides the amount of resources to invest in the prosecutorial process against the cartel (i.e., set up a team of legal and economic experts and/or hire consultants to scrutinize or refine the existing evidence or search for new evidence). If the AA successfully convicts the cartel, the colluding firms pay the full fine, which depends on the cartel's record.[198] Once dissolved (either by self-reporting or by the AA's own investigations), the AA intensely monitors the market and severely punishes recidivists. Thus, after their conviction by the AA, firms eternally compete.[199]

---

as well if there is a reasonable suspicion that evidence related to the inspection is being kept elsewhere. Sometimes these investigations are held simultaneously at the offices of multiple inspected firms in various countries.

[196] Equivalently, we may assume that some AAs are more effective in distinguishing cogent convicting evidence and poor/unreliable evidence (*profound AA v. superficial AA*). We will elaborate on this distinction distinction in Section 4.5.

[197] This can also be justified, to some extend, on the limitation period that comes with antitrust law violation. In essence, an AA cannot have an open investigation against firms forever. According to Article 25 of the EC Council Regulation 1/2003, the limitation period is (a) 3 years in the case of infringements of provisions concerning requests for information or the conduct of inspections; (b) 4 years in the case of all other infringements.

[198] This is consistent with the latest Fining Guidelines of the European Commission. See, *Guidelines on the method of setting fines imposed pursuant to Article 23(2)(a) of Regulation No 1/2003* (2006/C 210/02).

[199] In effect, this assumption deters firms from colluding and reporting information in each period, i.e., exploiting the LP. See, Spagnolo [2004] and Chen and Harrington [2007].

Critically, the invested resources reinforce the analytical foundations of the case at hand, and thereby enhance the accuracy of the AA's investigation. For instance, the AA may undertake further enquiries to discover new evidence, or to interpret the existing evidence. This is intuitive as the more resources the AA invests in a case, the higher the probability of cartel conviction, all else equal. To this end, we assume that the probability of conviction for a good AA is $\tilde{p} = \frac{e+p}{1+e}$, where $e \in [0, \bar{e}]$ is the effort level (bounded above by AA's budget) and $p \in (0,1)$ is a parameter that reflects the quality of the prima facie evidence (or the probability of inspection).[200] To simplify the analysis, we assume that $p = 0$ for a bad AA.[201] This implies that the probability of conviction for a bad AA is $\frac{e}{1+e}$. Essentially, a bad AA cannot convict the cartel unless it exerts more effort.

The adopted functional form for the probability of conviction has the following properties: 1) $\lim_{e\to 0}\tilde{p} = p$, 2) $\lim_{e\to\infty}\tilde{p} = 1$, 3) $\frac{\partial \tilde{p}}{\partial e} > 0$, 4) $\frac{\partial^2 \tilde{p}}{\partial e^2} < 0$, 5) $\frac{\partial \tilde{p}}{\partial p} > 0$ and 6) $\frac{\partial^2 \tilde{p}}{\partial e \partial p} < 0$.[202] Property 1 implies that with no further prosecutorial effort the probability of conviction hinges only on the quality of the prima facie evidence. Property 2 implies that when the AA invests infinitely many resources in the case at hand, the probability of conviction tends to 1. Properties 3 and 4 imply that the probability of conviction is increasing with the AA's effort. However, the marginal productivity of effort is decreasing with additional effort. In other words, the AA's technology of scrutinizing its existing evidence or detecting additional evidence exhibits diminishing marginal returns. Property 6 implies that the marginal productivity of AA's effort is decreasing with the quality of the evidence already scrutinized. The added value of the AA's prosecutorial effort depend on the evidence it has already scrutinized. In particular, the AA's effort become less productive in increasing the probability of conviction when the strength of the prima facie evidence is high.

More accuracy, however, comes at a cost. We assume that the AA's cost of effort is $C(e) = \lambda e$, where $\lambda > 1$ for a bad AA that sends a SO at t=0 and $\lambda = 1$ otherwise. Parameter λ represents the marginal cost (or the opportunity cost) of effort. We assume that this cost is higher for a bad AA, if it sends a SO at t=0. Intuitively, an AA with poor quality of prima facie evidence can invest its limited resources more productively to other activities or antitrust cases. Equivalently, it is too costly for a bad AA to divert resources from other activities and relocate them to the case at hand, given the poor quality of its prima facie evidence. On the other hand, at t=1 the opportunity cost of effort is equal for the two types of AA. This implies that it is not as costly for a bad AA to channel resources to

---

[200]The lower the probability of erroneously exonerating the cartel, i.e., the lower the type II error probability $\frac{1-p}{1+e}$.

[201]Thus, the probability of conviction for a bad AA is $\frac{e}{1+e}$.

[202]All calculations can be found in Appendix C.

the case since, otherwise, it will have to drop the case. This is grounded on two facts: i) the incriminating evidence is perishable and ii) the limitation period that comes with antitrust law violation (i.e., the AA cannot have an open investigation forever).[203]

A key assumption of the game is that there is no strategic interaction among cartel members in the market. Essentially, this allows us to abstract from the possibility of defections in the market by individual cartel members.[204] This can be justified on the fact that after the dawn raid at firms' premises, the market is tightly monitor, and, therefore, firms cannot change their market conduct without this smelling fishy to the AA. A sudden and unusual change in the firms' conduct, not justified by any change in market conditions and fundamental economic parameters, provides sufficient evidence to the AA to convict the cartel. Nonetheless, firms can (unilaterally) run to the AA to self-report and benefit from a lenient treatment, which materializes in a lower fine. We assume that only the first self-reporting firm is eligible to a fine discount.[205]

The emerging game is portrayed in Figure 5. A formal description of the game follows.

---

[203]This basically implies that the AA must prioritize the prosecution of the cartel cases which are more clear-cut (i.e. has better incriminating evidence). As we will see later, this may have a concrete impact on AA's reputation, and thus its effectiveness.

[204]Although there is no strategic interaction among cartel members in the market, *protection from punishment effect* (Spagnolo [2004]) or the *deviator amnesty effect* (Harrington [2008]) is present in our model. This is because we assume that the SO is issued at the end of each period; therefore colluding firms obtain the per-period collusive profits even when self-reporting occurs.

[205]This restriction on the eligibility to leniency induces firms into a race to the AA in order to be the first to self-report and pay a reduced fine. This effect of LPs is known as the *race to the courthouse effect.* Given the symmetry of firms (size, market shares, evidence they hold etc.) offering leniency to more than one firm would short-circuit this effect of leniency programmes on firms incentives to self-report their unlawful conduct. See, Spagnolo [2004].
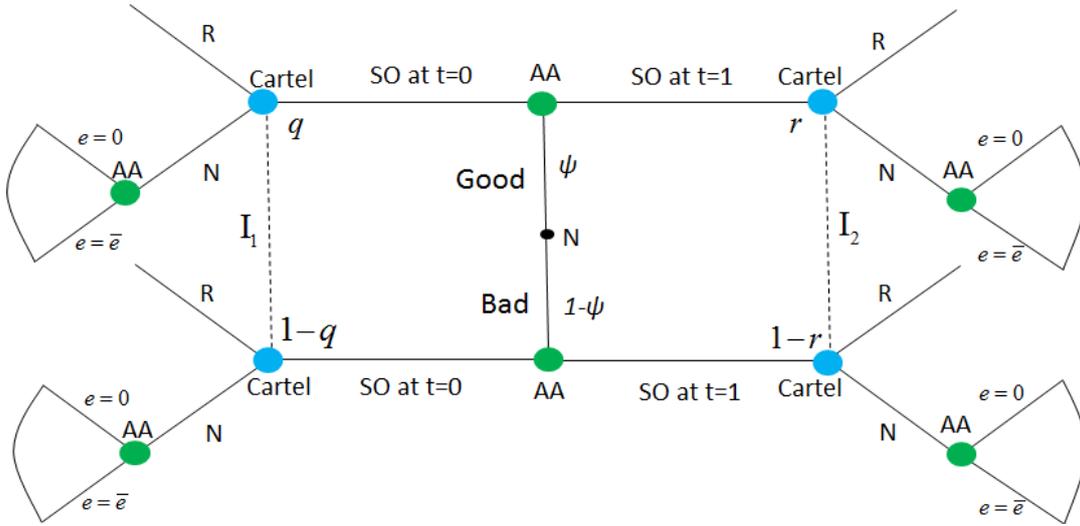
Figure 5: Game tree.

Players

The game consists of three players, 1) Nature, 2) AA and 3) Cartel.

Timing

The game unfolds as follows: 1) nature draws a type for the AA (**G**ood, **B**ad), 2) the AA observes its type and decides in which period to send a SO to the cartel 3) the cartel observes the AA's action but not its type, then update its beliefs and decides whether to self-report or not and 4) the AA decides its prosecutorial effort, if the cartel does not self-report its illegal act.

Strategies

The two types of AA have the same strategy space, $\{(\text{SO at t=0}, \text{SO at t=1}) \times e \in [0, \bar{e}]\}$. Recognizing that the AA will choose its effort level optimally $e = e_{tj}$, where $t = \{0, 1\}$ is the time wherein the AA issues the SO and $j = \{G, B\}$ is the AA's type, we can essentially distinguish four strategies:

1. Send a SO at t=0 and choose effort level $e_{0,G}$ if good type, Send a SO at t=0 and choose effort level $e_{0,B}$ if bad type;

2. Send a SO at t=0 and choose effort level $e_{0,G}$ if good type, Send a SO at t=1 and choose effort level $e_{1,B}$ if bad type;

96

3. Send a SO at t=1 and choose effort level $e_{1,G}$ if good type, Send a SO at t=0 and choose effort level $e_{0,B}$ if bad type;

4. Send a SO at t=1 and choose effort level $e_{1,G}$ if good type, Send a SO at t=1 and choose effort level $e_{1,B}$ if bad type.

The cartel has two information sets, $I_1 \equiv \{(\text{Good, SO at t=0}), (\text{Bad, SO at t=0})\}$ and $I_2 \equiv \{(\text{Good, SO at t=0}), (\text{Bad, SO at t=0})\}$, at each of which it has two actions. So, the cartel has also four strategies:

1. Self-report at I_1 and Self-report at I_2;

2. Self-report at I_1 and Not self-report at I_2;

3. Not self-report at I_1 and Self-report at I_2;

4. Not self-report at I_1 and Not self-report at I_2.

AA's payoffs

Before we derive the payoffs, two remarks are in order. The first is that the SO is issued by the AA at the end of each period, i.e., either at the end of t=0 or the end of t=1. The second is that the historical record of the cartel does not influence the AA's or cartel's decisions. Thus, we normalize all historical harm caused by the cartel to zero. This is without loss of generality since the strategic interaction between the AA and the cartel begins at t=0. The cartel's record may only be considered by the AA when calculating the fine.

The payoff (loss function) to the AA of either type, if it sends a SO at t=0 and the cartel self-reports, is:
$$W_{0,G}^R = W_{0,B}^R = L$$

where L is the per-period social harm caused by the cartel.[206] This is because the SO is sent at the end of t=0, which implies that the social harm caused by the cartel in t=0 cannot be deterred.

The payoff to a good AA when it sends a SO at t=0 and the cartel declines to self-reports is:
$$W_{0,G}^N|_{e=e_{0,G}} = L + \left(\frac{1-p}{1+e}\right)\left(\frac{\delta}{1-\delta}\right)L + e$$

---

[206] A remark for the notation. The superscript denotes the cartel's strategy ('R' is for Self-Report and 'N' is for Not Self-Report). The subscript comprises of two indicators. The first denotes the period wherein the AA sends a SO. It is equal to 0 when the AA sends a SO at t=0 and 1 when the AA sends a SO at t=1. The second denotes the type of the AA ('G' for good and 'B' for Bad).

where $e_{0,G} = \arg\min\limits_{e} W_{0,G}$. Parameter δ denotes the discount rate for future payoffs (common to all players). In words, the payoff for the AA comprises of three elements. The first is the social harm caused by the cartel in period t=0. The second is the present value of the cost of an erroneous decision, i.e., the expected social harm caused by the cartel in the case of a type II error. This cost is estimated by the product of type II error probability and the expected present value of the discounted social harm caused by the cartel in the future. The third is the cost of the AA's prosecutorial effort.

Similarly, the payoff to a bad AA when it sends a SO at t=0 and the cartel does not self-report is:

$$W_{0,B}^{N}|_{e=e_{0,B}} = L + \left(\frac{1}{1+e}\right)\left(\frac{\delta}{1-\delta}\right)L + \lambda e$$

where $e_{0,B} = \arg\min\limits_{e} W_{0,B}$. The cost of a type II error is higher in this case since the AA's prima facie evidence is poor.

The payoff to the AA, of either type, when it sends a SO at t=1 and the cartel self-reports is:

$$W_{1,G}^{R} = W_{1,B}^{R} = L(1+\delta)$$

The welfare loss when the AA sends a SO at t=1 is higher compared to the welfare loss if the AA sends a SO at t=0 since the cartel operates for an additional period. This is precisely the reason why the AA prefers to unravel the cartel sooner than later.

The payoff to a good AA when it sends a SO at t=1 and the cartel does not self-report is:

$$W_{1,G}^{N}|_{e=e_{1,G}} = L + \delta L + \left(\frac{1-p}{1+e}\right)\left(\frac{\delta^2}{1-\delta}\right)L + e$$

where $e_{1,G} = \arg\min\limits_{e} W_{1,G}$. The payoff consists of four elements. The first is the social harm caused by the cartel in period t=0. The second is the social harm caused by the cartel in period t=1. The third represents the expected present value of the discounted social cost of the cartel if it evades detection. It is estimated by the product of the type II error probability and the present value of the social harm caused by the cartel in the future. The fourth corresponds to the cost the AA incurs to prosecute the cartel.

Similarly, the payoff to a bad AA when it sends a SO at t=1 and the cartel

does not self-report is:

$$W_{1,B}^{N}|_{e=e_{1,B}} = L + \delta L + \left(\frac{1}{1+e}\right)\left(\frac{\delta^2}{1-\delta}\right)L + e$$

where $e_{1,G} = \arg\min_e W_{1,B}$.

Collusive firms' payoffs

The expected payoff to a firm, if both firms self-report to an AA which sends a SO at t=0 is:

$$U_{0,G}^{R} = U_{0,B}^{R} = \pi - (1 - \frac{\rho}{2})a\pi$$

The self-reporting firm obtains the period collusive profits π and pays a discounted fine, which is proportional to the collusive profits.[207] Parameter ρ is the fine discount rate and is divided by 2 because only the first self-reporting firm is eligible to receive lenient treatment.[208] Parameter $a \in (0, \frac{1}{\tilde{p}}]$ is the proportion of collusive profits paid as a fine.[209] We assume that a is exogenously set by law. As remarked before, once detected colluding firms eternally compete and obtain the competitive profits. To simplify the exposition the competitive profits are normalized to zero.

The payoff to a firm, if both firms decline to self-report to a good AA that sends a SO at t=0 is:

$$U_{0,G}^{N} = \pi - \left(1 - \frac{1-p}{1+e_{0,G}}\right)a\pi + \left(\frac{1-p}{1+e_{0,G}}\right)\left(\frac{\delta}{1-\delta}\right)\pi$$

The payoff consists of three elements. The first is the period per-firm collusive

---

[207]Proportional fine schemes closely reflect the current antitrust rules on both sides of the Atlantic. In particular, the relevant fine guidelines in the EU and the US specify fines as proportions of illegal (net) profits or damages caused to society. See, *US sentencing guidelines* (chapter 8: Sentencing of organizations), and *Guidelines on the method of setting fines imposed pursuant to article 23(2)(a) of Regulation no 1/2003*, Official Journal of the European Union (2006/C 210/02). Remarkably, most contributions in the literature on LPs (including the seminal papers by Motta and Polo [2003] and Spagnolo [2004]) consider constant fines that are independent of the accrued (net) illegal profits (or damages to society). Among the exceptions that consider proportional fines are the studies by Motchenkova [2004] and Houba et al. [2011].

[208]The first informant rule of the LP implies that the expected fine in case of self-reporting is $\alpha\pi + (1-\rho)a\pi = (1 - \frac{\rho}{2})a\pi$. Fines are capped so that firms' financial stability is not jeopardized. Empirically, fines hardly exceed the illegal profits earned by cartels. For instance, see Craycraft et al. [1997] and Combe and Monnier [2011].

[209]The expected fine cannot be higher than the (net) gain from the illegal conduct, i.e., $\pi \geq \tilde{p}F$ where $F = a\pi$.

profits, the second is the fine a firm pays if it is detected and successfully prosecuted, and the third is the present value of the discounted future collusive profits a cartelist gets if the AA commits a type II error.

The payoff to a firm, if both firms decline to self-report to a bad AA that sends a SO at t=0 is:

$$U_{0,B}^N = \pi - \left(1 - \frac{1}{1 + e_{0,B}}\right) a\pi + \left(\frac{1}{1 + e_{0,B}}\right)\left(\frac{\delta}{1 - \delta}\right)\pi$$

Notice that, contrary to a good AA, if the bad AA exerts no further prosecutorial effort, the firm obtains the discounted stream of full collusive profits. This is because, unless the bad AA exerts effort, the probability of a type II error is 1.

The payoff to each firm, if both firms self-reports to either type of AA that sends a SO at t=1 is:

$$U_{1,G}^R = U_{1,B}^R = \pi - a(1 - \frac{\rho}{2})(\pi + \frac{\pi}{\delta})$$

The cartel receives the period collusive profits and pays a discounted fine that takes into account the cartel's record.

The payoff to a firm, if both firms decline to self-report to a good AA that sends a SO at t=1 is:

$$U_{1,G}^N = \pi - \left(1 - \frac{1 - p}{1 + e_{1,G}}\right)\left[a\left(\pi + \frac{\pi}{\delta}\right)\right] + \left(\frac{1 - p}{1 + e_{1,G}}\right)\left(\frac{\delta}{1 - \delta}\right)\pi$$

The payoff comprises of three elements. The period collusive profits, the fine paid in case of conviction by the AA, which takes into account cartel's record, and the present value of future profits that accrue to the cartel if the AA commits a type II error.

Similarly, the payoff to a firm, if both firms decline to self-report to a bad AA that sends a SO at t=1 is:

$$U_{1,B}^N = \pi - \left(1 - \frac{1}{1 + e_{1,B}}\right)\left[a\left(\pi + \frac{\pi}{\delta}\right)\right] + \left(\frac{1}{1 + e_{1,B}}\right)\left(\frac{\delta}{1 - \delta}\right)\pi$$

Beliefs

Before observing the AA's decisions regarding the timing of issuing a formal SO, the cartel believes that the AA has good quality of evidence (AA is of good type) with probability ψ. With the complementary probability $1 - \psi$, the cartel believes that the AA has poor quality of evidence (AA is of bad type). We define the cartel's beliefs at its two information sets, I_1 and I_2, by $(q, 1 - q)$ and $(r, 1 - r)$,

respectively. Particularly, if the cartel receives a SO at t=0, then it believes that it is issued by a good (bad) AA with probability q $(1 - q)$:

$$prob(good\ AA|I_1) = q \text{ and } prob(bad\ AA|I_1) = 1 - q$$

If the cartel receives a SO at t=1, then it believes that it is issued by a good (bad) AA with probability r $(1 - r)$:

$$prob(good\ AA|I_2) = r \text{ and } prob(bad\ AA|I_2) = 1 - r$$

3.3   Base model: Complete information

In this section, we solve the game assuming that there is no information asymmetry between the AA and the cartel. For example, when the AA truthfully reveals the complete set of its evidence in the SO or when the cartel has acquired, in some way, perfect knowledge about the strength of the AA's prima facie evidence. The base model would allows to impose some restrictions on the parameters' space, to establish a meaningful and interesting signalling game.

3.3.1   Equilibrium analysis

To solve the complete information game we distinguish the two subgames originating from a good AA and a bad AA. Then, we consider the two possible scenarios: 1) the scenario whereby the AA sends a SO at t=0 and 2) the scenario whereby the AA sends a SO at t=1. To derive the equilibrium we use the backward induction algorithm.

Good AA sends a SO at t=0   The optimal effort $e_{0,G}$ for the good AA, if it sends a SO at t=0 and the cartel declines to self-report, is determined by the following optimization problem:

$$e_{0,G} = \arg\min_{e} \left\{ \left(\frac{\delta}{1 - \delta} \right) \left(\frac{1 - p}{1 + e} \right) L + e \right\} \tag{1}$$

In words, the AA chooses its effort intensity to minimize the sum of the expected cost of the type II error and the cost of its effort. The first-order condition (FOC) of (1) implies an optimal effort equal to $e_{0,G} = \sqrt{\frac{\delta(1-p)L}{1-\delta}} - 1$.[210] The optimal effort level equalizes the (social) marginal benefit (reduction in the social cost of the type II error) and the marginal cost of effort (which equals to 1). The AA's effort is positive if the quality of the prima facie evidence is relatively weak. In

---

[210]The SOC for a minimum is satisfied since $\frac{2\delta(1-p)L}{(1-\delta)(1+e)^3} > 0$.

particular, $e_{0,G} > 0$ if $p < p_1$, where $p_1 = \frac{\delta(1+L)-1}{\delta L}$. If, on the contrary, the prima facie evidence is sufficiently potent, i.e., $p > p_1$, the AA optimally prosecutes the cartel without additional prosecutorial effort. The cost of effort, from a welfare point of view, is too high in this case. Moreover, $p_1 \geq 0$ if $L \geq \frac{1-\delta}{\delta} \equiv L_1$. If the social harm instigated by the cartel is relatively low, i.e., lower than L_1, then the AA need not invest additional resources to the case. For the purposes of our study, we restrict $L \geq L_1$ so that effort is always positive if $p < p_1$.

$$e = \begin{cases} e_{0,G} & p < p_1 \text{ and } L > L_1 \\ 0 & \text{otherwise} \end{cases}$$

Simple comparative statics on $e_{0,G}$ reveal that the optimal effort level is inversely related to the quality of the prima facie evidence ($\frac{\partial e_{0,G}}{\partial p} < 0$) and positively related to the seriousness of the cartel ($\frac{\partial e_{0,G}}{\partial L} > 0$). Intuitively, when the AA holds prima facie evidence of higher quality, the type II error probability declines. In this case, a lower level of effort is required by the AA to achieve the optimal law enforcement rate. On the other hand, when the cartel causes a greater harm to society, the cost of type II error escalates. Thus, a higher effort level is required to achieve the optimal law enforcement rate.

The cartel anticipates the AA's effort $e_{0,G}$. If firms do not self-report, they receive an expected payoff equal to:

$$U_{0,G}^N = (1-a)\pi + \frac{[(\delta + a(1-\delta))](1-p)\pi}{\sqrt{\delta(1-\delta)(1-p)L}}$$

If both firms self-report, then each gets a payoff equal to:

$$U_{0,G}^R = \pi - (1 - \frac{\rho}{2})a\pi$$

Firms self-report if the payoff with self-reporting is higher than the one without self-reporting, i.e., $U_{0,G}^R \geq U_{0,G}^N$. The latter inequality holds if $\rho \geq \rho_1$, where $\rho_1 = \frac{2(1-p)[\delta(1-a)+a]}{a\sqrt{\delta(1-\delta)(1-p)L}}$. In words, if the fine discount rate for self-reporting is sufficiently high, then the firms self-report after receiving a SO at t=0 from a good AA. Notice that the AA should not offer positive rewards if the cartel generates a relatively high social harm. In particular, $\rho_1 \leq 1$ if $L \geq \frac{4(1-p)[\delta(1-a)+a]^2}{a^2\delta(1-\delta)} \equiv L_2$. If, on the other hand, the per period social harm caused by the cartel is relatively low, then the firms self-report only if it is provided with a positive reward. The reasoning is that when the social harm instigated by the cartel is relatively low, the benefit from a lower type II error contracts. Given that the productivity of effort diminishes, the AA's optimal level of effort declines. Thus, the expected payoff

to the cartel if it refuses to self-report increases. As a result, the AA must offer courageous incentives to the cartel, i.e., a positive reward, to induce self-reporting.

Lemma 1. *If the social harm instigated by the cartel is relatively low, i.e., $L < L_2$, then the AA must offer a positive reward to induce self-reporting:*

$$\rho_1 \begin{cases} \leq 1 & \text{if } L \geq L_2 \\ > 1 & \text{otherwise} \end{cases}$$

The optimal fine discount rate is decreasing in L, p and a. As the social harm caused by the cartel or the quality of the prima facie evidence improves, the productivity of AA's prosecutorial effort increases. Therefore, the AA intensifies its prosecutorial effort, and thus a lower fine discount rate is required to induce the cartel to self-report. Similarly, with a more severe fine the AA can achieve the optimal enforcement rate with a less generous fine discount. In effect, this is because α and ρ are substitute policy instruments. This is a standard result in the literature on LPs, and it suggests that the model of the paper is sensible.[211]

Good AA sends a SO at t=1 The optimal effort $e_{1,G}$ for the good AA, if it sends a SO at t=1 and the cartel refuses to self-report, is determined by the following optimization problem:

$$e_{1,G} = \arg\min_e \left\{ \left(\frac{\delta}{1-\delta} \right) \left(\frac{1-p}{1+e} \right) L + e \right\} \tag{2}$$

This is identical to the problem of the AA at t=0, as it is represented by equation (1) above. Thus, the optimal effort is $e_{1,G} = \sqrt{\frac{\delta(1-p)L}{1-\delta}} - 1$. As remarked before $e_{1,G} > 0$ if $p < p_1$. Moreover, $p_1 \geq 0$ if $L \geq L_1$.

The cartel anticipates the effort level $e_{1,G}$. If firms do not self-report, then each receives an expected payoff equal to:

$$U_{1,G}^N = [(1-a)\delta - a]\frac{\pi}{\delta} + \frac{\left[\delta^2(1-a) - a\right](1-p)\pi}{\delta\sqrt{\delta(1-\delta)(1-p)L}}$$

If, on the contrary, firms self-report, then each receives a payoff equal to:

$$U_{1,G}^R = \pi - a(1 - \frac{\rho}{2})(\pi + \frac{\pi}{\delta})$$

Firms receive the period collusive profits and pay a reduced fine, which takes into account the cartel's record. Given this, firms self-report if their payoff with

---

[211]See, for instance, Feess and Walzl [2004], Harrington [2008] and Silbye [2010].

self-reporting is higher than that without self-reporting, i.e., $U_{1,G}^R \geq U_{1,G}^N$. The inequality holds if $\rho \geq \rho_2$, where $\rho_2 = \frac{2(1-p)[\delta^2(1-a)-a]}{a(1+\delta)\sqrt{\delta(1-\delta)(1-p)L}}$. Thus, if the fine discount is sufficiently high, and in particular if $\rho \geq \rho_2$, the firms self-report after receiving a SO at t=1 from a good AA. The AA need not offer positive rewards if $L \geq \frac{4(1-p)[\delta^2(1-a)-a]^2}{a^2\delta(1-\delta)(1+\delta)^2} \equiv L_3$.

Lemma 2. *If the social harm instigated by the cartel is relatively low, i.e., $L < L_3$, then the AA needs to offer a positive reward to induce self-reporting:*

$$\rho_2 \begin{cases} \leq 1 & \text{if } L \geq L_3 \\ > 1 & \text{otherwise} \end{cases}$$

The optimal fine discount rate is decreasing in both L and a. The intuition is similar to the one outlined in Section 3.3.1.

AA's decision With the appropriate design of its LP the AA can induce the cartel to self-report at both periods. As shown above the optimal fine discount rates are ρ_1 for t=0 and ρ_2 for t=1. Given that, the AA has to decide in which period to issue and send a SO. If it sends a SO at t=0, and the cartel self-reports, the welfare loss will be $W_{0,G}^R = L$. If, on the other hand, the AA sends a SO at t=1, and the cartel self-reports, then the welfare loss will be $W_{1,G}^R = L + \delta L$. Given that $W_{0,G}^R < W_{1,G}^R$, the AA sends a SO at t=0 and dissolves the cartel sooner.

We summarize the above results in the following lemma.

Lemma 3. *If $e_{0,G} \leq \bar{e}$, then a good AA sends a SO at t=0 and the cartel self-reports. The optimal LP specifies fine discount rates $\rho_1 = \frac{2(1-p)[\delta(1-a)+a]}{a\sqrt{\delta(1-\delta)(1-p)L}}$ at t=0 and $\rho_2 = \frac{2(1-p)[\delta^2(1-a)-a]}{a(1+\delta)\sqrt{\delta(1-\delta)(1-p)L}}$ at t=1.*

It is straightforward to show that $\rho_1 > \rho_2$, which is consistent with the current practice, in many jurisdictions, of rewarding early cooperators with a more generous fine discount. The rationale for this remark is twofold. Firstly, the expected lifespan of the cartel is higher from the period 1 perspective, and, therefore, the stream of the expected collusive profits is higher. Thus, a higher fine discount must be offered to the cartel in order to be compensated for the foregone future collusive profits, when it self-reports at t=0. Secondly, the fine in the second period is higher because it takes into account the cartel's record. Thus, the AA can successfully force the cartel to self-report with a lower fine discount rate.[212]

Corollary 1. *A good AA must offer a higher fine discount to induce self-reporting at t=0 rather than at t=1, i.e., $\rho_1 > \rho_2$.*

---

[212]This is because ρ and α are substitute policy instruments.

Bad AA sends a SO at t=0   The optimal effort $e_{0,B}$ for the bad AA, if it sends a SO at t=0 and the cartel refuses to self-report, is determined by the following optimisation problem:

$$e_{0,B} = \arg\min_e \left\{ \left(\frac{\delta}{1-\delta} \right) \left(\frac{1}{1+e} \right) L + \lambda e \right\} \tag{3}$$

As already mentioned parameter λ represents the opportunity cost of effort for a bad AA that sends a SO at t=0. The FOC of (3) implies an optimal effort equal to $e_{0,B} = \sqrt{\frac{\delta L}{\lambda(1-\delta)}} - 1$. Notice that if λ is sufficiently high, then effort is too costly from a social point of view, and, therefore, the AA prefers to dismiss the case. In particular, if $\lambda \geq \frac{\delta L}{1-\delta} \equiv \bar{\lambda}$ then $e_{0,B} = 0$. If this is the case, the type II error probability equals to 1.

$$e_{0,B} = \begin{cases} 0 & \text{if } \lambda \geq \bar{\lambda} \\ \sqrt{\frac{\delta L}{\lambda(1-\delta)}} - 1 & \text{otherwise} \end{cases}$$

Assumption 5. *For a bad AA that sends a SO at t=0, $\lambda \geq \bar{\lambda}$.*

This assumption ensures a meaningful signalling game since the two types of AA have disparate interests, and thus behave differently. In effect, the restriction on λ implies that the cartel never self-reports when it receives a SO at t=0 from a bad AA, when firms know the weakness of the AA's evidence. Anticipating this reaction by the cartel, the AA postpones the SO for t=1 when the cost of effort is not prohibitively high. In other words, the AA's effort to prosecute the cartel in period t=0 would be fruitless since the cartel knows that the AA does not have good incriminating evidence. As remarked before, the restriction on λ can be justified on the ground that the AA with poor quality of prima facie evidence can allocate its limited resources to other antitrust cases more productively. Equivalently, it is very costly for the AA to divert resources from other activities to the case at hand.

Bad AA sends a SO at t=1   The optimal effort $e_{1,B}$ for a bad AA, if it sends a SO at t=1 and the cartel declines to self-report, is determined by the following optimisation problem:

$$e_{1,B} = \arg\min_e \left\{ \left(\frac{\delta}{1-\delta} \right) \left(\frac{1}{1+e} \right) L + e \right\} \tag{4}$$

Contrary to t=0 the opportunity cost of effort is equal to that of a good AA i.e., $\lambda = 1$. The AA cannot have an open investigation against the cartel forever. Thus, the AA knows that if it does not prosecute the cartel in t=1, it will have to

drop the case. Thus, the opportunity cost of effort is not as high as in t=0 where the AA has the option to postpone the SO for the next period. The FOC of (4) implies an optimal effort of $e_{1,B} = \sqrt{\frac{\delta L}{1-\delta}} - 1$, which is higher than that of a good AA that sends a SO within the same period, i.e., $e_{1,B} \geq e_{1,G}$. This is because a good AA has better quality of prima facie evidence than the bad AA, and thus a lower effort level is required to achieve the optimal level of law enforcement.[213] Notice that the optimal effort for a bad AA is positive if the social harm instigated by the cartel is sufficiently high. In particular, $e_{1,B} > 0$ if $L > L_1$.

The cartel anticipates the effort level $e_{1,B}$. Thus, if firms decline to self-report, they receive a payoff equal to:

$$U_{1,B}^N = [(1-a)\delta - a]\frac{\pi}{\delta} + \frac{\left[\delta^2(1-a) - a\right]\pi}{\delta\sqrt{\delta(1-\delta)L}}$$

If, on the contrary, the cartel self-reports then each firm obtains:

$$U_{1,B}^R = \pi - a(1 - \frac{\rho}{2})(\pi + \frac{\pi}{\delta})$$

Therefore, the firms self-report if $U_{1,B}^R \geq U_{1,B}^N$. The latter inequality holds if $\rho \geq \rho_3$, where $\rho_3 = \frac{2[\delta^2(1-a)-a]}{a(1+\delta)\sqrt{\delta(1-\delta)L}}$. In words, if the fine discount is sufficiently high ($\rho \geq \rho_3$), then the firms self-report after receiving a SO at t=1 from a bad AA. Notice that the AA need not offer positive rewards if $L \geq \frac{4[\delta^2(1-a)-a]^2}{a^2\delta(1-\delta)(1+\delta)^2} \equiv L_3$.

Lemma 4. *If the social harm instigated by the cartel is relatively low, i.e., $L < L_3$, then the AA needs to offer a positive reward to induce self-reporting:*

$$\rho_3 \begin{cases} \leq 1 & \text{if } L \geq L_3 \\ > 1 & \text{otherwise} \end{cases}$$

The optimal fine discount rate is decreasing in L and a. The intuition is as described in Section 3.3.1. Notice that $\rho_3 > \rho_2$. In words, an AA with weak prima facie evidence has to offer a more generous fine discount rate to induce the cartel to self-report. This is intuitive given that the probability of erroneously exonerating the cartel, and thus the cost of type II error, is higher for a bad AA. This observation is consistent with the LPs in real world - the more value added the evidence provided to the AA, the higher the fine discount offered in exchange.

---

[213]Technically, this is attributed to the fact that the second order cross partial derivative of the type II error probability is positive, i.e., $\frac{\partial \tilde{p}}{\partial e \partial p} > 0$. For a graphic illustration see Appendix D.

AA's decision With the appropriate design of its LP the AA can induce the cartel to self-report only in period t=1. As shown above the optimal fine discount rate is ρ_3. Given that, the AA decides in which period to send a SO. If the AA sends a SO at t=0 then the welfare loss will be $W_{0,B}^N = \frac{L}{1-\delta}$. If, on the other hand, the AA sends a SO at t=1 and the cartel self-reports, then the welfare loss will be $W_{1,B}^R = L + \delta L$. Since $W_{1,B}^R < W_{0,B}^N$, the AA prefers to send a SO at t=1.

We summarize the above results in the following lemma.

Lemma 5. *If $e_{1,B} \leq \bar{e}$, then a bad AA sends a SO at t=1 and the cartel self-reports. The optimal LP specifies a fine discount rate $\rho_3 = \frac{2[\delta^2(1-a)-a]}{a(1+\delta)\sqrt{\delta(1-\delta)L}}$.*

Given assumption 5 and using Lemmata 3 and 5, we derive the following proposition.

Proposition 1. *In the game with complete information a good AA sends a SO at t=0, while a bad AA sends a SO at t=1. Once the cartel receives a SO it self-reports. The optimal LP is determined by $\rho_1 = \frac{2(1-p)[\delta(1-a)+a]}{a\sqrt{\delta(1-\delta)(1-p)L}}$ at t=0 and $\rho_3 = \frac{2[\delta^2(1-a)-a]}{a(1+\delta)\sqrt{\delta(1-\delta)L}}$ at t=1.*

The equilibrium of the game with complete information is visualized in Figure 6 below.
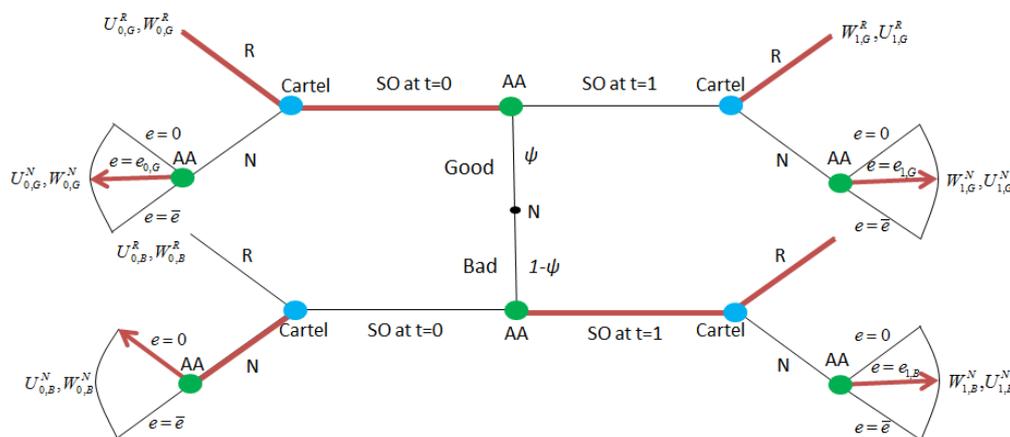
Figure 6: Complete information equilibrium: a good AA sends a SO at t=0 while a bad AA sends a SO at t=1. Firms self-report when they receive a SO by the AA.

3.4 Incomplete information: Equilibrium analysis

In this section, we consider the strategic interaction between the AA and the cartel when they are asymmetrically informed. More precisely, cartelists are uncertain about the quality of the AA's prima facie evidence. However, they hold some prior beliefs regarding the strength of the AA's evidence. Particularly, the firms believes that with probability ψ $(1-\psi)$ the AA holds good (poor) quality of prima facie evidence.

Following Harsanyi [1968] and Kreps and Wilson [1982], we transform the emerging incomplete information game to a game of imperfect information. To solve the game we use the concept of Perfect Bayesian Equilibrium (PBE).[214] Formally, a PBE is a strategy profile coupled with a set of beliefs such that: i) the strategy profile is sequentially rational given beliefs and 2) the system of beliefs is consistent given the strategy profile. Given that the PBE is a relatively weak equilibrium concept, we invoke Cho and Kreps [1987] intuitive criterion to restrict the beliefs for information sets that are off-the-equilibrium path.

The PBE are typically classified in two categories, *pooling* and *separating*.[215] In a pooling equilibrium, all senders choose the same message (action) regardless of their type. Therefore, the sender's message does not convey to the sender any information regarding its type. For instance, in our game, a pooling equilibrium is one where both types of the AA send the SO at the same period. This could be the case if the AA is unwilling to reveal the strength of its prima facie evidence to the cartel. A separating equilibrium is an equilibrium where senders with different types choose different messages (actions). In our game, a separating equilibrium is one where the two types of the AA send a SO at different periods of time. This could be the case if the AA wants to reveal the strength of its evidence to the cartel via the timing in which it sends the SO.

Considering the above, we examine whether or not it pays for the AA to reveal (signal) the strength of its prima facie evidence to the cartel via the time at which it sends the SO. We are particularly interested in the pooling equilibrium whereby both types of the AA send a SO at t=0 and unravel the cartel sooner, even when the evidence the AA retains is weak.

Subsequently we explore the four potential equilibria (in pure strategies):

---

[214] Fudenberg and Tirole [1991b] proved that in a multistage game of incomplete information and with independent types the set of PBE is identical to the set of Sequential equilibria if either each player has at most two possible types or there are only two periods. For a brief discussion of the comparison between a PBE and Sequential equilibria see [Fudenberg and Tirole, 1991a, pp. 345-350].

[215] There is also another category; partially pooling (also termed partially separating or hybrid), whereby all types in a given set of types send a common message, but different sets of types send different messages. In this case, the signal is partially revealing.

1. Pooling equilibrium whereby both types of the AA send a SO at t=0;

2. Pooling equilibrium whereby both types of the AA send a SO at t=1;

3. Separating equilibrium whereby the good AA sends a SO at t=0, while the bad AA sends a SO at t=1;

4. Separating equilibrium whereby the good AA sends a SO at t=1, while the bad AA sends a SO at t=0.

Before we delve into the analysis it is worth restating the beliefs the firms hold at the two information sets, I_1 and I_2, at which it has to make a decision in the game.

<div align="center">

Cartel's beliefs

	good AA	bad AA
SO at t=0	q	$1-q$
SO at t=1	r	$1-r$

</div>

3.4.1 Pooling equilibrium: AA sends a SO at t=0

In this subsection, we investigate whether a pooling equilibrium whereby both types of the send a SO at t=0 is feasible. Crucially, in a pooling equilibrium the time in which a SO is sent does not convey or reveal to the cartel any additional information with regard to the AA's true type. Thus, when both types of the AA send a SO at t=0, the cartel cannot infer the quality of the prima facie evidence the AA holds. Thus, posterior and prior beliefs coincide i.e., $q = \psi$ and $1 - q = 1 - \psi$. If the firms self-report then they obtain the period collusive profits and pay the reduced fine:

$$V_0^R = \pi - (1 - \frac{\rho}{2})a\pi$$

If, on the contrary, they do not self-report then given their beliefs regarding the AA's type, their expected payoff is:

$$V_0^N = \psi\, U_{0,G}^N + (1 - \psi)\, U_{0,B}^N$$

Given firms' beliefs regarding the AA's prima facie evidence, their payoff comprises of two elements. The first is the expected payoff a firm obtains if it declines to self-report to a good AA that sends a SO at t=0. The second is the expected payoff a firm obtains if it declines to self-report to a bad AA that sends a SO at t=0.

Considering the above the firms self-report at I_1 if $V_0^R \geq V_0^N$. This inequality holds if $\psi \geq \bar{\psi}$, where:

$$\bar{\psi} = \frac{\left[\delta + a(1 - \frac{\rho}{2})(1 - \delta)\right] \sqrt{\delta(1 - \delta)(1 - p)L}}{\left[\delta(1 - a) + a\right] \left[\sqrt{\delta(1 - \delta)(1 - p)L} - (1 - \delta)(1 - p)\right]}$$

In words, if the firms believe that the AA is of a good type with a relatively high probability, greater than the threshold value of $\bar{\psi}$, then they self-report after receiving a SO at t=0. Note that, when the social harm instigated by the cartel is sufficiently low, and in particular $L < \frac{4(1-p)[\delta(1-a)-a]^2}{a^2\rho^2\delta(1-\delta)} \equiv L_4$, then $\bar{\psi} > 1$. In this case, the pooling BNE whereby both types of AA send a SO at t=0 cannot be sustained.

Thus, assuming that i) the opportunity cost of effort for a bad AA at t=0 is sufficiently high i.e., $\lambda > \bar{\lambda}$, ii) the AA offers a fine discount ρ_3 at I_2, so that the firms self-report at $t = 2$ regardless of their beliefs and iii) the effort level $e_{1,B}$ is feasible (i.e., $e_{1,B} < \bar{e}$), we derive the following proposition:

Proposition 2. *In the incomplete information game the AA commits to a LP such that $\rho = \max\{\rho_1, \rho_3\}$. Both types of AA send a SO at t=0. The cartel self-reports whenever it receives a SO, if $\psi > \bar{\psi}$ and $\forall r \in [0, 1]$.*

This pooling PBE satisfies the *intuitive criterion* for $r = 0$. Intuitively, a good AA does not have any incentives to deviate from its strategy, given that it is a strictly dominant strategy for a good AA to send a SO at t=0. This implies that if the cartel members receive a SO at t=1, then it must have been sent by a bad AA. Therefore, the beliefs off-the-equilibrium path must be such that $(r, 1-r) = (0, 1)$.

Corollary 2. *The pooling PBE whereby the AA sends a SO at t=0 satisfies the intuitive criterion for $r = 0$.*

Comparative statics on $\bar{\psi}$ reveal that given firms' beliefs, it is easier for the AA to induce the firms to self-report at t=0, even if it is of a bad type, with a more generous LP ($\frac{\partial \bar{\psi}}{\partial \rho} < 0$). This implies that the optimal fine discount must be at the maximum feasible level. The lure of a higher fine discount can be used to facilitate AA's objective to induce the cartel to self-report at an earlier stage of the proceedings. Similarly, the more effective the AA's preliminary investigation in detecting good quality of incriminating evidence, the easier for the AA to compel the cartel to self-report at t=0. When the AA's preliminary investigation is very effective in exposing good quality of evidence. Moreover, it is easier to sustain this pooling equilibrium when the social harm caused by the cartel is higher ($\frac{\partial \bar{\psi}}{\partial L} < 0$). This is

attributed to the fact that the AA's optimal effort $e_{0,G}$ is higher when confronted with more serious cartels, implying a lower type II error. The magnitude of the fine also makes it easier to attain this pooling equilibrium ($\frac{\partial \bar{\psi}}{\partial a} < 0$). A higher fine increases the expected fine paid by the cartel in case of conviction by the AA. Finally, it is also more difficult to obtain this pooling equilibrium when the cartel is more robust, i.e., it ascribes a higher value to future profits ($\frac{\partial \bar{\psi}}{\partial \delta} > 0$).[216]

Corollary 3. *Given the cartel's prior beliefs regarding the AA's types, it is easier to sustain the pooling equilibrium whereby the AA sends a SO at t=0 and the cartel subsequently self-reports with higher i) social harm caused by the cartel, ii) fine level, iii) fine discount rate and iv) quality of prima facie evidence, and with lower vi) (expected) cartel duration and iv) discount rate.*

As shown above, a higher fine discount rate sharpens firms' incentives to denounce the cartel. In light of this, the analysis suggests that the optimal fine discount rate is the maximum feasible. This result is consistent with Becker [1968].

Corollary 4. *The optimal LP provides the maximum feasible fine to the first self-reporting firm.*

As remarked in Section 3.3.1, given that the cartel is harmful to society, the AA strictly prefers to send a SO at t=0 and dissolve the cartel earlier. Thus, the pooling equilibrium whereby both types of the AA send a SO at t=0 and the cartel subsequently self-reports is (weakly) superior, from a welfare point of view, to the complete information equilibrium. This is straightforward since with incomplete information a bad AA deceives the cartel to self-report in the first period. In contrast, the cartel never self-reports to a bad AA, and that is why in the setting with complete information the bad AA prefers to send a SO at t=1. An implication of this analysis is that the desistance effects of the LP may enhance with incomplete information. In other words, when the cartelists are uncertain about the quality of the AA's prima facie evidence, the AA may fruitfully deceive the cartel to self-report, even though the strength of its evidence is poor.

From the above analysis we arrive at the following recommendations. Firstly, the AA, should refrain from revealing the strength of its prima facie evidence. It is possible to elicit self-reporting even if the AA has weak prima facie evidence against the cartel. Secondly, the AA should manipulate the cartel's prior beliefs or act strategically in a manner that deceives the cartel regarding the strength of its

---

[216]Another interpretation for parameter δ is that it reflects how long it takes for the AA to convict the cartel if the latter does not self-report its unlawful conduct. In effect, a higher δ implies a lengthier period to convict the cartel. With this interpretation, a more competent and efficient AA, can successfully conclude the prosecution of the cartel sooner. Thus, such an AA can elicit self-reporting at an earlier stage of the proceedings despite having poor evidence.
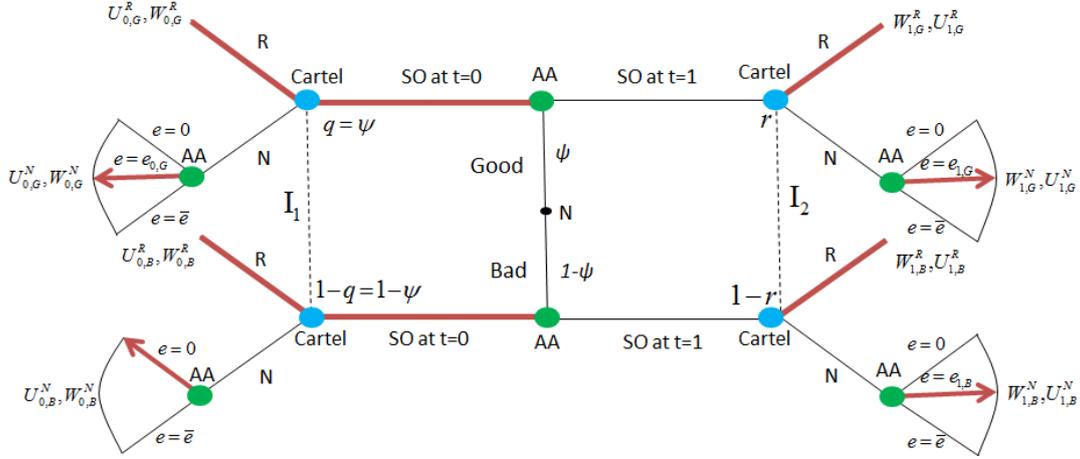
Figure 7: Pooling BNE: both types of AA send a SO at $t = 0$. Firms always self-report.

evidence. The more biased the cartel's beliefs, the easier to induce self-reporting at an earlier stage of the prosecutorial proceedings, despite having weak prima facie evidence. Thirdly, the AA should prioritize the prosecution of more serious cartels, i.e., those cartels that cause a greater harm to society. As remarked before, it is easier for the AA to compel more serious cartels to self-report, in spite of having poor quality of prima facie evidence.[217]

The above analysis is visualized in Figure 7 below.

3.4.2 Pooling equilibrium: AA sends a SO at t=1

As with the previous pooling equilibrium, the SO does not communicate to the cartel any additional information regarding the AA's types (or equivalently the AA's strength of evidence). Consequently, the cartel cannot revise its prior beliefs. In this case, posterior beliefs at information set I_2 are such that: $(r, 1 - r) = (\psi, 1 - \psi)$. Given these beliefs, if both firms self-report after receiving a SO at t=1, then each firm obtain the period collusive profits and pay a discounted fine, which takes into account the cartel's record:

$$V_1^R = \pi - a(1 - \frac{\rho}{2})\left(\pi + \frac{\pi}{\delta}\right)$$

---

[217]To illustrate the analysis consider the following numerical example. Suppose that $a = 1$, $\delta = 0.6$, $p = 0.8$, $\rho = 0.6$, $L = 50$, $\pi = 100$. Then, our equilibrium analysis predicts that $\bar{\psi} = 0.93$. If $\rho = 1$ (full amnesty) then $\bar{\psi} = 0.84$.

If the firms do not self-report at t=1, then the expected payoff is:

$$V_1^N = \psi U_{1,G}^N + (1 - \psi)U_{1,B}^N$$

The first (second) element is the expected payoff to the cartel given its beliefs that the SO is issued by a good (bad) AA. By construction of the model and to simplify the analysis we assume that the fine discount rate is such that the cartel always self-reports at I_2, regardless of the cartel's beliefs. This requires $\rho = \rho_3$. Moreover, from Section 3.4.1 we know that the cartel does not self-report at information set I_1 if the firms strongly believe that the SO was issued by a bad AA. In particular, the cartel declines to self-report if $q < \bar{\psi}$. To begin with, suppose that $q < \bar{\psi}$. Now, consider the AA's incentives given the (anticipated) cartelists' behavior at information sets I_1 and I_2. Notice that the bad AA would never have any incentive to deviate and send a SO at t=0. This is because the quality of prima facie evidence is poor, and thus the cost of prosecutorial effort is prohibitively high. This is not the case for a good AA though. Given the firms' beliefs and their anticipated behaviour at I_1, the AA would prefer to send a SO at t=0 if $W_{0,G}^N > W_{1,G}^N$. This inequality holds if $p < \bar{p}$, where

$$\bar{p} = \frac{L^2\delta^5 + (L^2 + 2L)\delta^4 + (1 + 2L - L^2)\delta^3 + (1 - 2L - L^2)\delta^2 + (2L - 1)\delta - 1}{4\delta L}$$

In words, if the AA has poor quality of prima facie evidence, then given the firms' beliefs, it will be better for the AA to postpone the SO for t=1. It is too costly and risky to convict the cartel in t=0. If, on the other hand, the AA holds good quality of prima facie evidence, then it is better to send a SO at t=0, even if the firms decline to self-report. The cost of postponing issuing the SO is greater than the risk of erroneously exonerating the cartel at t=0.

Thus, assuming that i) the opportunity cost of effort for a bad AA at t=0 is sufficiently high i.e., $\lambda > \bar{\lambda}$, ii) the AA offers a fine discount ρ_3 at I_2, so that the firms self-report at $t = 2$ regardless of their beliefs and iii) the effort level $e_{1,B}$ is feasible (i.e., $e_{1,B} < \bar{e}$), we derive the following proposition:

Proposition 3. *In the incomplete information game both types of the AA send a SO at t=1 and the cartel self-reports if the cartel's beliefs are such that $r = \psi$ and $q < \bar{\psi}$ and the quality of the AA's prima facie evidence is such that $p < \bar{p}$. Moreover, the LP specifies a fine discount ρ_3.*

Given the cartel's beliefs at the information set I_1 (off-the-equilibrium), the pooling BNE whereby both types of AA send a SO at t=1 satisfies the intuitive criterion. As mentioned before only a good AA would have incentives to deviate

from this pooling BNE. But if the firms' beliefs are such that $q < \bar{\psi}$, the good AA complies with the equilibrium strategy prescription.

Corollary 5. *The pooling PBE whereby the AA sends a SO at t=1 and the cartel self-reports satisfies the intuitive criterion.*

The above analysis is visualized in Figure 8 below.
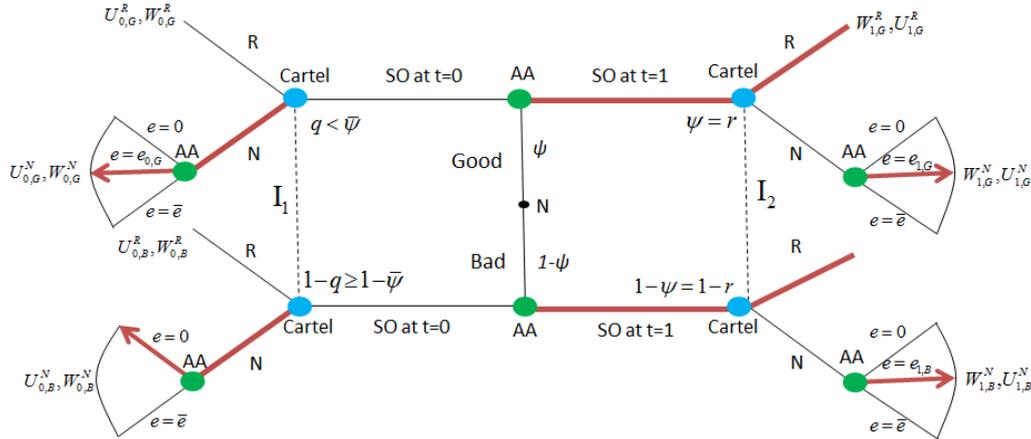
Figure 8: Pooling BNE: both types of AA send a SO at t=1. At t=1 firms' self-report only when they receive a SO from a good AA. Firms always self-report at t=2.

3.4.3 Separating equilibrium: good AA sends a SO at t=0, bad AA sends a SO at t=1

To begin with, suppose that the good AA sends a SO at t=0, while the bad AA sends a SO at t=1. Then, the firms update 'their beliefs, given that both information sets at which it has to make a decision, I_1 and I_2, are on the equilibrium path. Using the Bayes rule the firms' posterior beliefs are: $q = 1$ and $r = 1$. Given their beliefs, and provided that i) the LP is such that $\rho > \max\{\rho_1, \rho_3\}$, and ii) effort level $e_{1,B}$ is feasible, i.e., $e_{1,B} \leq \bar{e}$, the firms self-report regardless of the period in which the SO is sent by the AA, i.e., at both information sets I_1 and I_2. Notice, however, that given that the cartel self-reports at information set I_1, a bad AA has an incentive to deviate and send a SO at t=0 rather than at t=1, and thus unravel the cartel sooner. Therefore, this cannot be a separating BNE.

The above analysis is visualized in Figure 9 below.
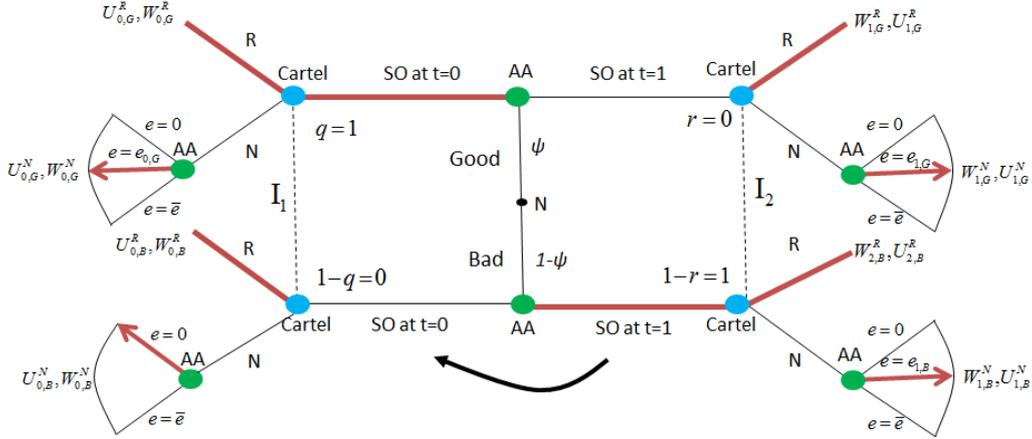
Figure 9: Separating equilibrium: a good type of AA sends a SO at $t = 0$, while a bad type of AA sends a SO at $t = 1$. Firms self-report regardless of the strength of the AA's evidence.

3.4.4 Separating equilibrium: good AA sends a SO at t=1, bad AA sends a SO at t=0

Suppose that the bad AA sends a SO at t=0, while the good AA sends a SO at t=1. Then, the cartel updates its beliefs, given that both information sets at which it has to make a decision, I_1 and I_2, are on the equilibrium path. Using the Bayes rule the firms' posterior beliefs are: $q = 0$ and $r = 1$. Given their beliefs, and provided that i) the LP offers a fine discount of ρ_1, and ii) effort levels $e_{1B} = 0$ and $e_{2G} \leq \bar{e}$, the firms decline to self-report at I_1, while it self-reports at I_2. Thus, if a SO is send at t=0 by a bad AA the cartel remains unconvicted since the probability of a type II error is 1. Notice, however, that given that the cartel self-reports at information set I_2, a bad AA has an incentive to deviate and send a SO at t=1, and thus unravel the cartel. Therefore, this cannot be a separating BNE.

The above analysis is visualized in Figure 9 below.

3.5 Concluding remarks

A prominent feature characterizing the strategic interaction between an AA and a cartel is the existence of asymmetric information. The latter can be two sided. On the one hand, firms know whether they participate in a cartel while, on the other hand, an AA knows the strength of its evidence against the cartel. Unlike previous literature, this paper explores the implications of a scenario where an AA has an informational advantage *vis-a-vis* a cartel. In particular, within the framework of
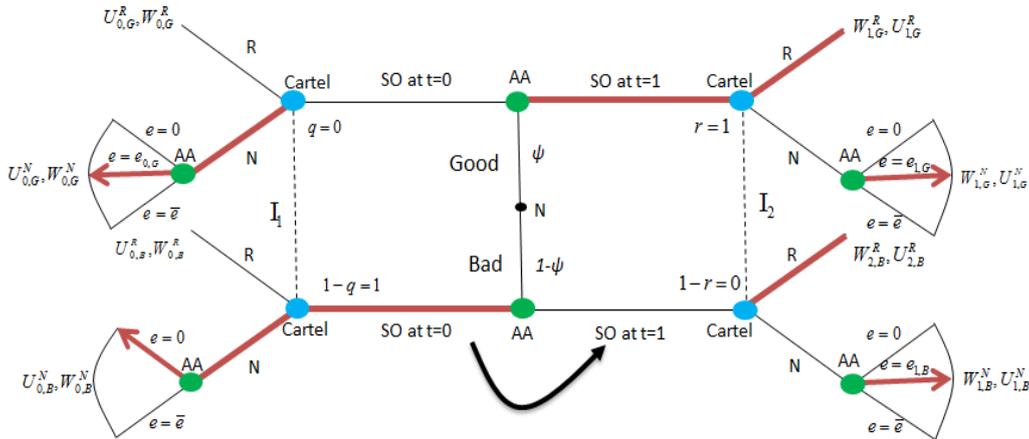
Figure 10: Separating equilibrium: a good AA sends a SO at t=1 while a bad AA sends a SO at t=0. Firms self-report only at t=2.

a signalling game, we address whether the AA can exploit its informational lead and deceive the cartel about the strength of its evidence, ultimately inducing self-reporting. A crucial assumption of our study is that the AA makes its decisions regarding the intensity of its investigations on a case-by-case basis, after reviewing some preliminary evidence concerning the inspected cartel. This is a realistic and desirable assumption and it reflects the actual practice of many AAs in various jurisdictions. Importantly, it allows to overcome potential dynamic inconsistency problems which may arise *ex-post* if the AA pre-commits to a certain prosecutorial effort.[218]

The role of the message in our signalling game is played by the time wherein the AA sends a SO and initiates the formal proceedings against the cartel. An AA with better quality of evidence prefers to send a SO to the cartel at an earlier stage of the prosecution proceedings. This is because the risk of erroneously exonerating the cartel, if it declines to self-report, is not as high for an AA which retains high quality of evidence compared to an AA which has poor quality of evidence. In this context, our equilibrium analysis shows that it is feasible for the AA to deceive the cartel and induce it to self-report even though it has deficient incriminating evidence.

A logical question that arises is why the AA does not scrutinize all possible evidence before issuing a SO. Apart form the obvious cost of doing so, the analysis of this study suggests that it is not essential for the AA to process all feasible

---

[218]Although issuing a SO is the standard procedural rule in the EU, the scenario we consider can be easily tailored to reflect the procedural rules in other jurisdictions. For instance, the grand jury plea bargaining system in the US, which strongly favors early settlements.

evidence in order to secure a conviction. By behaving in a strategic manner, the AA can elicit self-reporting by colluding firms even if it has poor prima facie evidence. We do not assert that an AA should write whatever comes to mind or lie in the SO or in the press releases[219]. What we claim is that the AA should be prudent in the way it lays down its evidence against the cartel. Additionally, it must be very careful in the approach it sets forth the arguments on which it will base its preliminary conclusion. Different ways of presenting or making an allegation may concretely impact on a cartel's beliefs, and, ultimately, its behavior.[220] This is because the cartel is uncertain about the strength of the AA's evidence. The fact that colluding firms receive a SO at an earlier stage of the prosecutorial process may confound them about the AA's evidence. They know that a good AA prefers to send a SO at an earlier stage of the prosecutorial process, but, at the same time, they know that a bad AA has potent incentives to mimic the conduct of a good AA. A strategic AA can take advantage of its informational lead. More formally, it can convert a separating equilibrium (which an early SO would uniquely signal strong evidence) into a pooling equilibrium. With the lure of a sufficiently generous fine discount, the AA can fruitfully undermine the cartel's vulnerable internal trust. As a result, colluding firms race to the AA in order to be the first to self-report and benefit from the LP.

The role of a cartel's (prior) beliefs is of pivotal importance. The more biased these beliefs are towards compelling evidence, the easier it is for the AA to induce the cartel to self-report at an earlier stage of the antitrust proceedings. A critical question that arises is where these beliefs come from and how they are formed. Realistically, these beliefs are molded on the basis of an AA's competency in detecting and processing the evidence collected from its own investigations. In this respect, we can distinguish between two types of AAs: profound and superficial. A profound AA, which opens only a few investigations, may be fruitful in exposing high quality of incriminating evidence. On the other hand, a superficial AA, which opens a large number of investigations, may not be as competent as a profound AA in detecting good quality of evidence or filtering the collected (crude and bulky) evidence.[221] Clearly, the depth of an AA's investigations influences the cartel's

---

[219]For instance, see the case of the large supermarkets cartel in the UK. In that particular case, the Office of Fair Trading was forced to make an unprecedented public apology and pay damages to Morrisons, after admitting that a press release (OFT, press release 134/07, 20/09/2007) deceptively *upgraded* the provisional findings into claims of price-fixing.

[220]A similar argument holds for the announcements an AA releases during the investigations (e.g. after a dawn raid). However, this issue is more subtle. An AA must protect suspected firms from speculations or rumors. At the same time, the AA may leave suspected firms with a nuance of ambiguity with respect to the detected evidence.

[221]A similar argument holds for experienced AAs v. inexperienced AAs. This is consistent with Hinloopen [2003] who argues that the AA can increase the likelihood of detecting cartels if it creates a reputation for exposing cartels.

beliefs. This reasoning is consistent with the findings of a recent experimental study by Dijkstra et al. [2011]. The authors show that self-reporting is significantly higher with a profound AA. Thus, a policy implication of our study is that AAs should open a few profound investigations rather than a large number of superficial investigations. A meticulous AA which profoundly performs its investigations may be able to establish a reputation of being effective in winning the cases it prosecutes to courts. When such reputation is established, the AA can exploit it by issuing an early SO even with weak evidence. As long as colluding firms self-report, so the weakness of evidence never becomes public knowledge, the AA's reputation will not be undermined. So, an early establishment of reputation with a few profound cases is essential.

The LP plays a critical role too. As in Motta and Polo [2003], the LP is a powerful tool to enhance welfare, even after an investigation is launched in the market. We claim that, contrary to the existing literature, firms' incentives to apply for leniency when an investigation is underway are potent. The argument is valid as long as the cartel members are unaware about the strength of the AA's incriminating evidence. In particular, we show that the more generous the LP, the more pronounced the cartel desistance effects. This implication is in accord with previous literature advocating for sufficiently generous fine discount rates to self-reporting firm(s). Moreover, it underpins the recent amendment of the EU LP so that to provide amnesty to the first firm to denounce the cartel, even if an investigation is underway.

The deterrence effects of LPs are not explicitly modeled in our study. Nonetheless, the argument or recommendation for profound market investigations (profound AAs), which impacts on cartel's beliefs, indirectly hints to the positive cartel deterrence effects. The underlying mechanism of our signalling game implies that the investigation resources of the AA can be efficiently used to increase the frequency (number) of investigations, and thus the effectiveness of the leniency programme in deterring cartels.

In our study, self-reporting occurs only if the AA credibly threatens cartels with a sufficiently high prosecutorial effort. Thus, the AA can successfully induce colluding firms to self-report if it has a sufficiently high budget to finance its investigations. Nonetheless, since the AA does not commit in advance to a certain prosecutorial effort, its budget may remain under-spent. The fact that the budget is unspent may create a temptation for a government to cut the AA's budget. The analysis of this paper suggests that this would undermine both self-reporting and deterrence. In addition, the AA can strategically overstate the resources to be invested in the case at hand.[222] Although the AA's budget is publicly available

---

[222]For instance, it can announce the total number of employees to handle the case including the administration staff, rather than the number of case handlers and experts.

and accessible, the resources of the AA to be directed in a particular case is not public information.

Our study also supports the recent direction of AAs in various jurisdictions to prioritize the prosecution of more serious cartels in terms of the gravity of the harm caused to society. All the same, the more harmful the cartel, the higher the cost of an erroneous decision by the AA. Thus, an AA aiming to minimize the social cost of cartels would intensify its prosecutorial effort when confronted with more serious cartels. This implies a higher probability of conviction, and thus a higher expected fine, which makes self-reporting more appealing.

As a final remark, we highlight the fact that the signalling game of this paper can apply to the US (and Canada as well) grand jury plea bargaining system and more generally to the enforcement of generic law (e.g. criminal law).

Chapter 4: Optimal Leniency Programme with Detection Avoidance

4.1 Introduction

The economic literature on law enforcement crystallizes that crime deterrence depends on the expected fine. The latter is determined by the product of the actual fine and the probability of detection.[223] A prominent limitation to the effectiveness of public policies in the area of crime deterrence is avoidance. In general, avoidance involves effort by violators aimed at decreasing the expected fine by reducing either the probability of detection or the magnitude of the fine if convicted.[224] Detection avoidance activities are costly for violators. Essentially, the cost of effort to avoid detection constitutes an additional component to the expected fine. To this extend, such activities influence the mechanics of deterrence, and, in consequence, the optimal enforcement policy mix.[225]

Although avoidance and its social cost play a pivotal role in the conventional literature on (generic) law enforcement, it has been given very little attention by the economic literature on antitrust law enforcement. Avoidance is undoubtedly a widespread phenomenon and it seems to be a reasonable strategy for firms, especially when the stakes from collusion are high. Just as an antitrust authority (or more generally an enforcement agency) invests resources to monitor markets and uncover cartels, colluding firms invest in avoiding that detection.

Detection avoidance include all activities that influence the probability of cartel detection (e.g., meetings in exotic resorts, concealment of incriminating evidence, use of encryption mechanisms, hire third parties to organize and administer the cartel, etc.). The firms that engage in such activities incur a direct cost. Importantly, the level of the cost depends on the intensity of firms' effort to avoid detection. This type of avoidance activities needs to be distinguished from strategic firms' behavior which may indirectly influence the probability of detection. For instance,

---

[223]For a survey of the literature see Garoupa [1997] and Polinsky and Shavell [2000].

[224]Avoidance activities encompass various expenditures engaged by violators in order to reduce their exposure to the risk of punishment. For instance, installing radar detectors to avoid speeding tickets, lobbying politicians to relax the enforcement of regulations, bribing an enforcement agent, destroying or concealing incriminating evidence, investing in costly litigations.

[225]For instance, Malik [1990] maintains that high fines may exacerbate socially costly detection avoidance activities. Thus, he argues that fines should not simply be set as high as possible, contrary to Becker [1968]. Moreover, Nussim and Tabbach [2009] show that a more severe fine may lead to more crime, since avoidance and crime are generally complements. Similarly, Jensen et al. [2012] show that a tougher fine policy may lead to more severe crime being committed. Finally, Innes [2001] establishes an additional benefit of self-reporting schemes when avoidance is possible, beyond those identified by Kaplow and Shavell [1994]. More precisely, those violators that self-report do not engage in socially wasteful detection avoidance activities.

firms anticipate that the detection probability is function of the price changes, so they design a prudent price sequence that avoid raising suspicions to consumers or antitrust authorities about the cartel.[226]

There is empirical evidence documenting that colluding firms undertake considerable efforts to avoid detection by antitrust authorities. According to Stephan [2010] efforts to conceal a cartel include: destruction of incriminating evidence and other abuses of evidentiary rules and procedures, cover-up of evidence, meetings in luxury hotels, cook of accounting books, use of encrypted messages, litigation effort, etc.[227]

Despite significant advancements to the economic literature on leniency programmes, the influence of firms' detection avoidance activities on the design of leniency programmes has not been adequately explored. A failure to take into account avoidance when designing leniency policies entails a real risk of corroding firms' incentives to self-report their unlawful conduct.

In this paper, we extend the analysis by Chen and Rey [2012] to incorporate firms' effort to avoid detection. The major objective is to explore the impact of firms' decisions to engage in detection avoidance activities on the leniency programme that achieves the optimal level of cartel deterrence. We show that the intensity of firms' avoidance effort increases with the frequency of investigations and the level of the fine and decreases with the cost of avoidance effort. Moreover, we demonstrate that the optimal fine discount rate depends positively on the severity of the fine and negatively on the probability of investigation and the cost of avoidance activities. Contrary to Chen and Rey [2012], we show that when the fine is relatively low and the cost of avoidance effort relatively high no leniency should be offered. Finally, we show that a leniency programme that fails to take into consideration firms' decisions to undertake effort to avoid detection is suboptimal, in the sense that it results in under-deterrence. These results are particularly interesting and useful for the policy designers. More broadly, the paper contributes to the theoretical literature on the design of optimal leniency programmes.

---

[226]In Harrington [2004, 2005] and Chen and Harrington [2007], the probability of cartel detection depends on firms' intertemporal pricing decisions. In particular, large price movements raise more concerns to buyers and antitrust authorities about collusion, implying a higher probability of detection.

[227]See, for instance, the recent decisions by the European Commission: COMP 39.437 - *TV and computer monitor tubes*; COMP 39.125 - *Car glass*; COMP 38,823 - *Elevators and escalators*; COMP 38.899 - *Gas insulated switchgear*; COMP 36.490 - *Graphite electrodes.*

4.1.1 Relevant literature

This paper is related to the economic literature on detection avoidance[228] and leniency programmes[229].

Malik [1990] provides the first systematic study of the implications of detection avoidance for the design of law enforcement. The author concludes that contrary to Becker [1968], the fine should not be set at its maximum feasible level. This is because a higher fine triggers more socially costly avoidance effort by offenders.[230]

The interplay between avoidance and crime is explored by Nussim and Tabbach [2009]. Contrary to Malik [1990] who assumes that offenders choice of crime is a binary decision, the author allows for different levels of crime to be committed. An essential insight of Nussim and Tabbach [2009] is that crime and avoidance are generally complements: more crime induces more avoidance and vice-versa. Intuitively, more crime increases the marginal benefits of investments in avoidance, which reduces the marginal cost of engaging to criminal activities.

Innes [2001] extends the analysis by Malik [1990] when criminals can self-report their illegal act in exchange of a reduced fine. His analysis highlights another benefit of self-reporting schemes, namely, avoidance costs are reduced. This is because those criminals who self-report do not engage in socially costly avoidance activities.[231]

A distinguishing characteristic of the papers discussed above is that they consider non-repeated single crimes committed by individuals. Below we discuss some papers which deal with repeated crimes committed by groups of individuals, such as cartels.

Avramovich [2010] explores the optimal allocation of colluding firms' resources to productive effort that reduces the cost of production and avoidance effort that decreases the probability of cartel detection. The author shows that the optimal allocation of firms' resources depends on the strictness of the antitrust policy (magnitude of the fine and frequency of investigations). A more severe antitrust policy has two effects. On the one hand, it increases the cost of collusion and, thus, enhance deterrence. On the other hand, it increases the resources devoted to unproductive avoidance activities (equivalently, it increases the productive inefficiency) and, thus, reduce welfare. Importantly, Avramovich [2010] demonstrates that a sufficiently generous leniency programme can attain the same level of deterrence

---

[228] For a lucid survey on detection avoidance and enforcement theory see Sanchirico [2011]

[229] For an excellent literature review on leniency programmes see Spagnolo [2008].

[230] However, Langlais [2008] shows that when criminals efforts to avoid detection and the enforcement agency's effort to detect crime are strategic complements, then the maximum feasible fine can be optimal. In particular, it is optimal to set the fine at its maxumum feasible level when criminals avoidance effort is more sensitive to the frequency than to the severity of the fine.

[231] Interestingly, Innes [2001] shows that Becker's rule for setting the fine at its maximum feasible level is restored.

with high fines, but at a lower cost, in terms of firms' productive inefficiencies.

In a recent paper, Jensen et al. [2012] investigate the relationship between punishment and the severity of the crime in the context of antitrust enforcement. Contrary to Nussim and Tabbach [2009], the severity of the crime is endogenous. The authors show that if colluding firms can engage in detection avoidance activities, then a tougher punishment may lead to higher cartel overcharges. Intuitively, when the overcharge and the fine are positively related, there is a trade-off between the benefit from a higher overcharge and the cost from a higher expected fine. Thus, if a higher fine leads to more avoidance activity, and, thereby, a lower probability of detection, then it may also lead the cartel to attain a higher overcharge in equilibrium. This result also holds when the portfolio of policy instruments includes a leniency programme. Similar to Aubert et al. [2006], Buccirossi and Spagnolo [2006b] and Motchenkova and Leliefeld [2010], the leniency programme provides a credible device to thwart defections from the illegal cartel.[232]

The economic literature on leniency programmes has proliferated over the last decade after the pioneer contributions by Motta and Polo [2003] and Spagnolo [2004]. A general finding of the relevant literature is that by eroding a cartel's internal trust, leniency programmes may evoke substantial desistance and deterrence effects (e.g. Motta and Polo [2003], Spagnolo [2004], Ellis and Wislon [2003] and Harrington [2008]). The direction and magnitude of these effects depend on the level of market concentration (Ellis and Wislon [2003]), the degree of evidence heterogeneity (Feess and Walzl [2004] and Silbye [2010]), the heterogeneity in firms' market shares (Motchenkova and Leliefeld [2010]), whether fines are fixed or proportional to the accumulated collusive profits (Motchenkova [2004]), whether the probability of cartel detection deterministically or stochastically changes over time (Hinloopen [2003] and Harrington [2008]) and whether the antitrust policy instruments (probability of detection and penalties) are endogenous to firms' pricing behaviour (Chen and Harrington [2007]). Nevertheless, leniency programmes are not unambiguously optimal. Wrongly designed leniency programmes may exacerbate collusive behaviour or provoke cartel formation (e.g. Motta and Polo [2003], Buccirossi and Spagnolo [2006a], Aubert et al. [2006], Motchenkova and Leliefeld [2010] and Jensen et al. [2012]). More recently, Chen and Rey [2012] explore the trade-off between greater deterrence and lower desistance. On the one hand, a more generous fine discount sharpens colluding firms' incentives to defect and self-report to the antitrust authority (desistance effect). On the other hand, firms can strategically exploit the generosity of the leniency programme and form a cartel that in the absence of the leniency programme would be unprofitable (deterrence effect).

---

[232]The authors demonstrate that this counterintuitive result holds as long as firms cannot retain illegal profits after detection.

The remainder of the paper is organized as follows: Section 4.2 describes the game and its main assumptions; Section 4.3 presents the equilibrium analysis and derives the optimal leniency programme; Section 4.5 concludes.

4.2 The model

Consider two symmetric firms which play an infinitely repeated game. In particular, the two firms, which maximize their expected payoff in each period, decide whether to collude or compete in the presence of antitrust enforcement. To simplify the analysis we consider a stationary antitrust policy which treats all markets alike.[233] The net payoff from collusion equals Π for each firm. A firm that deviates in the market obtains 2Π. We assume that firms use the same discount factor $\delta \in (0,1)$. To sustain collusion firms adopt grim-trigger strategies. That is, once a deviation occurs, firms revert to cut-throat competition forever. For simplicity, the competitive profits are normalized to zero. In a context without antitrust enforcement, collusion can be sustained if $\delta \geq \frac{1}{2}$. Thus, in order to have an interesting game we assume that $\delta \geq \frac{1}{2}$, implying that collusion is a real concern.

The starting point for the analysis with antitrust enforcement is that the cartel leaves incriminating evidence which can be exposed by the antitrust authority. To keep things simple we assume that the evidence lasts only for one period, thus the cartel cannot be prosecuted for its past actions.[234] Due to budget constraints the antitrust authority cannot monitor all markets. The probability of inspecting a particular market is $p \in (0,1)$. This probability reflects the antitrust authority's budget or the effectiveness of its investigations. Furthermore, each colluding firm can invest resources $a_i \geq 0$, $i = 1,2$, in avoidance activities to reduce the probability of cartel detection. Thus, the effective probability of detection of a cartel depends on the exogenous probability of firms' inspection and on the endogenous probability of exposing incriminating evidence after an inspection at firms' premises. In particular, the effective probability of detection is given by:

$$\tilde{p}(a_1, a_2, p) = \frac{p}{e^{a_1+a_2}} \tag{5}$$

---

[233]The assumption of stationarity implies that the probability of investigating a market in every period is the same. Moreover, the level of the fine for convicted cartels is also the same regardless of the number of previous convictions. Essentially, we abstract from the possibility of tighter monitoring of detected cartels and increased fines for repeated offenders.

[234]Existing literature reveals that relaxing this assumption gives rise to an additional effect of leniency programmes. A firm that deviates in the market can simultaneously self-report in order to preempt other firms from self-reporting in the future. Spagnolo [2004] calls this effect the 'protection form punishment effect' while Harrington [2008] calls it the 'deviator effect'. In principle, this effect mitigates the costs associated with a deviation from the collusive agreement and, thereby, make collusion easier to sustain. See also Chen and Harrington [2007].

It depends negatively on the intensity of firms' avoidance efforts (a_1 and a_2) and positively on the exogenous probability of market investigation (p). Moreover, the probability of detection has the following properties: (1) $\lim\limits_{a_i,a_j\to 0}\tilde{p} = p$, (2) $\lim\limits_{a_i,a_j\to\infty}\tilde{p} = 0$, (3) $\lim\limits_{a_i\to 0}\tilde{p} = \frac{p}{e^{a_j}} < p$, (4) $\frac{\partial\tilde{p}}{\partial a_i} < 0$, (5) $\frac{\partial^2\tilde{p}}{\partial a_i^2} > 0$, (6) $\frac{\partial^2\tilde{p}}{\partial a_i\partial a_j} > 0$ and (7) $\frac{\partial^2\tilde{p}}{\partial a_i\partial p} < 0$. In words, property (1) implies that when firms do not invest at all in avoidance activities the probability of detection equals the exogenous probability of inspection. Property (2) implies that when firms invest an infinite amount of resources in avoidance, the probability of detection tends to zero. Property (3) implies that firm's i avoidance effort is productive, in the sense that it decreases the effective probability of detection, even if firm j exerts no effort to avoid detection. The fact that one firm does not invest resources to avoid detection does not imply that the other firm's avoidance activity will be futile in reducing the risk of detection. Property (4) implies that the probability of detection decreases with the amount of resources a firm invests in avoidance effort. Property (5) implies that the marginal productivity of a firm's avoidance effort diminishes as it invests more resources in detection avoidance activities. In particular, the effectiveness of each unit of additional effort to avoidance activities for firm i decreases with the amount of effort already made by firm i. Property (6) implies that the productivity of firm's i avoidance effort increases in firm's j effort to avoid detection.[235] In effect, this assumption implies that firms' decisions with regard to the intensity of their avoidance efforts are strategic substitutes.[236] Properties (5) and (6) imply that the firms' detection avoidance technology exhibits diminishing returns. Property (7) implies that the effectiveness of antitrust enforcement and firms' avoidance

---

[235]One way to think about this property is the following. Suppose that both firms retain duplicate copies of hard incriminating evidence. The antitrust authority can detect incriminating evidence from both firms. Thus, firms must destroy the same pieces of evidence; otherwise the antitrust authority can discover the evidence destroyed by one firm from the other firm, even though with a lower probability. In this situation, firms' efforts to avoid detection are more fruitful when both firms destroy the same pieces of evidence. The more effort firm j exerts to destroy the incriminating evidence it retains, the more likely it is for firm i to destroy the same pieces of evidence as firm j does. Put it differently, the productivity of firm's i effort to destroy the incriminating evidence increases as firm j destroys more evidence.

[236]This contrasts with Avramovich [2010] who assumes that firm's i effort to avoid detection becomes less productive as firm j invests more resources in detection avoidance. The latter implies that firms' avoidance efforts are strategic complements. One justification for this assumption is that firms retain distinct pieces of incriminating evidence. Thus, the pieces of evidence destroyed by firm i cannot be discover by the antitrust authority's investigations in the firm's j premises. Moreover, the more pieces of evidence firm i destroys, the less evidence remains available to be discovered by the antitrust authority. Therefore, the productivity of a firm's effort to destroy evidence (avoid detection) is decreasing in both, it's own effort and the other firm's effort.

It is worth noting that the difference between the two papers lies in the adopted functional form of the effective probability of detection.

activities are related. The level of one factor influences the marginal effectiveness of the other factor. In particular, firms' avoidance efforts become less productive in reducing the probability of detection when the frequency of inspection increases, and vice versa.[237] In effect, this implies that firms' avoidance efforts and public effort in monitoring and detecting illegal cartels are strategic complements in the effective probability of detection. A consequence of this property is that the elasticity of the effective probability of detection with respect to the exogenous probability of detection is one, i.e., $\frac{\partial \tilde{p}}{\partial p} \frac{p}{\tilde{p}} = 1$.[238]

The cost of avoidance effort is:

$$C_i(a_i) = ke^{a_i}$$

where $k = 0$ if $a_i = 0$ and $k > 0$ if $a_i > 0$. Parameter k reflects the efficiency of firms' avoidance technology. A higher k implies a higher unit cost of effort. Note that the technology of avoidance exhibits decreasing returns to scale; the cost of avoidance effort is increasing at an increasing rate as firms invest more resources in detection avoidance activities, i.e., $\frac{\partial C_i(a_i)}{\partial a_i} > 0$ and $\frac{\partial^2 C_i(a_i)}{\partial a_i^2} > 0$, $\forall a_i > 0$.[239]

When the cartel is detected, either because of the antitrust authority's own investigations or because of a leniency application, firms must pay a fine. The leniency programme provides a fine discount θ only to the first self-reporting firm. Thus, the first informant pays only $(1 - \theta)F$, while the other firm pays the full fine F. If both firms simultaneously lodge a leniency application, then each firm pays an expected fine $(1 - \frac{\theta}{2})F$. The antitrust policy parameters are p, F and θ. To simplify the analysis and focus on the fine discount rate that achieves the optimal level of deterrence, p and F are exogenously fixed.

In each period, the game unfolds as follows:

---

[237] As we will see later, a consequence of this property of the effective probability of detection is that firms' avoidance efforts are influenced by both, the antitrust authority's enforcement efforts (reflected in p) and the level of the fine (F).

[238] Langlais [2008] show that when the private expenditures in avoidance and the public expenditures in the detection of criminals are complements, then setting the fine at its maximum feasible level is optimal contrary to Malik [1990]. Intuitively, although a higher fine increases wasteful avoidance, a proportionate decrease in enforcement efforts reduces avoidance even more. A requirement for this result is that criminals must be more sensitive to the risk of detection than to the severity of the fine, i.e., the elasticity of the criminal's effort to avoid detection with respect to public monitoring is at least a high as the elasticity of the criminal's effort with respect to the fine. The latter requirement is also satisfied in the model of this paper. This may offer a justification for letting F denote the maximum feasible fine in this paper.

[239] The adopted functional forms of the effective probability of detection and the cost function associated with detection avoidance effort ensure a well-defined optimization problem.

- Stage 0 (Cartel formation) - Firms decide whether to collude or not. If at least one firm chooses not to collude, firms compete and the game ends for that period; otherwise:

- Stage 1 (Avoidance) - Each firm decides how many resources to invest in avoidance activities.

- Stage 2 (Defection)- Each firm decides whether to comply or not to the illegal agreement. These decisions become known to rival firms at the end of the period.

- Stage 3 (Leniency application and Policy implementation) - Firms decide whether to self-report their illegal act to the antitrust authority. If at least one firm self-reports, the cartel is detected with probability 1. Otherwise, the cartel is detected with probability $\tilde{p}$.

4.3 Equilibrium Analysis

4.3.1 Normal collusion

When a leniency programme is not available, firms cannot benefit from self-reporting to the antitrust authority. Thus, in every period each firm receives collusive profits Π and pay a fine F if the cartel is detected by the antitrust authority. Moreover, firms pay the cost related to their avoidance effort. The expected discounted value of collusion is:

$$V_N = \frac{\Pi - \left(\frac{p}{e^{a_i+a_j}}\right) F - ke^{a_i}}{1-\delta}, \text{ for } i = 1, 2 \text{ and } i \neq j$$

where the subscript N stands for *normal collusion*. Similar to Chen and Rey [2012], we define '*normal collusion*' as the degree of collusion that can be sustained when a leniency programme is not included in the set of policy instruments.

Definition (Normal collusion). *Normal collusion is defined as the level of collusion that could be sustained in an environment with antitrust enforcement without a leniency programme.*

Firms choose their avoidance effort in order to maximize the value of collusion:

$$a_i^* = \arg\max_{a_i} V_N, \text{ for } i = 1, 2 \text{ and } i \neq j$$

From the above optimization problem we derive the best response function of firm i:

$$a_i^*(a_j) = \frac{1}{2}\left[\ln\left(\frac{pF}{k}\right) - a_j\right]$$

Note that $\frac{\partial a_i^*(a_j)}{\partial a_j} = -\frac{1}{2} < 0$, i.e., firms' detection avoidance activities are strategic substitutes. The higher firm's j avoidance effort, the lower the marginal contribution to firm's i profits of each unit of own's avoidance effort.

Given the symmetry of the game, the equilibrium avoidance effort of each firm is:

$$a^* = a_1 = a_2 = \frac{1}{3}\ln\left(\frac{pF}{k}\right) \tag{6}$$

This optimal level of avoidance effort balances the marginal contribution of each unit of avoidance effort to firm's profits, through a reduction in the probability of detection, and the marginal cost of an additional of effort.[240]

Note that if the fine is sufficiently low, then firms do not invest resources in avoidance activities at all. Moreover, observe that the intensity of firms' avoidance activity increases, at a diminishing rate, with the antitrust policy parameters p and F.[241] This is intuitive, as the higher the (exogenous) frequency of an investigation (or the magnitude of the fine), the more resources colluding firms invest in avoidance activities in order to decrease the expected fine. Furthermore, firms investment in avoidance activities decrease (at a increasing rate) with the cost of avoidance effort.

Lemma 1. *Firms' optimal avoidance effort is* $a^* = \begin{cases} \frac{1}{3}\ln\left(\frac{pF}{k}\right) & \text{if } F > \frac{k}{p} \\ 0 & \text{otherwise} \end{cases}$. *It decreases with k and increases with p and F.*

Taking the above into consideration, the present value of the expected payoff with *normal collusion* is:

$$V_N = \frac{\Pi - 2p^{\frac{1}{3}}F^{\frac{1}{3}}k^{\frac{2}{3}}}{1-\delta} \tag{7}$$

When a leniency programme is not available a firm can only deviate in the market. Thus, the value of deviation is:

---

[240] The second order sufficient conditions for a maximum are satisfied: $\frac{\partial^2 V_N}{\partial a_i^2} = \frac{\frac{-pF}{e^{a_i+a_j}} - ke^{a_i}}{1-\delta} < 0$,

$\forall a_i \geq 0$. Moreover, uniqueness is guaranteed: $\frac{\partial^2 V_N}{\partial a_i^2} + \frac{\partial^2 V_N}{\partial a_i a_j} = \frac{\frac{-pF}{e^{a_i+a_j}} - ke^{a_i}}{1-\delta} + \frac{-pF}{e^{a_i+a_j}(1-\delta)} < 0$,

$\forall a \geq 0$. Finally, stability is guaranteed: $\frac{\partial a_1^*(a_2)}{\partial a_2}\frac{\partial a_2^*(a_1)}{\partial a_1} = (-\frac{1}{2})(-\frac{1}{2}) < 1$. For the conditions that ensure uniqueness and stability see [Vives, 2001, p. 47 and p. 51].

[241] All comparative statics and calculations can be found in Appendix E.

$$V^d = 2\Pi - \left(\frac{p}{e^{a_i^* + a_j^*}}\right) F - ke^{a_i^*}, \ \text{for } i = 1, 2 \text{ and } i \neq j$$

Deviation from *normal collusion* is not profitable if the collusive profits are sufficiently high, i.e., $\Pi > \Pi_N \equiv \frac{2\delta p^{\frac{1}{3}} F^{\frac{1}{3}} k^{\frac{2}{3}}}{2\delta - 1}$. When the per-firm collusive profits are lower than Π_N, the temptation to renegade from the collusive agreement undermines the trust between firms and, thus, the cartel cannot be sustained. Notice that Π_N increases (at a diminishing rate) with p, F and k. In words, it is harder to sustain collusion with a higher frequency of market investigations, magnitude of the fine and cost of avoidance effort.

Lemma 2. *Normal collusion is sustained if $\Pi > \Pi_N$, where $\Pi_N = \frac{2\delta p^{\frac{1}{3}} F^{\frac{1}{3}} k^{\frac{2}{3}}}{2\delta - 1}$. It is harder to sustain normal collusion with higher p, F and k.*

4.3.2 A Leniency Programme is available

When a leniency programme is available, then a firm that deviates from the collusive agreement can run to the antitrust authority and self-report in exchange of a lower fine. The value for a deviant firm that self-reports is:

$$V^{dr} = 2\Pi - (1 - \theta)F$$

Thus, with a leniency programme, the optimal deviation depends on the level of the fine discount. In particular, a deviating firm would not only deviate in the market but also self-report if $V^{dr} > V^d$. This inequality holds when the fine discount rate is sufficiently high, i.e., $\theta > 1 - \frac{2p^{\frac{1}{3}} k^{\frac{1}{3}}}{F^{2/3}} \equiv \underline{\theta}$. Note that $\underline{\theta}$ increases (at a diminishing rate) with F while it decreases (at an increasing rate) with p and k. In words, the higher the frequency of market investigations or the cost of avoidance effort and the lower the fine, the more attractive a leniency application after a deviation in the market would be. Intuitively, an increase in the fine intensifies firms' engagement in avoidance activities that reduce the probability of detection. Given that avoidance effort is costly, a more generous fine discount should be offered by the antitrust authority in order to elicit self-reporting by colluding firms. With a higher rate of market investigations, the option to self-report becomes more tempting for firms. This is because with self-reporting firms pay a lower certain fine (rather than a higher expected fine) and, at the same time, save on avoidance activities. Similarly, an increase in the cost of avoidance effort makes firms to invest less resources in avoidance activities, implying a higher probability of detection. Thus, again, the option to self-report becomes more appealing for firms.

Lemma 3 (Desistance). *If $\theta > \underline{\theta}$, where $\underline{\theta} = 1 - \frac{2p^{\frac{1}{3}} k^{\frac{1}{3}}}{F^{2/3}}$, then a deviating firm*

is better-off to self-report to the antitrust authority after deviating in the market. Self-reporting is more appealing to firms when p and k are low and F is high.

To proceed with the analysis, suppose that $\theta > \underline{\theta}$. This implies that after a firm deviates in the market, it self-reports to the antitrust authority. In this case, *normal collusion* is sustained if $V_N \geq V^{dr}$. This inequality holds if collusion is sufficiently lucrative, i.e., $\Pi \geq \frac{2p^{\frac{1}{3}}k^{\frac{2}{3}}F^{\frac{1}{3}} - (1-\theta)(1-\delta)F}{2\delta - 1} \equiv \Pi_N^{LP}(\theta)$. It is straightforward to see that a more generous leniency programme (i.e., higher θ) makes it more challenging to sustain *normal collusion*.[242]

$$\frac{\partial \Pi_N^{LP}(\theta)}{\partial \theta} = \frac{(1-\delta)F}{2\delta - 1} > 0$$

Lemma 4. *When a leniency programme is available and provided that the fine discount rate is sufficiently high, i.e., $\theta > \underline{\theta}$, where $\underline{\theta} = 1 - \frac{2p^{\frac{1}{3}}k^{\frac{1}{3}}}{F^{2/3}}$, collusion is sustained if $\Pi \geq \Pi_N^{LP}(\theta)$, where $\Pi_N^{LP}(\theta) = \frac{2p^{\frac{1}{3}}k^{\frac{2}{3}}F^{\frac{1}{3}} - (1-\theta)(1-\delta)F}{2\delta - 1}$. Normal collusion is harder to sustain with a more generous fine discount.*

It is straightforward to see that $\Pi_N^{LP}(\theta) > \Pi_N$. This implies that the effectiveness of antitrust policy is enhanced with the introduction of a leniency programme.

4.3.3   Alternative collusive strategies

The set of collusive strategies expands when a leniency programme is available. For instance, firms can adopt an alternative strategy, namely, to collude and systematically self-report to the antitrust authority. To simplify the exposition, we consider a strategy where firms collude and report in every period. We label this strategy as '*Collude and Report*'. When firms adopt this strategy, then each of them appropriates the collusive profits Π and pays a discounted fine with probability $\frac{1}{2}$, since only the first self-reporting firm is eligible to a lenient treatment. Thus, the present value of the expected payoff if firms collude and self-report in every period is:

$$V_R = \frac{\Pi - (1 - \frac{\theta}{2})F}{1 - \delta}$$

If a firm deviates from this strategy, it obtains a payoff equal to:

$$V_R^d = 2\Pi - (1 - \frac{\theta}{2})F$$

---

[242]The impact of leniency programmes on a cartel's sustainability has been identified, among others, by Spagnolo [2004], Harrington [2008] and Houba et al. [2011].

Considering the above, a deviation from '*Collude and Report*' is not profitable if collusive profits are sufficiently high. In particular '*Collude and Report*' can be sustained if $V_R \geq V_R^d$. This inequality holds if $\Pi \geq \frac{(2-\theta)\delta F}{2(2\delta-1)} \equiv \Pi_R(\theta)$. It is straightforward to see that a more generous leniency programme makes it easier to sustain this alternative collusive strategy.

$$\frac{\partial \Pi_R(\theta)}{\partial \theta} = -\frac{\delta F}{2(2\delta-1)} < 0$$

Note that if the fine discount is sufficiently high, i.e., $\theta > 2\left(1 - \frac{2p^{\frac{1}{3}}k^{\frac{2}{3}}}{F^{\frac{2}{3}}}\right) \equiv \bar{\theta}$, then this alternative collusive strategy is more robust than normal collusion without a leniency programme, i.e., $\Pi_R(\theta) < \Pi_N$. The argument is that a sufficiently generous leniency programme can be exploited by firms and establish a cartel which in the absence of a leniency programme would be unprofitable. Thus, a leniency programme can enhance cartel formation contrary to its objective.[243]Observe that '*Collude and Report*' is easier to sustain with a higher probability of investigation (p) and cost of avoidance effort (k) and lower fine (F).

Lemma 5 (Deterrence). *'Collude and Report' is more robust than normal collusion without a leniency programme if $\theta > \bar{\theta}$, where $\bar{\theta} = 2\left(1 - \frac{2p^{\frac{1}{3}}k^{\frac{2}{3}}}{F^{\frac{2}{3}}}\right)$.*

4.4 The Optimal fine discount rate

The above analysis highlights the trade-off between cartel desistance and cartel deterrence.[244] On the one hand, a more generous leniency programmes makes cartels more fragile, and, thus, more difficult to sustain (see Lemma 4). On the other hand, a sufficiently generous leniency programme may enable some firms to establish a cartel that would be unprofitable without a leniency programme (see Lemma 5).

From the analysis of Sections 4.3.1, 4.3.2 and 4.3.3, *normal collusion* can be sustained regardless of whether a leniency programme is available if:

$$\Pi \geq \Pi_N(\theta) = \max\{\Pi_N, \Pi_N^{LP}(\theta)\}$$

while 'Collude and Report' can be sustained if:

---

[243]The pro-collusive effect of leniency programmes was initially pointed out by Motta and Polo [2003].

[244]See, Motta and Polo [2003], Chen and Harrington [2007], Harrington [2008] and Chen and Rey [2012].

$$\Pi \geq \Pi_R(\theta)$$

Thus, the deterrence threshold is:

$$\Pi(\theta) \equiv \min\{\Pi_N(\theta), \Pi_R(\theta)\}$$

In words, those cartels with profits less than $\Pi(\theta)$ are deterred. Notice that, by definition, the leniency programme does not facilitate *normal collusion* nor does it provokes new cartels to form. Given that the antitrust authority's objective is to deter cartels in as many industries as possible, the optimal fine discount rate should maximize the deterrence threshold:

$$\hat{\theta} = \arg\max_{\theta} \Pi(\theta)$$

As remarked above, the introduction of a leniency programme makes collusion more fragile if $\theta > \underline{\theta}$ and does not foster alternative forms of collusion if $\theta < \bar{\theta}$. Since $\theta < \bar{\theta}$, it is optimal to offer a fine discount rate $\theta \in (\underline{\theta}, \bar{\theta})$, so that to deter collusion in industries where, absent a leniency programme, *normal collusion* could be sustained. Moreover, given that $\Pi_N^{LP}(\theta)$ increases in θ (i.e., destabilizes *normal collusion*) but, at the same time, decreases Π_R (i.e., facilitate '*Collude and Report*' strategies), the optimal leniency programme is such that $\Pi_N^{LP}(\theta) = \Pi_R(\theta)$. This implies that the optimal discount rate is:

$$\hat{\theta} = \frac{2\left(F - 2p^{\frac{1}{3}}F^{\frac{1}{3}}k^{\frac{2}{3}}\right)}{(2-\delta)F} \tag{8}$$

To visualize the analysis above consider Figure 11 below. The horizontal axis of the graph represents the fine discount rate, while the vertical axis represents the per-firm collusive profits. The green line depicts Π_N, the purple line Π_N^{LP} and the red line Π_R.

Notice that Π_N^{LP} has a positive slope as an increase in the fine discount rate (more generous leniency programme) reinforce firms' incentives to deviate in the market and simultaneously self-report to the antitrust authority, i.e., it makes it harder to sustain the cartel. Moreover, Π_R has a negative slope as an increase in the fine discount rate makes it more profitable for firms to collude and report in every period. Observe that Π_N is independent of the fine discount rate, since it reflects the threshold for normal collusion in absence of a leniency programme. Without a leniency programme all cartels with per-firm collusive profits lower than Π_N are deterred.

When $\theta > \underline{\theta}$, then a deviant firm prefer to self-report to the antitrust authority and therefore the threshold for cartel deterrence is given by Π_N^{LP}. Moreover, when

$\theta > \bar{\theta}$ firms can exploit the generosity of the leniency programme by adopting collude and report strategies. For $\theta > \bar{\theta}$, this strategy adversely affects deterrence since for $\Pi_N > \Pi_R$.

The optimal fine discount rate $\hat{\theta}$ is determined by the intersection of Π_N^{LP} and Π_R. Note that the threshold of deterrence with an optimally designed leniency programme which takes into account firms' decisions to engage in detection avoidance activities $\hat{\Pi}_R(\hat{\theta})$ is greater than the threshold for normal collusion Π_N. Thus, an optimally designed leniency policy is effective in deterring cartels.
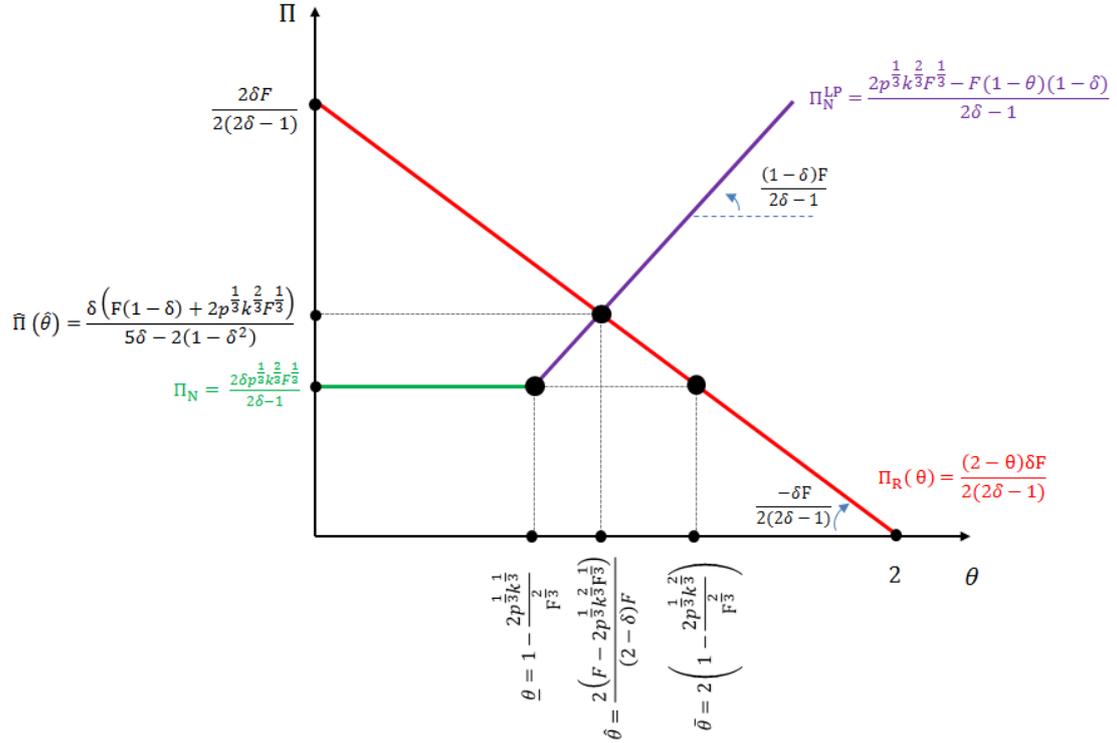

Figure 11: Optimal fine discount rate when firms can engage in detection avoidance activities.

Note that the optimal fine discount rate is increasing (at a diminishing rate) with the severity of the fine. Intuitively, when the fine increases, firms profoundly invest in costly avoidance activities. Thus, a more generous fine discount must be offered by the antitrust authority in order to elicit self-reporting by firms. Moreover, notice that the optimal fine discount rate is decreasing (at an increasing rate) with the frequency of market investigations and the cost of avoidance effort. If the antitrust authority conducts fewer market investigations (i.e., the audit rate decreases), then firms would invest less in costly avoidance activities. In this case,

the antitrust authority can attain the optimal level of deterrence with a lower fine discount rate. Similarly, when the cost of avoidance activities increases, firms' efforts to avoid detection becomes less intense. Thus, the antitrust authority can attain the optimal level deterrence with a less generous leniency programme.

Moreover, observe that it is desirable to offer a positive reward to a self-reporting firm, i.e., $\hat{\theta} > 1$, if a) $p < \frac{\delta^3 F^2}{64k^2}$ or b) $F > \frac{8k\sqrt{p}}{\delta^{\frac{3}{2}}}$ or c) $k < \frac{\delta^{\frac{3}{2}}F}{8\sqrt{p}}$. In words, if a) the frequency of market investigations is low so that the probability of detection is relatively low or b) the cost of avoidance effort is relatively low so that firms' invest heavily in avoidance activities or c) the magnitude of the fine is sufficiently high so that the expected fine is relatively high, then the leniency programme should be more generous. In contrast to Chen and Rey [2012], it may be optimal not to offer a lenient treatment if the frequency of inspection is relatively high $(p > \frac{1}{2})$ and the fine is relatively low $(F \leq 2k\sqrt{2p})$.

Taking into consideration that: i) $\frac{k}{p} \lessgtr 2k\sqrt{2p}$ if $p \gtrless \frac{1}{2}$ and ii) $2k\sqrt{2p} < \frac{8k\sqrt{p}}{\delta^{\frac{3}{2}}}$, we derive the following proposition:

Proposition 1. *The optimal leniency programme when firms can invest resources to avoid detection is characterized by* $\hat{\theta} = \frac{2\left(F - 2p^{\frac{1}{3}}F^{\frac{1}{3}}k^{\frac{2}{3}}\right)}{(2-\delta)F}$. *The optimal leniency programme is such that:*

1. $\hat{\theta} = 0$ (no leniency) if $p > \frac{1}{2}$ and $F \in (\frac{k}{p}, 2k\sqrt{2p}]$

2. $\hat{\theta} \in (0, 1]$ if:

    (a) $p = \frac{1}{2}$ and $F \in (\frac{k}{p}, \frac{8k\sqrt{p}}{\delta^{\frac{3}{2}}}]$ or

    (b) $p > \frac{1}{2}$ and $F \in (2k\sqrt{2p}, \frac{8k\sqrt{p}}{\delta^{\frac{3}{2}}})$

3. $\hat{\theta} > 1$ (positive reward) if:

    (a) $p < \frac{1}{2}$ and $F > \frac{k}{p}$ or

    (b) $p \geq \frac{1}{2}$ and $F > \frac{8k\sqrt{p}}{\delta^{\frac{3}{2}}}$

The threshold for collusive profits that characterizes the effectiveness of the leniency programme is:

$$\hat{\Pi}(\hat{\theta}) = \frac{\delta\left[F(1-\delta) + 2p^{\frac{1}{3}}F^{\frac{1}{3}}k^{\frac{2}{3}}\right]}{5\delta - 2(1+\delta^2)}$$

134

This threshold determines the effectiveness of the antitrust policy when the portfolio of policy instruments includes a leniency programme. It implies that cartels with profits less than $\hat{\Pi}(\hat{\theta})$ are deterred.

Corollary 1. *The level of optimal deterrence when colluding firms can engage in avoidance activities is determined by $\hat{\Pi}(\hat{\theta})$, where $\hat{\Pi}(\hat{\theta}) = \frac{\delta\left[F(1-\delta)+2p^{\frac{1}{3}}F^{\frac{1}{3}}k^{\frac{2}{3}}\right]}{5\delta-2(1+\delta^2)}$.*

It is straightforward to see that $\hat{\Pi}(\hat{\theta}) > \Pi_N$. This implies that the optimal leniency programme enhances the efficacy of the antitrust policy, as more cartels are deterred when a leniency programme is available.

Moreover, notice that $V_N > V_R(\hat{\theta})$. This implies that when the fine discount rate is set at the optimal level, then firms would not adopt '*Collude and Report*' strategies. Although this type of strategies could be sustained, firms would prefer to collude without reporting to the antitrust authority.

4.4.1 Comparison with Chen and Rey (2012)

If the antitrust authority does not take into account firms' decisions to engage in costly efforts to avoid detection, then, as shown by Chen and Rey [2012], the optimal fine discount rate is:

$$\check{\theta} = \frac{2(1-p)}{2-\delta}$$

It is straightforward to see that $\check{\theta} \gtrless \hat{\theta}$ if $F \lessgtr \frac{2\sqrt{2}k}{p}$. In words, if the antitrust authority fails to take into consideration firms' effort to avoid detection, then the leniency programme would be suboptimal. In particular, if the leniency programme provides excessive leniency (i.e., $\check{\theta} > \hat{\theta}$) then it would be: a) easier to sustain '*Collude and Report*', since the gains from defecting in the market are reduced ($\Pi_R(\theta)$ decreases with θ) and, b) harder to sustain *normal collusion*, since the gains from a deviation in the market along with a leniency application increase ($\Pi_N^{LP}(\theta)$ increases with θ). Therefore, cartels with profits $\Pi \in \left(\check{\Pi}_R(\check{\theta}), \hat{\Pi}_R(\hat{\theta})\right)$ which would otherwise not have been sustained, can establish a viable cartel by adopting '*Collude and Report*' strategies. Thus, a wrongly designed leniency programme that provides a fine discount rate greater than the optimal level (i.e., $\theta > \hat{\theta}$) results in under-deterrence. This is illustrated in Figure 12 below.

On the other hand, if the leniency programme provides insufficient leniency (i.e., $\check{\theta} < \hat{\theta}$), then it would be: a) easier to sustain *normal collusion*, since the gains from deviating in the market with a simultaneous self-report to the antitrust authority diminish ($\Pi_N^{LP}(\theta)$ increases with θ) and, b) harder to '*Collude and Report*' since the gains from defecting in the market increase ($\Pi_R(\theta)$ decreases with
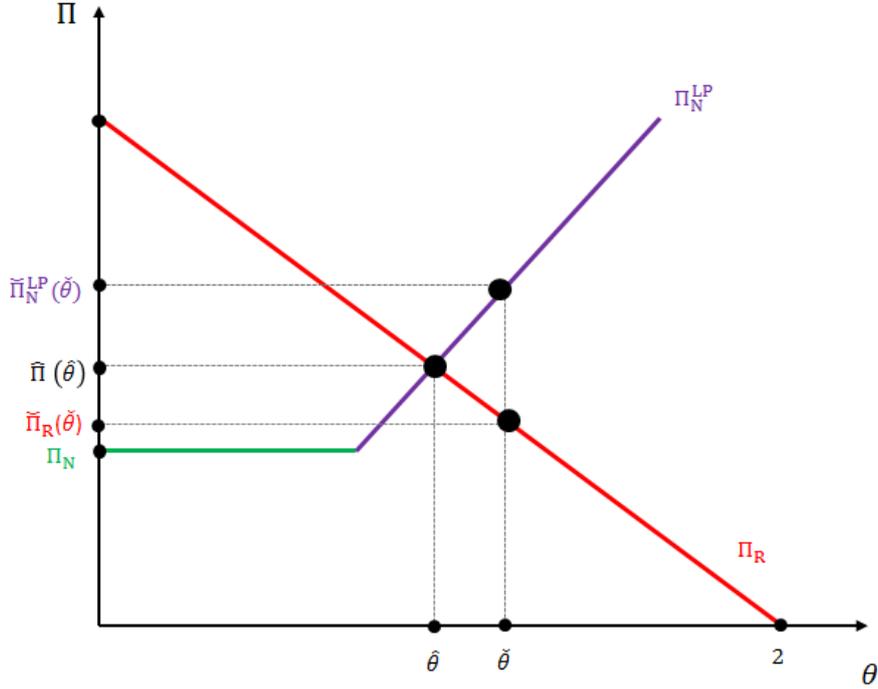
Figure 12: When $\check{\theta} > \hat{\theta}$, the leniency programme results in under-deterrence. With a fine discount rate that disregards firms' efforts to avoid detection $\check{\theta}$, cartels with profits $\Pi \in \left(\check{\Pi}_R(\check{\theta}), \hat{\Pi}_R(\hat{\theta}) \right)$ can be sustained by resorting to 'collude and report' strategies.

θ). Therefore, cartels with profits $\Pi \in \left(\check{\Pi}_N^{LP}(\check{\theta}), \hat{\Pi}(\hat{\theta}) \right)$ which otherwise would have been deterred can now be sustained. This is illustrated in Figure 13 below.

The above discussion is condensed in the following proposition.

Proposition 2. *The optimal fine discount rate is suboptimal when the antitrust authority does not take into consideration firms' decisions to engage in avoidance activities. The leniency programme results in under-deterrence. Normal collusion is facilitated when the fine is relatively low, i.e., $F < \frac{2\sqrt{2}k}{p}$. Firms exploit the generosity of the leniency programme and establish cartels that otherwise would have been unprofitable when the fine is relatively high, i.e., $F > \frac{2\sqrt{2}k}{p}$.*

4.5 Conclusions

This paper derives the optimal leniency programme when colluding firms have the opportunity to engage in activities to avoid detection by an antitrust authority.

136
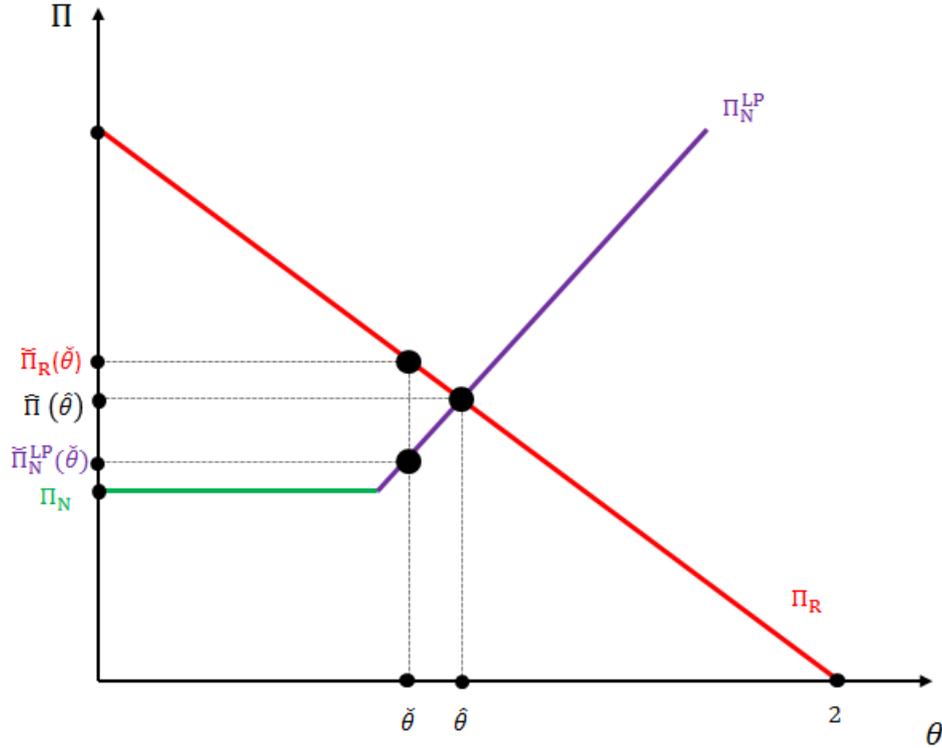
Figure 13: When $\check{\theta} < \hat{\theta}$, the leniency programme results in under-deterrence. With a fine discount rate that disregards firms' efforts to avoid detection $\check{\theta}$, cartels with profits $\Pi \in \left(\check{\Pi}_N^{LP}(\check{\theta}), \hat{\Pi}(\hat{\theta}) \right)$ can be sustained since a deviation in the market with a simultaneous self-report to the antitrust authority becomes less profitable.

We have modified the basic model by Chen and Rey [2012] in order to incorporate an endogenous probability of detection which depends on the frequency of inspection and the intensity of firms' detection avoidance activities. This model provides a very simple theoretical framework to analyze the optimal leniency programme. Importantly, it helps to illuminate the trade-off between deterrence and desistance. By extending the basic model developed by Chen and Rey [2012], we show that firms' avoidance efforts have a profound impact on the design of optimal leniency policies. In particular, we show that the optimal fine discount rate depends positively on the severity of the fine and negatively on the probability of investigation and the cost of avoidance activities. Moreover, we demonstrate that when the fine is relatively low and the cost of avoidance effort relatively high no leniency should be offered. This differs from the results previously obtained by Chen and Rey [2012] who show that it is always optimal to offer *some* leniency to

self-reporting firms. Furthermore, the paper derives the conditions under which the antitrust authority must offer a positive reward in order to attain the optimal level of deterrence. Finally, we show that a leniency programme that fails to take into account firms' decisions to undertake effort to avoid detection may result in under-deterrence. Depending on the severity of the fine, either *normal collusion* is facilitated or cartel formation is fostered.

This paper contributes to the literature on leniency programmes in antitrust and, more generally, to the literature on generic law enforcement for crimes committed by groups of violators. By characterizing the leniency programme when avoidance is taken into consideration, the paper derives interesting policy implications about the interplay between the policy instruments (fine, frequency of inspections and leniency programme). Furthermore, it highlights the potential risks that emanate from a wrongly designed leniency programme when policymakers fail to take into consideration firms' efforts to avoid detection. Thus, a policy recommendation that derives from the analysis of this paper is that policy designers should not underrate firms' efforts to avoid detection.

Some remarks are in order before we conclude. The repeated structure of the game essentially means that the antitrust authority does not distinguish between first time offenders and repeated offenders. In their paper, Chen and Rey [2012] extend the basic setting in order to examine the effects of a discriminatory leniency programme. Nonetheless, they do not find support for prohibiting access to the leniency for repeated offenders. Thus, we do not expect our results to change, at least qualitatively, by allowing a discriminatory leniency policy. A further remark is that in this paper we consider avoidance effort which influences the probability of cartel detection. However, the model could be easily adapted in order to consider effort or activities that influence the level of the fine rather than the probability of detection. Such effort may include defense expenditures or litigation effort (lawyers, consultants, advisors, experts etc.).

Several assumptions of this paper were made to simplify the exposition. For instance, we assume that colluding firms (or cartels) are assumed to be homogeneous with respect to the cost of avoidance effort, the cost of firms' avoidance activities is independent of the cartels' profits (or overcharges), and, the antitrust authority has complete information with regard to the firms' detection avoidance technologies. The implications of relaxing these assumptions remain open to future research.

Bibliography

References

Abreu, D. (1986). Extremal equilibria of oligopolistic supergames. *Journal of Economic Theory*, 39(1):191–225.

Abreu, D. (1988). On the theory of infinitely repeated games with discounting. *Econometrica*, 56(2):383–96.

Athey, S. and Bagwell, K. (2001). Optimal collusion with private information. *RAND Journal of Economics*, 32(3):428–65.

Athey, S., Bagwell, K., and Sanchirico, C. (2004). Collusion and price rigidity. *Review of Economic Studies*, 71(2):317–349.

Aubert, C., Rey, P., and Kovacic, W. (2006). The impact of leniency and whistle-blowing programs on cartels. *International Journal of Industrial Organization*, 24(6):1241–1266.

Aumann, R. and Shapley, L. (1976). Long term competition-a game theoretic analysis. UCLA Economics Working Papers 676, UCLA Department of Economics.

Avramovich, M. (2010). Prosecuting cartels: do high fines always contribute to social welfare? *mimeo*.

Avramovich, M. (2011). Rewarding whistle-blowers: implications to deterrence and to principal-agent contracts.

Bagwell, K. and Staiger, R. (1997). Collusion over the business cycle. *RAND Journal of Economics*, 28(1):82–106.

Bain, J. (1948). Output quotas in imperfect cartels. *The Quarterly Journal of Economics*, 62(4):617–622.

Bain, J. (1956). *Barriers to New Competition*. Cambridge, Mass.: Harvard University Press.

Beccaria, C. (1986). *One Crimes and Punishments and other Writings*. Indianapolis: Hackett Publishing Company.

Becker, G. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76.

Becker, G. (1993). Nobel lecture: The economic way of looking at behavior. *Journal of Political Economy*, 101(3):385–409.

Bentham, J. (1931). *Theory of Legislation*. New York: Harcourt, Brace.

Bernheim, B. and Whinston, M. (1990). Multimarket contact and collusive behavior. *RAND Journal of Economics*, 21(1):1–26.

Besanko, D. and Spulber, D. (1989). Antitrust enforcement under asymmetric information. *Economic Journal*, 99(396):408–425.

Bigoni, M., Fridolfsson, S., Coq, C., and Spagnolo, G. (2012). Fines, leniency and rewards in antitrust: An experiment. *RAND Journal of Economics (forthcoming)*.

Bloch, F. (1996). Sequential formation of coalitions in games with externalities and fixed payoff division. *Games and Economic Behavior*, 14(1):90–123.

Bloch, F. (2002). Coalitions and networks in industrial organization. *Manchester School*, 70(1):36–55.

Bloch, F. (2005). Group formation in economics: Networks, clubs and coalitions. chapter Group and Network Formation in Industrial Organization: A Survey, pages 335–353. Cambridge University Press.

Block, M., F., and Sidak, J. (1981). The deterrent effect of antitrust enforcement. *Journal of Political Economy*, 89(3):429–45.

Bolotova, Y. (2009). Cartel overcharges: An empirical analysis. *Journal of Economic Behavior & Organization*, 70(1-2):321–341.

Bolotova, Y., Connor, J., and Miller, D. (2007). Factors influencing the magnitude of cartel overcharges: An empirical analysis of food-industry cartels. *Agribusiness*, 23(1):17–33.

Bos, I. and Pot, E. (2012). On the possibility of welfare-enhancing hard core cartels. *Journal of Economics*, 107(3):199–216.

Boyer, M. and Kotchoni, R. (2011). The econometrics of cartel overcharges. *CIRANO - Scientific Publication No. 2011s-35*.

Brenner, S. (2009). An empirical study of the european corporate leniency program. *International Journal of Industrial Organization*, 27(6):639–645.

Brock, W. and Scheinkman, J. (1985). Price setting supergames with capacity constraints. *Review of Economic Studies*, 52(3):371–82.

Bryant, P. and Eckard, E. (1991). Price fixing: The probability of getting caught. *The Review of Economics and Statistics*, 73(3):531–36.

Buccirossi, P. and Spagnolo, G. (2006a). Leniency policies and illegal transactions. *Journal of Public Economics*, 90(6-7):1281–1297.

Buccirossi, P. and Spagnolo, G. (2006b). Optimal fines in the era of whistleblowers. CEPR Discussion Papers 5465, C.E.P.R. Discussion Papers.

Buccirossi, P. and Spagnolo, G. (2007). The political economy of antitrust (contributions to economic analysis, volume 282). chapter Optimal Fines in the Era of Whistle-blowers - Should Price Fixers Still Go to Prison?, pages 81–122. Emerald Group Publishing Limited.

Chamberlin, E. (1929). Duopoly: Value where sellers are few. *The Quarterly Journal of Economics*, 44(1):63–100.

Chamberlin, E. (1933). *The Theory of Monopolistic Competition*. Cambridge, MA: Harvard University Press.

Chang, M.-H. (1991). The effects of product differentiation on collusive pricing. *International Journal of Industrial Organization*, 9(3):453–469.

Chang, M.-H. and Harrington, J. (2010). The impact of a corporate leniency program on antitrust enforcement and cartelization. Economics Working Paper Archive 548, The Johns Hopkins University, Department of Economics.

Chen, J. and Harrington, J. (2007). The impact of the corporate leniency program on cartel formation and the cartel price path. *The Political Economy of Antitrust*, pages 59–80.

Chen, Z. and Rey, P. (2012). On the design of leniency programs. Technical report, IDEI Working Paper n. 452.

Cho, I.-K. and Kreps, D. (1987). Signaling games and stable equilibria. *The Quarterly Journal of Economics*, 102(2):179–221.

Clay, K. and Troesken, W. (2002). Strategic behavior in whiskey distilling, 1887-1895. *The Journal of Economic History*, 62(4):999–1023.

Cloutier, M. (2011). An empirical investigation of the u.s. corporate leniency programme. Technical report, Queens University.

Cohen, M. and Scheffman, D. (1989). The antitrust sentencing guideline: Is punishment worth the costs? *Journal of Criminal Law*, 27:330–366.

Combe, E. and Monnier, C. (2011). Fines against hard-core cartels in europe: the myth of over-enforcement. *Antitrust Bulletin*, 56(2).

Combe, E., Monnier, C., and Legal, R. (2008). Cartels: The probability of getting caught in the european union. *BEER paper n. 12*.

Compte, O. (1998). Communication in repeated games with imperfect private monitoring. *Econometrica*, 66(3):597–626.

Compte, O., Jenny, F., and Rey, P. (2002). Capacity constraints, mergers and collusion. *European Economic Review*, 46(1):1–29.

Connor, J. (2001). "our customers are our enemies": The lysine cartel of 1992-1995. *Review of Industrial Organization*, 18(1):5–21.

Connor, J. (2003). Private international cartels: Effectiveness, welfare, and anti-cartel enforcement. *Purdue Agricultural Economics Working Paper No. 03-12*.

Connor, J. (2005). Price-fixing overcharges: Legal and economic evidence. Technical report.

Connor, J. and Bolotova, Y. (2006). Cartel overcharges: Survey and meta-analysis. *International Journal of Industrial Organization*, 24(6):1109–1137.

Connor, J. and Lande, R. (2008). Cartel overcharges and optimal cartel fines. *SSRN eLibrary*.

Cowling, K. and Mueller, D. (1978). The social costs of monopoly power. *Economic Journal*, 88(352):727–48.

Cramton, P. and Palfrey, T. (1990). Cartel enforcement with uncertainty about costs. *International Economic Review*, 31(1):17–47.

Crawford, V. and Sobel, J. (1982). Strategic information transmission. *Econometrica*, 50(6):1431–51.

Craycraft, C., Craycraft, J., and Gallo, J. (1997). Antitrust sanctions and a firm's ability to pay. *Review of Industrial Organization*, 12:171–183.

Cyrenne, P. (1999). On antitrust enforcement and the deterrence of collusive behaviour. *Review of Industrial Organization*, 14(3):257–272.

d'Aspremont, C., Jacquemin, A., Gabszewicz, J., and Weymark, J. (1983). On the stability of collusive price leadership. *Canadian Journal of Economics*, 16(1):17–25.

Davidson, C. and Deneckere, R. (1990). Excess capacity and collusion. *International Economic Review*, 31(3):521–41.

Deltas, G., Salvo, A., and Vasconcelos, H. (2009). Welfare-enhancing collusion and coupons. *Discussion Paper*.

Deneckere, R. (1983). Duopoly supergames with product differentiation. *Economics Letters*, 11(1-2):37–42.

Dijkstra, P., Haan, M., and Schoonbeek, L. (2011). Leniency programs and the design of antitrust: Experimental evidence with unrestricted communication. *mimeo*.

Donsimoni, M.-P. (1985). Stable heterogeneous cartels. *International Journal of Industrial Organization*, 3(4):451–467.

Donsimoni, M.-P., Economides, N., and Polemarchakis, H. (1986). Stable cartels. *International Economic Review*, 27(2):317–27.

Doyle, C. and Martijn, H. (2005). Efficient cartelization through buyer groups. Technical Report 2009-03, Amsterdam Center for Law & Economics.

Ellis, C. and Wislon, W. (2003). Cartels, price fixing, and corporate leniency policy: what doesn't kill is make us stronger. *mimeo*.

Escrihuela-Villar, M. (2004). Cartel sustainability and cartel stability. Working Paper 44, Fondazione Eni Enrico Mattei.

Farrell, J. (1987). Cheap talk, coordination, and entry. *RAND Journal of Economics*, 18(1):34–39.

Farrell, J. and Maskin, E. (1989). Renegotiation in repeated games. *Games and Economic Behavior*, 1(4):327–360.

Farrell, J. and Rabin, M. (1996). Cheap talk. *The Journal of Economic Perspectives*, 10(3):103–118.

Feess, E. and Heesen, E. (2002). Self-reporting and ex post asymmetric information. *Journal of Economics*, 2:141–153.

Feess, E. and Markus, M. (2005). Quid-pro-quo or winner-takes-it-all? an analysis of corporate leniency programs and lessons to learn for eu and us policies. Technical report.

Feess, E. and Walzl, M. (2004). Self-reporting in optimal law enforcement when there are criminal teams. *Economica*, 71(283):333–348.

Feizi, M. (2010). Dynamic (in)consistent antitrist enforcement and cartel infringement: a differential game approach. Goethe University Frankfurt. mimeo.

Fershtman, C. and Pakes, A. (2000). A dynamic oligopoly with collusion and price wars. *RAND Journal of Economics*, 31(2):207–236.

Feuerstein, S. (2005). Collusion in industrial economics: A survey. *Journal of Industry, Competition and Trade*, 5(3):163–198.

Friedman, J. W. (1971). A non-cooperative equilibrium for supergames. *Review of Economic Studies*, 38(113):1–12.

Froeb, L., Koyak, R., and Werden, G. (1993). What is the effect of bid-rigging on prices? *Economics Letters*, 42(4):419–423.

Fudenberg, D. and Maskin, E. (1986). The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, 54(3):533–54.

Fudenberg, D. and Tirole, J. (1985). Preemption and rent equilization in the adoption of new technology. *Review of Economic Studies*, 52(3):383–401.

Fudenberg, D. and Tirole, J. (1991a). *Game Theory*. The MIT Press.

Fudenberg, D. and Tirole, J. (1991b). Perfect bayesian equilibrium and sequential equilibrium. *Journal of Economic Theory*, 53(2):236–260.

Gallo, J., Dau-Schmidt, K., Craycraft, J., and Parker, C. (2000). Department of justice antitrust enforcement, 1955-1997: An empirical study. *Review of Industrial Organization*, 17(75-133).

Garoupa, N. (1997). The theory of optimal law enforcement. *Journal of Economic Surveys*, 11(3):267–95.

Green, E. and Porter, R. (1984). Noncooperative collusion under imperfect price information. *Econometrica*, 52(1):87–100.

Griffin, J. (1989). *Economics in Theory and Practice: An Eclectic Approach*, chapter Previous Cartel Experience: Any Lessons for OPEC?, pages 179–206. Kluwer Academic Publishers.

Häckner, J. (1996). Optimal symmetric punishments in a bertrand differentiated products duopoly. *International Journal of Industrial Organization*, 14(5):611–630.

Haltiwanger, J. and Harrington, J. (1991). The impact of cyclical demand movements on collusive behavior. *RAND Journal of Economics*, 22(1):89–106.

144

Harberger, A. C. (1954). Monopoly and resource allocation. *The American Economic Review*, 44(2):77–87.

Harrington, J. (2004). Cartel pricing dynamics in the presence of an antitrust authority. *RAND Journal of Economics*, 35(4):651–673.

Harrington, J. (2005). Optimal cartel pricing in the presence of an antitrust authority. *International Economic Review*, 46(1):145–169.

Harrington, J. (2008). Optimal corporate leniency programs. *The Journal of Industrial Economics*, 56(2):215–246.

Harsanyi, J. (1968). Games with incomplete information played by "bayesian" players, i-iii. part iii. the basic probability distribution of the game. *Management Science*, 14(7):486–502.

Harsanyi, J. and Selten, R. (1988). *A General Theory of Equilibrium Selection in Games*. The MIT Press.

Hicks, J. (1935). Annual survey of economic theory: The theory of monopoly. *Econometrica*, 3(1):1–20.

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

Hinloopen, J. and Soetevent, A. (2008). Laboratory evidence on the effectiveness of corporate leniency programs. *RAND Journal of Economics*, 39(2):607–616.

Houba, H., Motchenkova, E., and Wen, Q. (2010). Antitrust enforcement with price-dependent fines and detection probabilities. *Economics Bulletin*, 30(3):2017–2027.

Houba, H., Motchenkova, E., and Wen, Q. (2011). Antitrust enforcement and marginal deterrence. Tinbergen Institute Discussion Papers 11-166/1, Tinbergen Institute.

Hovenkamp, H. and Leslie, C. (2011). The firm as a cartel manager. *Vanderbilt Law Review*, 64:813–873.

Howard, J. and Kaserman, D. (2004). Proof of damages in construction industry bid-rigging cases. *Antitrust Bulletin*, 34:359–393.

Hviid, M. (1992). Endogenous cartel formation with private information. *Canadian Journal of Economics*, 25(4):972–82.

Hviid, M. and Stephan, A. (2009). *Cases in European Competition Policy*, chapter The graphite electrodes cartel: fines which deter?, pages 137–158. Cambridge University Press.

Innes, R. (1999). Remediation and self-reporting in optimal law enforcement. *Journal of Public Economics*, 72(3):379–393.

Innes, R. (2000). Self-reporting in optimal law enforcement when violators have heterogeneous probabilities of apprehension. *Journal of Legal Studies*, 29(1):287–300.

Innes, R. (2001). Violator avoidance activities and self-reporting in optimal law enforcement. *Journal of Law, Economics and Organization*, 17(1):239–56.

Jenny, F. and Weber, A.-P. (1983). Aggregate welfare loss due to monopoly power in the french economy: Some tentative estimates. *Journal of Industrial Economics*, 32(2):113–30.

Jensen, S., Kvaløy, O., Olsen, T., and Sørgard, L. (2012). When tougher punishment leads to a more severe crime: an application to antitrust enforcement. *mimeo*.

Kamerschen, D. (1966). An estimation of the welfare losses from monopoly in the american economy. *Western Economic Journal*, 4:221–236.

Kamien, M. I. and Zang, I. (1990). The limits of monopolization through acquisition. *The Quarterly Journal of Economics*, 105(2):465–499.

Kandori, M. and Matsushima, H. (1998). Private observation, communication and collusion. *Econometrica*, 66(3):627–652.

Kaplow, L. and Shavell, S. (1994). Optimal law enforcement with self-reporting of behavior. *Journal of Political Economy*, 102(3):583–606.

Kihlstrom, R. and Vives, X. (1992). Collusion by asymmetrically informed firms. *Journal of Economics & Management Strategy*, 1(2):371–96.

Klein, G. (2010). Cartel destabilization and leniency programs: Empirical evidence. Technical report.

Kreps, D. and Wilson, R. (1982). Sequential equilibria. *Econometrica*, 50(4):863–94.

Kühn, K.-U. (2001). Fighting collusion by regulating communication between firms. *Economic Policy*, 16(32):167–204.

Kwoka, J. (1999). The price eﬀects of bidding conspiracies: evidence from real estate auctions. *The Antitrust Bulletin*, 42:503–516.

Labini, P. (1969). *Barriers to New Competition*. Cambridge, Mass.: Harvard University Press.

Lambson, V. (1987). Optimal penal codes in price-setting supergames with capacity constraints. *Review of Economic Studies*, 54(3):385–97.

Lambson, V. (1994). Some results on optimal penal codes in asymmetric bertrand supergames. *Journal of Economic Theory*, 62(2):444–468.

Lambson, V. (1995). Optimal penal codes in nearly symmetric bertrand supergames with capacity constraints. *Journal of Mathematical Economics*, 24(1):1–22.

Landes, W. M. (1983). Optimal sanctions for antitrust violations. *The University of Chicago Law Review*, 50(2):pp. 652–678.

Langlais, E. (2008). Detection avoidance and deterrence: Some paradoxical arithmetic. *Journal of Public Economic Theory*, 10(3):371–382.

Leibenstein, H. (1966). Allocative efficiency vs. "x-efficiency". *The American Economic Review*, 56(3):392–415.

Levenstein, M. (1995). Mass production conquers the pool: Firm organization and the nature of competition in the nineteenth century. *The Journal of Economic History*, 55(3):575–611.

Levenstein, M. and Suslow, V. (2006). What determines cartel success? *Journal of Economic Literature*, 44(1):43–95.

Malik, A. (1990). Avoidance, screening and optimum enforcement. *RAND Journal of Economics*, 21(3):341–353.

Malik, A. (1993). Self-reporting and the design of policies for regulating stochastic pollution. *Journal of Environmental Economics and Management*, 24(3):241–257.

Marceau, N. and Mongrain, S. (2000). Amnesties and co-operation. *International Tax and Public Finance*, 7:259–273.

Maskin, E. and Tirole, J. (2001). Markov perfect equilibrium. *Journal of Economic Theory*, 100:191–219.

Miklós-Thal, J. (2009). Optimal collusion under cost asymmetry. *Economic Theory*, 46(1):99–125.

Miller, D. and Connor, J. (2010). The predictability of global cartel fines. *Concurrences: Review of Competition Law, Vol. 2, 2010*.

Miller, N. (2009). Strategic leniency and cartel enforcement. *American Economic Review*, 99(3):750–68.

Modigliani, F. (1958). New developments on the oligopoly front. *Journal of Political Economy*, 66:215.

Motchenkova, E. (2004). Effects of leniency programs on cartel stability. *Tilburg University, Tilburg Law and Economic Center*.

Motchenkova, E. and Kort, P. (2006). Analysis of the properties of the current penalty schemes for violations of antitrust law. *Journal of Optimization Theory and Applications*, 182(2):431–451.

Motchenkova, E. and Laan, R. (2011). Strictness of leniency programs and asymmetric punishment effect. *International Review of Economics*, 58(4):401–431.

Motchenkova, E. and Leliefeld, D. (2010). Adverse effects of corporate leniency programs in view of industry asymmetry. *Journal of Applied Economic Sciences*, 5(2(12)/Sum):114–128.

Motta, M. (2004). *Competition Policy*. Cambridge Books. Cambridge University Press.

Motta, M. and Polo, M. (1999). Leniency programs and cartel prosecution. Open access publications from european university institute, European University Institute.

Motta, M. and Polo, M. (2003). Leniency programs and cartel prosecution. *International Journal of Industrial Organization*, 21(3):347–379.

Neven, D. (1989). Strategic entry deterrence: Recent developments in the economics of industry. *Journal of Economic Surveys*, 3(3):213–33.

Nussim, J. and Tabbach, A. (2009). Deterrence and avoidance. *International Review of Law and Economics*, 29(4):314–323.

Odudu, O. (2006). *The Boundaries of EC Competition Law The Scope of Article 81*. Oxford University Press.

Ormosi, P. (2011). How big is a tip of the iceberg? A parsimonious way to estimate cartel detection rate. *SSRN eLibrary*.

Pénard, T. and Souam, S. (1996). Optimal antitrust policy under asymmetric information and budget constraints. Technical report, University of Paris I, Panthéon-Sorbonne.

Phlips, L. (1995). *Competition Policy: A Game Theoretic Perspective*. Cambridge University Press.

Polinsky, M. and Shavell, S. (1979). The optimal tradeoff between the probability and magnitude of fines. *American Economic Review*, 69(5):880–91.

Polinsky, M. and Shavell, S. (2000). The economic theory of public enforcement of law. *Journal of Economic Literature*, 38(1):45–76.

Porter, R. and Zona, J. (1999). Ohio school milk markets: An analysis of bidding. *RAND Journal of Economics*, 30(2):263–288.

Porter, R. H. (1983). Optimal cartel trigger price strategies. *Journal of Economic Theory*, 29(2):313–338.

Posner, R. (1970). A statistical study of antitrust enforcement. *Journal of Law and Economics*, 13(2):365–419.

Posner, R. (2001). *Antitrust Law*. The University of Chicago Press, $2^{nd} edition$.

Raith, M. (1996). Product differentiation, uncertainty and the stability of collusion. STICERD - Economics of Industry Papers 16, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE.

Reinganum, J. (1981). On the diffusion of new technology: A game theoretic approach. *Review of Economic Studies*, 48(3):395–405.

Roberts, K. (1985). Cartel behaviour and adverse selection. *Journal of Industrial Economics*, 33(4):401–13.

Ross, T. (1992). Cartel stability and product differentiation. *International Journal of Industrial Organization*, 10(1):1–13.

Rotemberg, J. and Saloner, G. (1986). A supergame-theoretic model of price wars during booms. *The American Economic Review*, 76(3):390–407.

Rotemberg, J. and Saloner, G. (1990). Collusive price leadership. *The Journal of Industrial Economics*, 39(1):93–111.

Rothschild, R. (1999). Cartel stability when costs are heterogeneous. *International Journal of Industrial Organization*, 17(5):717–734.

Rubinstein, A. (1979). Equilibrium in supergames with the overtaking criterion. *Journal of Economic Theory*, 21(1):1–9.

Salant, S., Switzer, S., and Reynolds, R. (1983). Losses from horizontal merger: The effects of an exogenous change in industry structure on cournot-nash equilibrium. *The Quarterly Journal of Economics*, 98(2):185–199.

Sanchirico, C. (2011). Procedural law and economics. Encyclopedia of Law and Economics, chapter Detection avoidance and enforcement theory. Edward Elgar Publishing, second edition edition.

Sauvagnat, J. (2010). Prosecution and leniency programs: a fool's game. TSE Working Papers 10-188, Toulouse School of Economics (TSE).

Schinkel, M. and Tuinstra, J. (2006). Imperfect competition law enforcement. *International Journal of Industrial Organization*, 24(6):1267–1297.

Schwartzman, D. (1960). The burden of monopoly. *The Journal of Political Economy*, 68(6):627–630.

Scott, J. (2008). Multimarket contacts. *Issues in Competition Law and Policy, ABA Section of Antitrust Law, Vol. 2, p. 1553, 2008.*

Selten, R. (1973). A simple model of imperfect comptition where 4 are few and 6 are many. *International Journal of Game Theory*, 2(3):141–201.

Shaffer, S. (1995). Stable cartels with a cournot fringe. *Southern Economic journal*, 61(3):744–754.

Shapiro, C. (1989). Theories of oligopoly behavior. In Schmalensee, R. and Willig, R., editors, *Handbook of Industrial Organization*, volume 1 of *Handbook of Industrial Organization*, chapter 6, pages 329–414. Elsevier.

Silbye, F. (2010). Asymmetric evidence and optimal leniency programs. In *Topics in competitition policy: cartels, leniency and price discrimination*, pages 41–67. University of Copenhagen.

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. Clarendon Press, Oxford, Edition 1976.

Snyder, C. (1996). A dynamic theory of countervailing power. *RAND Journal of Economics*, 27(4):747–769.

Spagnolo, G. (1999). On interdependent supergames: Multimarket contact, concavity, and collusion. *Journal of Economic Theory*, 89(1):127–139.

Spagnolo, G. (2000). Optimal leniency programs. *F.E.E.M. Nota di Lavoro No. 42.00.*

Spagnolo, G. (2004). Divide et impera: Optimal deterrence mechanisms against cartels and organized crime. Econometric Society 2004 North American Winter Meetings, Econometric Society.

Spagnolo, G. (2008). Handbook of antitrust economics. chapter Leniency and Whistleblowers in Antitrust, pages 259–304. MIT Press.

Staiger, R. and Wolak, F. (1992). Collusive pricing with capacity constraints in the presence of demand uncertainty. *RAND Journal of Economics*, 23(2):203–220.

Stephan, A. (2010). Hear no evil, see no evil: Why antitrust compliance programmes may be ineffective at preventing cartels. *The Company Lawyer*, 31(4):231–239.

Stigler, G. (1950). Monopoly and oligopoly by merger. *The American Economic Review*, 40(2):23–34.

Stigler, G. (1964). A theory of oligopoly. *Journal of Political Economy*, 72:44–61.

Stigler, G. (1970). The optimum enforcement of laws. *Journal of Political Economy*, 78(3):526–36.

Stigler, G. (1983). *The Organization of Industry.* Chicago: University of Chicago Press.

Symeonidis, G. (2002). Cartel stability with multiproduct firms. *International Journal of Industrial Organization*, 20(3):339–352.

Tirole, J. (1988). *The Theory of Industrial Organization.* MIT Press, Cambridge, Massachusetts.

Turner, D. (1962). The definition of agreement under the sherman act: Conscious parallelism and refusals to deal. *Harvard Law Review*, 75(4):655–706.

van Damme, E. (1989). Renegotiation-proof equilibria in repeated prisoners' dilemma. *Journal of Economic Theory*, 47(1):206–217.

Vaska, M. (1985). Conscious parallelism and price fixing: Defining the boundary. *The University of Chicago Law Review*, 52(2):508–535.

Verboven, F. (1998). Localized competition, multimarket operation, and collusive behavior. *International Economic Review*, 39(2):371–98.

Vickers, J. (1995). Concepts of competition. *Oxford Economic Papers*, 47(1):1–23.

Vives, X. (2001). *Oligopoly Pricing: Old Ideas and New Tools*. The MIT Press.

Werdem, G. and Simon, M. (1987). Why price fixers should go to prison. *Antitrust Bulletin*, 32(4):917–937.

Werden, G. J. (2003). The effect of antitrust policy on consumer welfare: What crandall and winston overlook. *JU.S. Department of Justice Antitrust Division Discussion Paper No. EAG 03-2*.

Williamson, O. (1968). Economies as an antitrust defense: The welfare tradeoffs. *The American Economic Review*, 58(1):18–36.

Wils, W. (2008). *Efficiency and Justice in European Antitrust Enforcement*. Hart Publishing.

Wilson, R. (1992). Strategic models of entry deterrence. In Aumann, R. and Hart, S., editors, *Handbook of Game Theory with Economic Applications*, volume 1 of *Handbook of Game Theory with Economic Applications*, chapter 10, pages 305–329. Elsevier.

Worcester, D. (1973). New estimates of the welfare loss to monopoly, united states, 1956-1969. *Southern Economic Journal*, 40:234–245.

Yi, S.-S. (1997). Stable coalition structures with externalities. *Games and Economic Behavior*, 20(2):201–237.

Zhou, J. (2011). Evaluating leniency and modeling cartel durations: time-varying policy impacts and sample selection. *SSRN eLibrary*.

Zhou, J. (2012). Evaluating leniency with missing information on undetected cartels: Exploring time-varying policy impacts on cartel duration. Technical report, Bonn University.

Appendix A - Proof of Proposition 1

Lemma 6.   *1. $p_1 \geq 0$ iff $\Delta \leq \Delta_1$, where $\Delta_1 \equiv \frac{\pi}{F}\frac{2(\delta+\beta-\beta\delta)}{(\delta+\beta-\beta\delta+1)}$*

2. $\Delta_1 > 1$ iff $\frac{\pi}{F} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$

3. $p_1 + \Delta \leq 1$ iff $\Delta \geq \Delta_2$, where $\Delta_2 \equiv \frac{2(\beta+\delta-\beta\delta)}{(1-\beta)(1-\delta)}\left(\frac{\pi}{F}-1\right)$

4. $\Delta_2 > 1$ iff $\frac{\pi}{F} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$

5. $\Delta_1 \geq \Delta_2$ iff $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$.

<u>Proof of Lemma 6</u>
(1) To show: $p_1 > 0$ iff $\Delta < \Delta_1$, where $\Delta_1 \equiv \frac{\pi}{F}\frac{2(\delta+\beta-\beta\delta)}{(\delta+\beta-\beta\delta+1)}$.

$$
\begin{aligned}
p_1 &\geq 0 \Leftrightarrow \\
\frac{\pi}{F} - \frac{(\delta+\beta-\beta\delta+1)\Delta}{2(\delta+\beta-\beta\delta)} &\geq 0 \Leftrightarrow \\
\Delta &\leq \frac{\pi}{F}\frac{2(\delta+\beta-\beta\delta)}{(\delta+\beta-\beta\delta+1)} \equiv \Delta_1
\end{aligned}
$$

$\therefore QED.$
(2) To show: $\Delta_1 > 1$ iff $\frac{\pi}{F} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$.

$$
\begin{aligned}
\Delta_1 &> 1 \Leftrightarrow \\
\frac{\pi}{F}\frac{2(\delta+\beta-\beta\delta)}{(\delta+\beta-\beta\delta+1)} &> 1 \Leftrightarrow \\
\frac{\pi}{F} &> \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}
\end{aligned}
$$

$\therefore QED.$
(3) To show: $p_1 + \Delta \leq 1$ iff $\Delta > \Delta_2$, where $\Delta_2 \equiv \frac{2(\beta+\delta-\beta\delta)}{(1-\beta)(1-\delta)}\left(\frac{\pi}{F}-1\right)$.

$$\frac{\pi}{F} - \frac{(\delta + \beta - \beta\delta + 1)\Delta}{2(\delta + \beta - \beta\delta)} + \Delta \leq 1 \Leftrightarrow$$

$$(\frac{\pi}{F} - 1) \leq \Delta(\frac{\delta + \beta - \beta\delta + 1}{2(\delta + \beta - \beta\delta)} - 1) \Leftrightarrow$$

$$\Delta \geq \frac{2(\beta + \delta - \beta\delta)}{(1 - \beta)(1 - \delta)}(\frac{\pi}{F} - 1) \equiv \Delta_2$$

$\therefore QED.$

(4) To show: $\Delta_2 > 1$ iff $\frac{\pi}{F} > \frac{\delta + \beta - \beta\delta + 1}{2(\delta + \beta - \beta\delta)}$.

$$\Delta_2 > 1 \Leftrightarrow$$

$$\frac{2(\beta + \delta - \beta\delta)}{(1 - \beta)(1 - \delta)}(\frac{\pi}{F} - 1) > 1 \Leftrightarrow$$

$$\frac{\pi}{F} - 1 > \frac{(1 - \beta)(1 - \delta)}{2(\beta + \delta - \beta\delta)} \Leftrightarrow$$

$$\frac{\pi}{F} > \frac{(1 - \beta)(1 - \delta)}{2(\beta + \delta - \beta\delta)} + 1 \Leftrightarrow$$

$$\frac{\pi}{F} > \frac{\delta + \beta - \beta\delta + 1}{2(\delta + \beta - \beta\delta)}$$

$\therefore QED.$

(5) To show: $\Delta_1 > \Delta_2$ iff $\frac{\pi}{F} \in [1, \frac{\delta + \beta - \beta\delta + 1}{2(\delta + \beta - \beta\delta)}]$.

$$\Delta_1 > \Delta_2 \Leftrightarrow$$

$$\frac{\pi}{F}\frac{2(\delta + \beta - \beta\delta)}{(\delta + \beta - \beta\delta + 1)} > \frac{2(\beta + \delta - \beta\delta)}{(1 - \beta)(1 - \delta)}(\frac{\pi}{F} - 1) \Leftrightarrow$$

$$\frac{\pi}{F} < \frac{1 + \delta + \beta - \beta\delta}{2(\delta + \beta - \beta\delta)}$$

$$\text{true given that } \frac{\pi}{F} \in [1, \frac{\delta + \beta - \beta\delta + 1}{2(\delta + \beta - \beta\delta)}]$$

$\therefore QED.$

If $\frac{\pi}{F} \in [1, \frac{\delta + \beta - \beta\delta + 1}{2(\delta + \beta - \beta\delta)}]$ then both $\Delta_1 \leq 1$ (from Lemma 6(2)) and $\Delta_2 \leq 1$ (from Lemma 6(4)). Moreover, $\Delta_1 > \Delta_2$ (from Lemma 6(5)). Hence, if $\Delta < \Delta_2$ then $p_1 > 1 - \Delta$ so that for all $p \in [0, 1 - \Delta]$ (KEEP, KEEP) is a NE. If $\Delta \in [\Delta_2, \Delta_1]$

then $p \in [0, 1 - \Delta]$ and thus for all $p < p_1$ (KEEP, KEEP) is a NE. Moreover, if $\Delta > \Delta_1$ then $p_1 < 0$ so that (KEEP, KEEP) is not a NE. The intuition for the latter result is the following. If the additional probability with which the collusive firms are detected in the presence of hard evidence is sufficiently high then each firm has a unilateral incentive to deviate from the agreement and destroy the hard evidence. By destroying the hard evidence the probability of detection and therefore of paying the fine decreases by $\frac{\Delta}{2}$. This decrease is higher when Δ is high. If on the other hand $\frac{\pi}{F} > \frac{\delta + \beta - \beta\delta + 1}{2(\delta + \beta - \beta\delta)}$ then both $\Delta_1 > 1$ (from Lemma 6(2)) and $\Delta_2 > 1$ (from Lemma 6(4)), implying that $p_1 > 0$ and $p_1 > 1 - \Delta$. In this case (KEEP, KEEP) is a NE of the *Keep-Destroy* subgame for all $p \in [0, 1 - \Delta]$. Hence, if the ratio of collusive profits to fine is relatively high then no firm has an incentive to deviate from (KEEP, KEEP). Any deviation in that case will result in deprivation of future profits which is very costly given that the level of the fine is relatively low compared to the collusive profits.

Lemma 7. *(KEEP, KEEP) is a NE of the Keep-Destroy game if:*

1. $\frac{\pi}{F} \in [1, \frac{\delta + \beta - \beta\delta + 1}{2(\delta + \beta - \beta\delta)}]$ *and either a)* $\Delta < \Delta_2$ *and* $p \in [0, 1 - \Delta]$ *or b)* $\Delta \in [\Delta_2, \Delta_1]$ *and* $p \in [0, p_1]$ *or*

2. $\frac{\pi}{F} > \frac{\delta + \beta - \beta\delta + 1}{2(\delta + \beta - \beta\delta)}$

Lemma 8.   *1.* $p_3 \geq 0$ *iff* $\Delta < \Delta_3$, *where* $\Delta_3 = \beta\frac{\pi}{F}$

2. $\Delta_3 > 1$ *iff* $\frac{\pi}{F} > \frac{1}{\beta}$

3. $p_3 \leq 1 - \Delta$ *iff* $\Delta \geq \Delta_4$, *where* $\Delta_4 = \frac{\beta}{1-\beta}(\frac{\pi}{F} - 1)$

4. $\Delta_4 > 1$ *iff* $\frac{\pi}{F} > \frac{1}{\beta}$

5. $\Delta_2 > \Delta_4$

6. $\Delta_3 > \Delta_2$

7. $\Delta_1 > \Delta_3$

8. $\Delta_1 > \Delta_3 > \Delta_2 > \Delta_4$

9. $p_1 > p_3$

10. $\frac{\delta + \beta - \beta\delta + 1}{2(\beta + \delta - \beta\delta)} < \frac{1}{\beta}$

11. $\Delta_3 > \Delta_4$ *iff* $\frac{\pi}{F} < \frac{1}{\beta}$.

Proof of Lemma 8

(**1**) To show: $p_3 \geq 0$ iff $\Delta \leq \Delta_3$, where $\Delta_3 \equiv \beta \frac{\pi}{F}$.

$$
\begin{aligned}
p_3 &\geq 0 \\
\frac{\pi}{F} - \frac{\Delta}{\beta} &\geq 0 \Leftrightarrow \\
\beta \frac{\pi}{F} &\geq \Delta \Leftrightarrow \\
\Delta &\leq \beta \frac{\pi}{F} \equiv \Delta_3
\end{aligned}
$$

$\therefore QED.$

(**2**) To show: $\Delta_3 > 1$ iff $\frac{\pi}{F} > \frac{1}{\beta}$.

$$
\begin{aligned}
\beta \frac{\pi}{F} &> 1 \Leftrightarrow \\
\frac{\pi}{F} &> \frac{1}{\beta}
\end{aligned}
$$

$\therefore QED.$

(**4**) To show: $p_3 \leq 1 - \Delta$ iff $\Delta \geq \Delta_4$, where $\Delta_4 = \frac{\beta}{1-\beta}(\frac{\pi}{F} - 1)$.

$$
\begin{aligned}
p_3 &\leq 1 - \Delta \Leftrightarrow \\
\frac{\pi}{F} - \frac{\Delta}{\beta} &\leq 1 - \Delta \Leftrightarrow \\
\frac{\pi}{F} - 1 &\leq \Delta(\frac{1}{\beta} - 1) \Leftrightarrow \\
\Delta &\geq \frac{\beta}{1-\beta}(\frac{\pi}{F} - 1) \equiv \Delta_4
\end{aligned}
$$

$\therefore QED.$

(**5**) To show: $\Delta_2 > \Delta_4$.

$$\Delta_2 > \Delta_4$$

$$\frac{2(\beta + \delta - \beta\delta)}{1 - (\beta + \delta - \beta\delta)}(\frac{\pi}{F} - 1) > \frac{\beta}{1 - \beta}(\frac{\pi}{F} - 1)$$

$$2(\beta + \delta - \beta\delta)(1 - \beta) > \beta[1 - (\beta + \delta - \beta\delta)] \Leftrightarrow$$

$$(\beta + \delta - \beta\delta)(2 - 2\beta + \beta) > \beta \Leftrightarrow$$

$$(\beta + \delta - \beta\delta)(2 - \beta) > \beta \Leftrightarrow$$

$$2\beta + 2\delta(1 - \beta) - \beta^2 - \beta\delta(1 - \beta) > \beta \Leftrightarrow$$

$$\beta(1 - \beta) + \delta(2 - \beta)(1 - \beta) > 0$$

$$true$$

$\therefore QED.$

(6) To show: $\Delta_3 > \Delta_2$ iff $\frac{\pi}{F}[1, \frac{\delta + \beta - \beta\delta + 1}{2(\delta + \beta - \beta\delta)}]$.
By contradiction. Suppose that:

$$\Delta_2 > \Delta_3$$

$$\frac{2(\beta + \delta - \beta\delta)}{1 - (\beta + \delta - \beta\delta)}(\frac{\pi}{F} - 1) > \beta\frac{\pi}{F}$$

$$2(\beta + \delta - \beta\delta)(\frac{\pi}{F} - 1) > \beta[1 - (\beta + \delta - \beta\delta)]\frac{\pi}{F}$$

$$\frac{\pi}{F}(2(\beta + \delta - \beta\delta) - \beta + \beta(\beta + \delta - \beta\delta) > 2(\beta + \delta - \beta\delta)$$

$$\frac{\pi}{F} > \frac{2(\beta + \delta - \beta\delta)}{\beta^2(1 - \delta) + \beta(1 - \delta) + 2\delta}$$

Notice that $\frac{2(\beta + \delta - \beta\delta)}{\beta^2(1-\delta) + \beta(1-\delta) + 2\delta} > 1$.

$$\frac{2(\beta + \delta - \beta\delta)}{\beta^2(1 - \delta) + \beta(1 - \delta) + 2\delta} > 1$$

$$2(\beta + \delta - \beta\delta) > \beta^2(1 - \delta) + \beta(1 - \delta) + 2\delta$$

$$2\delta + 2\beta(1 - \delta) < \beta^2(1 - \delta) + \beta(1 - \delta) + 2\delta$$

$$\beta(1 - \delta) - \beta^2(1 - \delta) > 0$$

$$\beta(1 - \beta)(1 - \delta) > 0$$

$$\text{always true given that } \beta, \delta < 1$$

But given that $\frac{\pi}{F}[1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$ this leads to a contradiction since:

$$\frac{2(\beta + \delta - \beta\delta)}{\beta^2(1 - \delta) + \beta(1 - \delta) + 2\delta} > \frac{\delta + \beta - \beta\delta + 1}{2(\delta + \beta - \beta\delta)}$$

$$4(\beta + \delta - \beta\delta)(\delta + \beta - \beta\delta) > (\beta^2(1 - \delta) + \beta(1 - \delta) + 2\delta)(\delta + \beta - \beta\delta + 1)$$
$$\text{contradiction}$$

$\therefore QED.$
(**7**)
To show: $\Delta_1 > \Delta_3$.

$$\Delta_1 > \Delta_3$$
$$\frac{\pi}{F}\frac{2(\delta + \beta - \beta\delta)}{(\delta + \beta - \beta\delta + 1)} > \beta\frac{\pi}{F}$$
$$2(\delta + \beta - \beta\delta) > \beta + \beta(\delta + \beta - \beta\delta)$$
$$2\beta + 2\delta(1 - \beta) - \beta - \beta\delta - \beta(1 - \delta) > 0$$
$$2\delta(1 - \beta) > 0$$
$$\beta(1 - \beta) + (1 - \beta)[\delta(2 - \beta)] > 0$$
$$\text{true}$$

$\therefore QED.$
(**8**) To show: $\Delta_1 > \Delta_3 > \Delta_2 > \Delta_4$ iff $\frac{\pi}{F}[1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$.

From Lemmata 5, 6 and 7.

$\therefore QED.$

158

(**9**) To show: $p_1 > p_3$.

$$
\begin{aligned}
p_1 &> p_3 \\
\frac{\pi}{F} - \frac{(\delta + \beta - \beta\delta + 1)\Delta}{2(\delta + \beta - \beta\delta)} &> \frac{\pi}{F} - \frac{\Delta}{\beta} \\
\frac{\Delta}{\beta} &> \frac{(\delta + \beta - \beta\delta + 1)\Delta}{2(\delta + \beta - \beta\delta)} \\
(2 - \beta)(\delta + \beta - \beta\delta) - \beta &> 0 \\
\beta + 2\delta(1 - \beta) - \beta^2 - \beta\delta(1 - \beta) &> 0 \\
\beta(1 - \beta) + \delta(2 - \beta)(1 - \beta) &> 0 \\
\delta &> -\frac{\beta}{2 - \beta} < 0
\end{aligned}
$$

which is always true given that $\delta \in [0, 1]$

$\therefore QED.$

(**10**) To show: $\frac{\delta + \beta - \beta\delta + 1}{2(\beta + \delta - \beta\delta)} < \frac{1}{\beta}$.

$$
\begin{aligned}
\frac{\delta + \beta - \beta\delta + 1}{2(\beta + \delta - \beta\delta)} &< \frac{1}{\beta} \\
\beta(\delta + \beta - \beta\delta) + \beta &< 2(\beta + \delta - \beta\delta) \\
\beta^2 + \beta\delta(1 - \beta) + \beta &< 2\beta + 2\delta(1 - \beta) \\
\beta(1 - \beta) + 2\delta(1 - \beta) - \beta\delta(1 - \beta) &> 0 \\
\beta + \delta(2 - \beta) &> 0 \\
&\text{true}
\end{aligned}
$$

$\therefore QED.$

(**11**) To show: $\Delta_3 > \Delta_4$ iff $\frac{\pi}{F} < \frac{1}{\beta}$.

$$
\begin{aligned}
\Delta_3 &> \Delta_4 \Leftrightarrow \\
\beta\frac{\pi}{F} &> \frac{\beta}{1-\beta}(\frac{\pi}{F}-1) \Leftrightarrow \\
\frac{\pi}{F} &> \frac{1}{1-\beta}(\frac{\pi}{F}-1) \Leftrightarrow \\
(1-\beta)\frac{\pi}{F} &> \frac{\pi}{F}-1 \Leftrightarrow \\
\frac{\pi}{F}(1-\beta-1) &< -1 \Leftrightarrow \\
1 &> \beta\frac{\pi}{F} \Leftrightarrow \\
\frac{\pi}{F} &< \frac{1}{\beta}
\end{aligned}
$$

$\therefore QED.$

From Lemmata **7** and **8** we get the following result.

Lemma 9. *The Keep-Destroy subgame has two NE if:*

1. *For $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$*

    (a) *$\Delta < \Delta_2$ and $p \in [0, 1-\Delta]$*
    (b) *$\Delta \in [\Delta_2, \Delta_1]$ and $p \in [0, p_1]$*

2. *For $\frac{\pi}{F} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$, $\forall\Delta$ and $\forall p$.*

Corollary 2. *For $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$ the unique NE is (*DESTROY, DESTROY*) if:*

1. *$\Delta \in [\Delta_2, \Delta_1]$ and $p \in [p_1, 1-\Delta]$ or*

2. *$\Delta > \Delta_1$.*

From Lemma **9** we distinguish three cases:

1. $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$ and $\Delta < \Delta_2$

2. $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$, $p < p_1$ and $\Delta \in [\Delta_2, \Delta_1]$

3. $\frac{\pi}{F} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$

We will analyze each of the three cases below.

Case 1. $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$ *and* $\Delta < \Delta_2$

Given that $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$, then from Lemma 8(8) $\Delta_3 > \Delta_2 > \Delta_4$. Given that $\Delta_3 > \Delta_2$, from Lemma 8(1) $p_3 \geq 0$. Moreover, if $\Delta < \Delta_4$ then $p_3 > 1 - \Delta$ so that for all p (KEEP, KEEP) Pareto dominates (DESTROY, DESTROY). If on the other hand $\Delta \in [\Delta_4, \Delta_2]$ then from 8(3) $p_3 \leq 1 - \Delta$ and thus for $p < p_3$ (KEEP, KEEP) Pareto dominates (DESTROY, DESTROY), while for $p \geq p_3$ (DESTROY, DESTROY) Pareto dominates (KEEP, KEEP). These results are summarized in the following lemma.

Lemma 10. *For* $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$ *and* $\Delta < \Delta_2$ *the Pareto dominant NE of the Keep-Destroy subgame is:*

1. *(*KEEP, KEEP*) if:*

    (a) $\Delta < \Delta_4 (< \Delta_2)$ *and* $\forall p \in [0, 1 - \Delta]$

    (b) $\Delta \in [\Delta_4, \Delta_2]$ *and* $p < p_3$

2. *(*DESTROY, DESTROY*) if* $\Delta \in [\Delta_4, \Delta_2]$ *and* $p \geq p_3$.

Case 2. $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$, $p \leq p_1$ *and* $\Delta \in [\Delta_2, \Delta_1]$

From Lemma 8(8) we know that $\Delta_1 > \Delta_3 > \Delta_2 > \Delta_4$. Hence, $\Delta_3 \in [\Delta_2, \Delta_1]$. Given that $\Delta_4 < \Delta_2$ for all $\Delta \in [\Delta_2, \Delta_1]$, $p_3 \leq 1 - \Delta$. Then, for $\Delta \in [\Delta_2, \Delta_3]$, $p_3 \in [0, p_1]$, while for $\Delta \in (\Delta_3, \Delta_1]$, $p_3 < 0$. Hence, for $\Delta \in [\Delta_2, \Delta_3]$ if $p < p_3$ (KEEP, KEEP) Pareto dominates (DESTROY, DESTROY), while for $p \in [p_3, p_1]$ (DESTROY, DESTROY) Pareto dominates (KEEP, KEEP). On the other hand, if $\Delta \in (\Delta_3, \Delta_1]$ then for all $p \leq p_1$ (DESTROY, DESTROY) Pareto dominates (KEEP, KEEP).

From the above analysis and from Corollary 2 we get the following result.

Lemma 11. *For* $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$, $p < p_1$ *and* $\Delta \in [\Delta_2, \Delta_1]$ *the Pareto dominant NE of the Keep-Destroy subgame is:*

1. *(*KEEP, KEEP*) if* $\Delta \in [\Delta_2, \Delta_3]$ *and* $p < p_3$

2. *(*DESTROY, DESTROY*) if:*

    (a) $\Delta \in [\Delta_2, \Delta_3]$ *and* $p \in [p_3, p_1]$

(b) $\Delta \in [\Delta_2, \Delta_1]$ *and* $p > p_1$

(c) $\Delta \in (\Delta_3, \Delta_1]$ *and* $p \leq p_1$

(d) $\Delta > \Delta_1$

Conjoining Lemmata 10 and 11, we obtain the following lemma.

Lemma 12. *For* $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$ *the Pareto dominant NE of the Keep-Destroy subgame is:*

1. *(*KEEP, KEEP*) if:*

    (a) $\Delta < \Delta_4$ *and* $\forall p \in [0, 1 - \Delta]$

    (b) $\Delta \in [\Delta_2, \Delta_3]$ *and* $p < p_3$

2. *(*DESTROY, DESTROY*) if:*

    (a) $\Delta \in [\Delta_4, \Delta_3]$ *and* $p \geq p_3$

    (b) $\Delta > \Delta_3$ *and* $\forall p \in [0, 1 - \Delta]$

Case 3. $\frac{\pi}{F} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$.

In that case we know from Lemma 9 that there are two NE for all $p \in [0, 1 - \Delta]$ and for all $\Delta \in [0, 1)$. From Lemmata 8(2), 8(4), 8(9) and 8(10) for $\frac{\pi}{F} \in [\frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}, \frac{1}{\beta}]$, $\Delta_4 < \Delta_3 < 1$. Hence, if $\Delta < \Delta_4$, then $p_3 > 1 - \Delta$ and thus (KEEP, KEEP) Pareto dominates (DESTROY, DESTROY). If $\Delta \in [\Delta_4, \Delta_3]$ then for $p < p_3$ (KEEP, KEEP) Pareto dominates (DESTROY, DESTROY), while for $p \geq p_3$ (DESTROY, DESTROY) Pareto dominates (KEEP, KEEP). Moreover, for all $\Delta > \Delta_3$, $p_3 < 0$ so that (DESTROY, DESTROY) Pareto dominates (KEEP, KEEP). If on the other hand $\frac{\pi}{F} > \frac{1}{\beta}$ then both we know from Lemma 8(2) and 8(4) that both $\Delta_3 \geq 1$ and $\Delta_4 > 1$, implying that $p_3 \geq 0$ and $p_3 > 1 - \Delta$. In that case for all p and for all Δ (KEEP, KEEP) Pareto dominates (DESTROY, DESTROY).

We can summarize the above results into the following lemma.

Lemma 13. *For* $\frac{\pi}{F} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$ *the Pareto dominant NE of the Keep-Destroy subgame is:*

1. *(*KEEP, KEEP*)*

    (a) For $\frac{\pi}{F} \in [\frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}, \frac{1}{\beta}]$

        i. $\Delta < \Delta_4$ *and* $\forall p \in [0, 1 - \Delta]$

ii. $\Delta \in [\Delta_4, \Delta_3]$ *and* $p < p_3$

(b) *For* $\frac{\pi}{F} > \frac{1}{\beta}$

2. (DESTROY, DESTROY)

(a) *For* $\frac{\pi}{F} \in [\frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}, \frac{1}{\beta}]$

i. $\Delta \in [\Delta_4, \Delta_3]$ *and* $p \geq p_3$

ii. $\Delta > \Delta_3$ *and* $\forall p \in [0, 1-\Delta]$

From Lemmata 12 and 13 we reach the following proposition.

Proposition 1. *For given policy and other parameters (p, Δ, F, β, π) the Pareto dominant SPE of the baseline model without LP is:*

1. (KEEP, KEEP) *iff:*

(a) $\frac{\pi}{F} \in [1, \frac{1}{\beta}]$ *and*

i. $\Delta \in [0, \Delta_4]$ *and* $p \in [0, 1-\Delta]$ *or*

ii. $\Delta \in [\Delta_4, \Delta_3]$ *and* $p \in [0, p_3)$ *or*

(b) $\frac{\pi}{F} > \frac{1}{\beta}$, $\Delta \in [0, 1)$ *and* $p \in [0, 1-\Delta]$

2. (DESTROY, DESTROY) *iff* $\frac{\pi}{F} \in [1, \frac{1}{\beta}]$ *and*

(a) $\Delta \in [\Delta_4, \Delta_3]$ *and* $p \in [p_3, 1-\Delta]$ *or*

(b) $\Delta \in (\Delta_3, 1)$ *and* $p \in [0, 1-\Delta]$

where $p_3 \equiv \frac{\pi}{F} - \frac{\Delta}{\beta}$, $\Delta_3 \equiv \beta\frac{\pi}{F}$ *and* $\Delta_4 \equiv \frac{\beta}{1-\beta}(\frac{\pi}{F} - 1)$.

Appendix B - Proofs for Section 2.4

Proof of Lemma 6

Lemma 14.   *1. For $\delta \frac{\pi}{F} > 1$ the revelation subgame has two NE.*

2. For $\delta \frac{\pi}{F} \leq 1$ and $p + \Delta \leq \delta \frac{\pi}{F}$ the revelation subgame has also two NE.

Corollary 3. *For $\delta \frac{\pi}{F} \in [\frac{1}{2}, 1]$ and $p + \Delta > \delta \frac{\pi}{F}$ the revelation subgame has one NE (*KEEP AND REPORT, KEEP AND REPORT*).*

As discussed before in the case where we have two NE we apply the Pareto dominance criterion. It is easy to show that $V_{KR} > V_{KNR}$ iff $p + \Delta > \frac{1}{2}$.

Lemma 15. *(*KEEP AND REPORT, KEEP AND REPORT*) Pareto dominates (*KEEP AND NOT REPORT, KEEP AND NOT REPORT*) in the revelation subgame if $p + \Delta > \frac{1}{2}$.*

Therefore, from Lemmata 14 and 15 we obtain the following result:

Lemma 16.   *1. For $\delta \frac{\pi}{F} \in [\frac{1}{2}, 1]$*

   *(a) (*KEEP AND REPORT, KEEP AND REPORT*) Pareto dominates (*KEEP AND NOT REPORT, KEEP AND NOT REPORT*) iff $\frac{1}{2} < p + \Delta \leq \delta \frac{\pi}{F}$*

   *(b) (*KEEP AND NOT REPORT, KEEP AND NOT REPORT*) Pareto dominates (*KEEP AND REPORT, KEEP AND REPORT*) iff $p + \Delta \leq \frac{1}{2}$*

   2. For $\delta \frac{\pi}{F} > 1$

   *(a) (*KEEP AND REPORT, KEEP AND REPORT*) Pareto dominates (*KEEP AND NOT REPORT, KEEP AND NOT REPORT*) iff $p + \Delta > \frac{1}{2}$*

   *(b) (*KEEP AND NOT REPORT, KEEP AND NOT REPORT*) Pareto dominates (*KEEP AND REPORT, KEEP AND REPORT*) iff $p + \Delta \leq \frac{1}{2}$*

Conflation of Corollary 3 and Lemma 16 above boils down to the following lemma.

Lemma 17. *The Pareto dominant NE of the Revelation subgame is:*

*1. (*KEEP AND NOT REPORT, KEEP AND NOT REPORT*) iff $p + \Delta \leq \frac{1}{2}$;*

*2. (*KEEP AND REPORT, KEEP AND REPORT*) iff $p + \Delta > \frac{1}{2}$.*

Proof of Lemma 14

Lemma 18.   1. $p_4 < \delta \frac{\pi}{F}$

2. $p_4 < 0$ iff $\frac{\pi}{F} > \frac{1}{2\beta}$

3. $p_4 < \frac{1}{2} - \Delta$ iff $\Delta < \Delta_5$, where $\Delta_5 = \frac{\beta}{1-\beta}\frac{\pi}{F} - \frac{\beta}{2(1-\beta)}$

4. $\Delta_5 > 1$ iff $\frac{\pi}{F} > \frac{2-\beta}{2\beta}$

Proof of Lemma 18

(**1**) To show: $p_4 < \delta \frac{\pi}{F}$.

$$
\begin{aligned}
p_4 &< \delta \frac{\pi}{F} \\
\frac{F - 2\beta\Pi_M}{2(1-\beta)F} &< \delta \frac{\pi}{F} \\
\beta &> \frac{1 - 2\delta\frac{\pi}{F}}{2(1-\delta)\frac{\pi}{F}} \equiv \beta_1
\end{aligned}
$$

Notice that:

$$
\begin{aligned}
\beta_1 &> 0 \\
\frac{1 - 2\delta\frac{\pi}{F}}{2(1-\delta)\frac{\pi}{F}} &> 0 \\
\frac{\pi}{F} &< \frac{1}{2\delta} \\
\text{contradiction since } \frac{\pi}{F} &\geq 1
\end{aligned}
$$

$\therefore$ Given that $\beta_1 < 0$ it is always true that $p_4 < \delta\frac{\pi}{F}$.

(**2**) To show: $p_4 < 0$ iff $\frac{\pi}{F} > \frac{1}{2\beta}$.

$$
\begin{aligned}
p_4 &< 0 \Longleftrightarrow \\
\frac{F - 2\beta\Pi_M}{2(1-\beta)F} &< 0 \Longleftrightarrow \\
\frac{\pi}{F} &> \frac{1}{2\beta}
\end{aligned}
$$

$\therefore QED.$

(**3**) To show: $p_4 < \frac{1}{2} - \Delta$ iff $\Delta < \Delta_5$, where $\Delta_5 = \frac{\beta}{1-\beta}\frac{\pi}{F} - \frac{\beta}{2(1-\beta)}$.

$$
\begin{aligned}
p_4 &< \frac{1}{2} - \Delta \Longleftrightarrow \\
\Delta &< \frac{\beta}{1-\beta}\frac{\pi}{F} - \frac{\beta}{2(1-\beta)} \equiv \Delta_5
\end{aligned}
$$

$\therefore QED.$

(**4**) To show: $\Delta_5 > 1$ iff $\frac{\pi}{F} > \frac{2-\beta}{2\beta}$.

$$
\begin{aligned}
\Delta_5 &> 1 \Longleftrightarrow \\
\frac{\beta}{1-\beta}\frac{\pi}{F} - \frac{\beta}{2(1-\beta)} &> 1 \Longleftrightarrow \\
\frac{\pi}{F} &> \frac{2-\beta}{2\beta}
\end{aligned}
$$

$\therefore QED.$

Consider first Lemma 9(1), that is, $\delta\frac{\pi}{F} < 1 - \Delta$ so that the *Keep-Destroy* subgame has two NE for $p \in [\frac{1}{2} - \Delta, \delta\frac{\pi}{F}]$. Then, given Lemmata 10 and 18 above we can now proceed to the analysis to find the NE at the *Keep-Destroy* subgame. To begin with suppose that $\beta < \frac{1}{2}$. Then, $1 < \frac{1}{2\beta} < \frac{2-\beta}{2\beta}$. From Lemma 18(2) if $\frac{\pi}{F} > \frac{1}{2\beta}$, $p_4 < 0$ implying that (KEEP AND REPORT, KEEP AND REPORT) Pareto dominates (DESTROY, DESTROY) for all $p \in [\max\{0, \frac{1}{2} - \Delta\}, \delta\frac{\pi}{F}]$. If however $\frac{\pi}{F} < \frac{1}{2\beta}$ then from Lemma 18(3) $p_4 \in [0, \frac{1}{2} - \Delta]$ if $\Delta < \Delta_5$. Therefore, if $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$ and $\Delta < \Delta_5$ then (KEEP AND REPORT, KEEP AND REPORT) Pareto dominates (DESTROY, DESTROY) for all $p \in [\frac{1}{2} - \Delta, \delta\frac{\pi}{F}]$. If on the other hand $\Delta > \Delta_5$ then for all $p \in [\max\{0, \frac{1}{2} - \Delta\}, p_4]$ (DESTROY, DESTROY) Pareto dominates (KEEP AND REPORT, KEEP AND REPORT), whereas for all $p \in (p_4, \delta\frac{\pi}{F}]$ (KEEP AND REPORT, KEEP AND REPORT) Pareto dominates (DESTROY, DESTROY).

Suppose now that $\beta \in [\frac{1}{2}, \frac{2}{3}]$. Then $\frac{2-\beta}{2\beta} > 1 > \frac{1}{2\beta}$. From Lemma 18(2) if $\frac{\pi}{F} \in [1, \frac{2-\beta}{2\beta}]$ then $p_4 < 0$ so that (KEEP AND REPORT, KEEP AND REPORT) Pareto dominates (DESTROY, DESTROY) for all $p \in [\frac{1}{2} - \Delta, \delta\frac{\pi}{F}]$. Moreover, if $\frac{\pi}{F} > \frac{2-\beta}{2\beta}$ then again $p_4 < 0$ so that (KEEP AND REPORT, KEEP AND REPORT) Pareto dominates (DESTROY, DESTROY) for all $p \in [\max\{0, \frac{1}{2} - \Delta\}, \delta\frac{\pi}{F}]$.

166

Finally, suppose that $\beta > \frac{2}{3}$. Then $1 > \frac{2-\beta}{2\beta} > \frac{1}{2\beta}$. From Lemma 18(2) for all $\frac{\pi}{F} \geq 1$, $p_4 < 0$ implying that (Keep and Report, Keep and Report) Pareto dominates (Destroy, Destroy) for all $p \in [\max\{0, \frac{1}{2} - \Delta\}, \delta\frac{\pi}{F}]$.

Consider now Lemma 9(2), that is $\delta\frac{\pi}{F} > 1$ so that for all $p \in [\frac{1}{2} - \Delta, 1 - \Delta]$ there are two NE. In that case we have to check whether p_4 is greater or lower than $1 - \Delta$.

Lemma 19. *1. $p_4 < 1 - \Delta$ iff $\Delta < \Delta_6$, where $\Delta_6 = \frac{1-2\beta}{2(1-\beta)} + \frac{\beta}{1-\beta}\frac{\pi}{F}$*

2. $\Delta_6 > 1$ iff $\frac{\pi}{F} > \frac{1}{\beta}$

3. $\Delta_6 > \Delta_5$

Proof of Lemma 19
(1) To show: $p_4 < 1 - \Delta$ iff $\Delta < \Delta_6$ where $\Delta_6 = \frac{1-2\beta}{2(1-\beta)} + \frac{\beta}{1-\beta}\frac{\pi}{F}$.

$$
\begin{aligned}
p_4 &< 1 - \Delta \Leftrightarrow \\
\frac{F - 2\beta\Pi_M}{2(1-\beta)F} &< 1 - \Delta \Leftrightarrow \\
\Delta &< \frac{1-2\beta}{2(1-\beta)} + \frac{\beta}{1-\beta}\frac{\pi}{F} \equiv \Delta_6
\end{aligned}
$$

$\therefore QED.$
(**2**) To show: $\Delta_6 > 1$ iff $\frac{\pi}{F} > \frac{1}{\beta}$.

$$
\begin{aligned}
\Delta_6 &> 1 \Leftrightarrow \\
\frac{1-2\beta}{2(1-\beta)} + \frac{\beta}{1-\beta}\frac{\pi}{F} &> 1 \Leftrightarrow \\
\frac{\pi}{F} &> \frac{1}{\beta}
\end{aligned}
$$

$\therefore \Delta_6 > 1$ if $\frac{\pi}{F} > \frac{1}{\beta}$.

(**3**) To show: $\Delta_6 > \Delta_5$.

$$
\begin{aligned}
\Delta_6 &> \Delta_5 \\
\frac{1-2\beta}{2(1-\beta)} + \frac{\beta}{1-\beta}\frac{\pi}{F} &> \frac{\beta}{1-\beta}\frac{\pi}{F} - \frac{\beta}{2(1-\beta)} \\
\frac{1-2\beta}{2(1-\beta)} &> -\frac{\beta}{2(1-\beta)} \\
&\quad\; \frac{1-2\beta+\beta}{2(1-\beta)} \\
\frac{1}{2} &> 0
\end{aligned}
$$

$\therefore QED.$

Taking into consideration Lemma 19 above we can find the Pareto dominant NE of the *Keep-Destroy* subgame for different combinations of parameters' values. To begin with suppose that $\beta < \frac{1}{2}$. Then, $1 < \frac{1}{2\beta} < \frac{2-\beta}{2\beta} < \frac{1}{\beta}$. In that case, if $\frac{\pi}{F} > \frac{1}{2\beta}$ then $p_4 < 0$ implying that (KEEP AND REPORT, KEEP AND REPORT) Pareto dominates (DESTROY, DESTROY) for all $p \in [\max\{0, \frac{1}{2} - \Delta\}, 1 - \Delta]$. If on the other hand, $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$ then we can distinguish the following three cases:

1. From Lemma 18(3) if $\Delta < \Delta_5$ then $p_4 < \frac{1}{2} - \Delta$ so that for all $p \in [\frac{1}{2} - \Delta, 1 - \Delta]$ (KEEP AND REPORT, KEEP AND REPORT) Pareto dominates (DESTROY, DESTROY).

2. From Lemmata 18(3), 19(1) and 19(3) if $\Delta \in [\Delta_5, \Delta_6]$ then $p_4 > \frac{1}{2} - \Delta$ so that for $p \in [0, p_4]$ (DESTROY, DESTROY) Pareto dominates (KEEP AND REPORT, KEEP AND REPORT), while for $p \in (p_4, 1 - \Delta]$ (KEEP AND REPORT, KEEP AND REPORT) Pareto dominates (DESTROY, DESTROY).

3. From Lemma 19(1) if $\Delta > \Delta_6$ then $p_4 > 1 - \Delta$ so that for all $p \in [0, 1 - \Delta]$ (DESTROY, DESTROY) Pareto dominates (KEEP AND REPORT, KEEP AND REPORT).

Suppose now that $\beta \in [\frac{1}{2}, \frac{2}{3}]$. Then, $\frac{1}{2\beta} < 1 < \frac{2-\beta}{2\beta} < \frac{1}{\beta}$. From Lemma 18(2) for all $\frac{\pi}{F} \geq 1$ then $p_4 < 0$ so that (KEEP AND REPORT, KEEP AND REPORT) Pareto dominates (DESTROY, DESTROY) for all $p \in [\max\{0, \frac{1}{2} - \Delta\}, 1 - \Delta]$.

Finally, if $\beta > \frac{2}{3}$ then $\frac{1}{2\beta} < \frac{2-\beta}{2\beta} < 1 < \frac{1}{\beta}$. From 18(2) then for all $\frac{\pi}{F} \geq 1$, $p_4 < 0$ implying that (KEEP AND REPORT, KEEP AND REPORT) Pareto dominates (DESTROY, DESTROY) for all $p \in (\max\{0, \frac{1}{2} - \Delta\}, 1 - \Delta]$.

We can now summarize the above results in the lemma below.

168

Lemma 20. *The Pareto dominant NE of the Keep-Destroy subgame for* $p + \Delta > \frac{1}{2}$ *and* $\delta \frac{\pi}{F} \leq 1$ *is:*

1. (KEEP AND REPORT, KEEP AND REPORT) *iff:*

   (a) $\beta < \frac{1}{2}$, $\frac{\pi}{F} > \frac{1}{2\beta}$ *and* $p \in [\max\{0, \frac{1}{2} - \Delta\}, 1 - \Delta]$;

   (b) $\beta < \frac{1}{2}$, $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta < \Delta_5$ *and* $p \in [\frac{1}{2} - \Delta, 1 - \Delta]$;

   (c) $\beta < \frac{1}{2}$, $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta > \Delta_5$ *and* $p \in [p_4, 1 - \Delta]$;

   (d) $\beta > \frac{1}{2}$, $\frac{\pi}{F} \geq 1$ *and* $p \in [\max\{0, \frac{1}{2} - \Delta\}, 1 - \Delta]$.

2. (DESTROY, DESTROY) *iff* $\beta < \frac{1}{2}$, $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta > \Delta_5$ *and* $p \in [\max\{0, \frac{1}{2} - \Delta\}, p_4]$.

Lemma 21. *The Pareto dominant NE of the Keep-Destroy subgame for* $p > \frac{1}{2} - \Delta$ *and* $\delta \frac{\pi}{F} > 1$ *is:*

1. (KEEP AND REPORT, KEEP AND REPORT) *iff:*

   (a) $\beta < \frac{1}{2}$, $\frac{\pi}{F} > \frac{1}{2\beta}$ *and* $p \in [\max\{0, \frac{1}{2} - \Delta\}, 1 - \Delta]$;

   (b) $\beta < \frac{1}{2}$, $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta < \Delta_5$ *and* $p \in (\frac{1}{2} - \Delta, 1 - \Delta]$;

   (c) $\beta < \frac{1}{2}$, $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta \in [\Delta_5, \Delta_6]$ *and* $p \in (p_4, 1 - \Delta]$;

   (d) $\beta > \frac{1}{2}$, $\frac{\pi}{F} \geq 1$ *and* $p \in [\max\{0, \frac{1}{2} - \Delta\}, 1 - \Delta]$.

2. (DESTROY, DESTROY) *iff:*

   (a) $\beta < \frac{1}{2}$ *and* $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$ *and* $\Delta \in [\Delta_5, \Delta_6]$ *and* $p \in [\max\{0, \frac{1}{2} - \Delta\}, p_4]$;

   (b) $\beta < \frac{1}{2}$ *and* $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$ $\Delta > \Delta_6$ *and* $p \in [0, 1 - \Delta]$.

From Lemmata 20 and 21 we get the following Lemma.

Lemma 22. *The Pareto dominant SPE of the model with LP for* $p + \Delta > \frac{1}{2}$ *is:*

1. (KEEP AND REPORT, KEEP AND REPORT) *if:*

   (a) $\beta < \frac{1}{2}$ *and*

       i. $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta < \Delta_5$ *and* $p \in [\frac{1}{2} - \Delta, 1 - \Delta]$

       ii. $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta \in [\Delta_5, \Delta_6]$ *and* $p \in (p_4, 1 - \Delta]$

169

iii. $\frac{\pi}{F} > \frac{1}{2\beta}$ *and* $p \in [\max\{0, \frac{1}{2} - \Delta\}, 1 - \Delta]$

(b) $\beta > \frac{1}{2}$, $\frac{\pi}{F} \geq 1$ *and* $p \in [\max\{0, \frac{1}{2} - \Delta\}, 1 - \Delta]$.

2. *(*DESTROY, DESTROY*) if* $\beta < \frac{1}{2}$, $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$ *and*

    (a) $\Delta \in [\Delta_5, \Delta_6]$ *and* $p \in [\max\{0, \frac{1}{2} - \Delta\}, p_4]$

    (b) $\Delta > \Delta_6$ *and* $p \in [0, 1 - \Delta]$.

Proof of Lemma 15

Lemma 23.   *1.* $p_5 \geq 0$ *iff* $\Delta \leq \Delta_7$, *where* $\Delta_7 = \frac{\pi}{F} \frac{2(\delta + \beta - \beta\delta)}{\delta + \beta - \beta\delta + 1}$

2. $\Delta_7 \leq 1$ *iff* $\frac{\pi}{F} \in [1, \frac{\delta + \beta - \beta\delta + 1}{2(\delta + \beta - \beta\delta)}]$

3. $\frac{\delta + \beta - \beta\delta + 1}{2(\delta + \beta - \beta\delta)} > \frac{1}{2\delta}$

4. $p_5 + \Delta < \frac{1}{2}$ *iff* $\Delta > \Delta_8$, *where* $\Delta_8 = \frac{(2\frac{\pi}{F} - 1)(\beta + \delta - \beta\delta)}{1 - (\beta + \delta - \beta\delta)}$

5. $\Delta_8 > 1$

6. $\Delta_8 > \Delta_7$

Proof of Lemma 23

(1) To show: $p_5 > 0$ iff $\Delta < \Delta_7$, where $\Delta_7 = \frac{\pi}{F} \frac{2(\delta + \beta - \beta\delta)}{\delta + \beta - \beta\delta + 1}$.

$$
\begin{aligned}
p_5 &> 0 \Leftrightarrow \\
\frac{\pi}{F} - \frac{(\delta + \beta - \beta\delta + 1)\Delta}{2(\delta + \beta - \beta\delta)} &> 0 \Leftrightarrow \\
\Delta &< \frac{\pi}{F} \frac{2(\delta + \beta - \beta\delta)}{\delta + \beta - \beta\delta + 1} \equiv \Delta_7
\end{aligned}
$$

$\therefore QED.$

(2) To show: $\Delta_7 < 1$ iff $\frac{\pi}{F} \in [1, \frac{\delta + \beta - \beta\delta + 1}{2(\delta + \beta - \beta\delta)}]$.

$$
\begin{aligned}
\Delta_7 &< 1 \Leftrightarrow \\
\frac{\pi}{F} \frac{2(\delta + \beta - \beta\delta)}{\delta + \beta - \beta\delta + 1} &< 1 \Leftrightarrow \\
\frac{\pi}{F} &< \frac{\delta + \beta - \beta\delta + 1}{2(\delta + \beta - \beta\delta)}
\end{aligned}
$$

170

$\therefore QED.$

(3) To show: $\frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)} > \frac{1}{2\delta}$.

$$\frac{\delta + \beta - \beta\delta + 1}{2(\delta + \beta - \beta\delta)} > \frac{1}{2\delta}$$

$$\beta < \frac{\delta^2}{(1-\delta)^2}$$

Given that $\delta \in [\frac{1}{2}, 1]$, $\lim_{\delta\to\frac{1}{2}} \frac{\delta^2}{(1-\delta)^2} = 1$ and $\lim_{\delta\to 1} \frac{\delta^2}{(1-\delta)^2} = \infty$. Thus, given that $\beta \in [0,1]$, $\beta < \frac{\delta^2}{(1-\delta)^2}$. Another way to see this is the following. Observe that $\frac{\delta+\beta-\beta\delta+1}{\delta+\beta-\beta\delta} > 1$ and that $\max_{\delta\in[\frac{1}{2},1]} \frac{1}{2\delta} = 1$.

$\therefore QED.$

(4) To show: $p_5 + \Delta < \frac{1}{2}$ iff $\Delta > \Delta_8$, where $\Delta_8 = \frac{(2\frac{\pi}{F}-1)(\beta+\delta-\beta\delta)}{\Delta_8 1 - (\beta+\delta-\beta\delta)}$.

$$p_5 + \Delta < \frac{1}{2} \Leftrightarrow$$

$$\frac{\pi}{F} - \frac{(\delta+\beta-\beta\delta+1)\Delta}{2(\delta+\beta-\beta\delta)} + \Delta < \frac{1}{2} \Leftrightarrow$$

$$\Delta > \frac{(2\frac{\pi}{F}-1)(\beta+\delta-\beta\delta)}{1-(\beta+\delta-\beta\delta)} \equiv \Delta_8$$

$\therefore QED.$

(5) To show: $\Delta_8 > 1$.

$$\frac{(2\frac{\pi}{F}-1)(\beta+\delta-\beta\delta)}{1-(\beta+\delta-\beta\delta)} > 1$$

$$(2\frac{\pi}{F}-1)(\beta+\delta-\beta\delta) > 1-(\beta+\delta-\beta\delta)$$

$$\frac{\pi}{F} > \frac{1-(\beta+\delta-\beta\delta)+(\beta+\delta-\beta\delta)}{2(\beta+\delta-\beta\delta)}$$

$$\frac{\pi}{F} > \frac{1}{2(\beta+\delta-\beta\delta)}$$

$$\frac{1}{2(\beta + \delta - \beta\delta)} \quad < \quad 1$$

$$1 \quad < \quad 2(\beta + \delta - \beta\delta)$$

$$\beta + \delta - \beta\delta \quad > \quad \frac{1}{2}$$

$$\text{true given that } \delta \quad \geq \quad \frac{1}{2}$$

$\therefore \Delta_8 > 1$.

(6) To show: $\Delta_7 > \Delta_8$.

$$\Delta_7 \quad > \quad \Delta_8$$

$$\frac{\pi}{F}\frac{2(\delta + \beta - \beta\delta)}{\delta + \beta - \beta\delta + 1} \quad > \quad \frac{(2\frac{\pi}{F} - 1)(\beta + \delta - \beta\delta)}{1 - (\beta + \delta - \beta\delta)}$$

$$\frac{\pi}{F} \quad < \quad \frac{1 + \delta + \beta - \beta\delta}{4(\delta + \beta - \beta\delta)}$$

Notice however that given that we are in the area where $\frac{\pi}{F} > \frac{1}{2\delta}$ for this to hold it must be the case that:

$$\frac{1 + \delta + \beta - \beta\delta}{4(\delta + \beta - \beta\delta)} \quad > \quad \frac{1}{2\delta}$$

$$\delta(1 - \beta) + (1 - \delta)(2 - \beta) \quad < \quad 0$$

This, however, implies that $\frac{\pi}{F} < \frac{1+\delta+\beta-\beta\delta}{4(\delta+\beta-\beta\delta)} < \frac{1}{2\delta}$ which is a contradiction.
$\therefore QED$.

Notice that Lemma 23(5) imply that $p_5 > \frac{1}{2} - \Delta$.

Corollary 4. $p_5 > \frac{1}{2} - \Delta$.

From Lemma 23(1) $p_5 > 0 \ \forall \Delta$ iff $\frac{\pi}{F} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$, while $p_5 > 0$ for all $\Delta \leq \Delta_7$ iff $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$. Therefore, for $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$ and $\Delta < \Delta_7$ and for all $p \in [0, \frac{1}{2} - \Delta]$ (KEEP AND NOT REPORT, KEEP AND NOT REPORT) is a NE of the *Keep-Destroy* subgame. If however $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$ then from Lemmata 23(1) and 23(2) and $\Delta > \Delta_7$ then $p_5 < 0$. In the latter case (KEEP NOT REPORT, KEEP NOT REPORT) is not a NE. On the other hand, from Lemmata 23(1), 23(2), 23(4) and 23(5) if $\frac{\pi}{F} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$ then both $\Delta_8 > 1$ and $\Delta_7 > 1$, so that for all $p \in [0, \frac{1}{2} - \Delta]$ (KEEP AND NOT REPORT, KEEP AND NOT REPORT) form a NE of the *Keep-Destroy* subgame.

Lemma 24. *The Pareto dominant NE of the* Keep-Destroy *subgame for* $p < \frac{1}{2} - \Delta$ *is:*

1. $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$ *and* $\Delta \leq \Delta_7$ *and* $p \in [0, \frac{1}{2} - \Delta]$

2. $\frac{\pi}{F} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$ *and* $p \in [0, \frac{1}{2} - \Delta]$

Corollary 5. *If* $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$ *and* $\Delta > \Delta_7$ *the unique NE of the Keep-Destroy subgame for* $p \leq \frac{1}{2} - \Delta$ *is (*DESTROY, DESTROY*).*

Lemma 25.   *1.* $p_3 \geq 0$ *iff* $\Delta \leq \Delta_3$*, where* $\Delta_3 \equiv \beta\frac{\pi}{F}$*.*

2. $\Delta_3 > 1$ *iff* $\frac{\pi}{F} > \frac{1}{\beta}$

3. $p_3 \leq \frac{1}{2} - \Delta$ *iff* $\Delta \geq \Delta_5$*, where* $\Delta_5 \equiv \frac{\beta}{1-\beta}(\frac{\pi}{F} - \frac{1}{2})$

4. $\Delta_5 > 1$ *iff* $\frac{\pi}{F} > \frac{2-\beta}{2\beta}$ *(from Lemma 18(4))*

5. $\Delta_5 > \Delta_3$ *iff* $\frac{\pi}{F} > \frac{1}{2\beta}$

6. $\frac{1}{\beta} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$

7. $\frac{2-\beta}{2\beta} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$

8. $\frac{2-\beta}{2\beta} < \frac{1}{\beta}$

9. $\frac{1}{2\beta} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$ *if* $\beta < 0.414$ *or if* $\beta \in [0.414, 0.5]$ *and* $\delta > \frac{\beta^2}{(1-\beta)^2}$

<u>Proof of Lemma 25</u>
(1) To show: $p_3 > 0$ if $\Delta < \Delta_3$, where $\Delta_3 = \beta\frac{\pi}{F}$.

$$
\begin{aligned}
p_3 &> 0 \Leftrightarrow \\
\frac{\pi}{F} - \frac{\Delta}{\beta} &> 0 \Leftrightarrow \\
\Delta &< \beta\frac{\pi}{F} \equiv \Delta_3
\end{aligned}
$$

$\therefore QED.$

(2) To show: $\frac{1}{\beta} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$.

Let $x \equiv \delta + \beta - \beta\delta$. Notice that $x < 1$. Then,

$$
\begin{aligned}
\frac{1}{\beta} &> \frac{x+1}{2x} \\
2x &> \beta x + x \\
x(2-\beta) &> \beta
\end{aligned}
$$

$$
\begin{aligned}
\frac{1}{\beta} &> \frac{\delta + \beta - \beta\delta + 1}{2(\delta + \beta - \beta\delta)} \\
(2\delta + \beta(1-\delta))(1-\beta) &> 0
\end{aligned}
$$

Given that $x > \beta$ it is the case also that $x(2-\beta) > \beta$.
$\therefore QED.$

(3) To show: $p_3 < \frac{1}{2} - \Delta$ if $\Delta > \Delta_5$, where $\Delta_5 = \frac{\beta}{1-\beta}(\frac{\pi}{F} - \frac{1}{2})$.

$$
\begin{aligned}
p_3 &< \frac{1}{2} - \Delta \Leftrightarrow \\
\Delta &> \frac{\beta}{1-\beta}(\frac{\pi}{F} - \frac{1}{2}) \equiv \Delta_5
\end{aligned}
$$

$\therefore QED.$

(4) To show: $\Delta_5 > 1$ if $\frac{\pi}{F} > \frac{2-\beta}{2\beta}$.

$$
\begin{aligned}
\Delta_5 &> 1 \Leftrightarrow \\
\frac{\beta}{1-\beta}(\frac{\pi}{F} - \frac{1}{2}) &> 1 \Leftrightarrow \\
\frac{\pi}{F} &> \frac{2-\beta}{2\beta}
\end{aligned}
$$

$\therefore QED.$
(5) To show: $\Delta_3 > 1$ if $\frac{\pi}{F} > \frac{1}{\beta}$.

$$
\begin{aligned}
\Delta_3 &> 1 \Leftrightarrow \\
\beta\frac{\pi}{F} &> 1 \Leftrightarrow \\
\frac{\pi}{F} &> \frac{1}{\beta}
\end{aligned}
$$

174

$\therefore QED.$

(6) To show: $\frac{2-\beta}{2\beta} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$.

By contradiction. Suppose that $\frac{2-\beta}{2\beta} < \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$. Then:

$$\frac{2-\beta}{2\beta} < \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$$
$$\delta < \frac{\beta(2\beta-1)}{2\beta^2-4\beta+2}$$

Now, observe that $\frac{\beta(2\beta-1)}{2\beta^2-4\beta+2} < \frac{1}{2}$.

$$\frac{\beta(2\beta-1)}{2\beta^2-4\beta+2} < \frac{1}{2}$$
$$\beta^2 + \beta < 2$$
$$\text{always true}$$

This implies that $\delta < \frac{\beta(2\beta-1)}{2\beta^2-4\beta+2} < \frac{1}{2}$ which is a contradiction, given assumption 2.

$\therefore QED.$

(7) To show: $\frac{2-\beta}{2\beta} < \frac{1}{\beta}$.

$$\frac{2-\beta}{2\beta} < \frac{1}{\beta}$$
$$\beta > 0$$

$\therefore QED.$

(8) To show: $\Delta_5 > \Delta_3$ iff $\frac{\pi}{F} > \frac{1}{2\beta}$.

$$\Delta_5 > \Delta_3 \Leftrightarrow$$
$$\frac{\beta}{1-\beta}(\frac{\pi}{F} - \frac{1}{2}) > \beta\frac{\pi}{F} \Leftrightarrow$$
$$\frac{\pi}{F} > \frac{1}{2\beta}$$

$\therefore QED.$

(9) To show: $\frac{1}{2\beta} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$ if $\beta < 0,414$ or if $\beta \in [0.414, 0, 5]$ and $\delta > \frac{\beta^2}{(1-\beta)^2}$.

175

Let $x \equiv \delta + \beta - \beta\delta$. Then:

$$\frac{1}{2\beta} > \frac{\delta + \beta - \beta\delta + 1}{2(\delta + \beta - \beta\delta)}$$
$$x > \beta x + \beta$$
$$x > \frac{\beta}{1 - \beta}$$

Now notice that $\frac{\beta}{1-\beta} > 1$ if $\beta > \frac{1}{2}$. However, this leads to a contradiction since $x < 1$. Therefore. if $\beta > \frac{1}{2}$ then $\frac{1}{2\beta} < \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$.

Now suppose that $\beta < \frac{1}{2}$. In this case,

$$\frac{1}{2\beta} > \frac{\delta + \beta - \beta\delta + 1}{2(\delta + \beta - \beta\delta)}$$
$$\delta > \frac{\beta^2}{(1 - \beta)^2}$$

Notice that $\frac{\beta^2}{(1-\beta)^2} < \frac{1}{2}$ if $\beta < 0,414$. Therefore, for $\beta < 0,414$ $\frac{1}{2\beta} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$. Moreover, for $\beta \in [0.414, 0, 5]$ and $\delta > \frac{\beta^2}{(1-\beta)^2}$. Conversely, if $\beta \in [0.414, 0, 5]$ and $\delta < \frac{\beta^2}{(1-\beta)^2}$ then $\frac{1}{2\beta} < \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$.

$\therefore QED$.

Case 1. $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$ *so that* $\Delta_7 < 1$

1. From Lemmata 25(8) and 25(9) if $\beta < 0.414$ or $\beta \in [0.414, 0.5]$ and $\delta > \frac{\beta^2}{(1-\beta)^2}$ then $\frac{1}{\beta} > \frac{2-\beta}{2\beta} > \frac{1}{2\beta} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)} > 1$. Therefore, given that $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$ and taking into account Lemmata 25(1), 25(2), 25(3), 25(4) and 25(5) we may distinguish three subcases:

   (a) if $\Delta < \Delta_5$ then $p_3 \geq 0$ and $p_3 > \frac{1}{2} - \Delta$. Hence (KEEP AND NOT RE-PORT, KEEP AND NOT REPORT) Pareto dominates (DESTROY, DE-STROY) for all $p \in [0, \frac{1}{2} - \Delta]$.

   (b) if $\Delta \in [\Delta_5, \Delta_3]$ then $p_3 \geq 0$ and $p_3 \leq \frac{1}{2} - \Delta$. Hence for $p \in [0, p_3)$ (KEEP AND NOT REPORT, KEEP AND NOT REPORT) Pareto dom-inates (DESTROY, DESTROY), while for $p \in [p_3, \frac{1}{2} - \Delta]$ (DESTROY, DESTROY) Pareto dominates (KEEP AND NOT REPORT, KEEP AND NOT REPORT).

(c) if $\Delta \in [\Delta_3, \frac{1}{2}]$ then $p_3 < 0$ and $p_3 \leq \frac{1}{2} - \Delta$. Hence (DESTROY, DESTROY) Pareto dominates (KEEP AND NOT REPORT, KEEP AND NOT REPORT) for all $p \in [0, \frac{1}{2} - \Delta]$. Notice that we restrict $\Delta < \frac{1}{2}$ so that the set $[0, \frac{1}{2} - \Delta]$ is non-empty.

2. From Lemmata 25(7), 25(8) and 25(9) if $\beta \in [0.414, 0.5]$ and $\delta < \frac{\beta^2}{(1-\beta)^2}$ then $\frac{1}{\beta} > \frac{2-\beta}{2\beta} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)} > \frac{1}{2\beta} > 1$.

(a) If $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$ then from Lemma 25(5) $\Delta_3 > \Delta_5$. Taking into consideration Lemmata 25(1), 25(2), 25(3) and 25(4) we may distinguish the following three cases:

   i. if $\Delta < \Delta_5$ then $p_3 \geq 0$ and $p_3 > \frac{1}{2} - \Delta$. Hence (KEEP AND NOT REPORT, KEEP AND NOT REPORT) Pareto dominates (DESTROY, DESTROY) for all $p \in [0, \frac{1}{2} - \Delta]$.

   ii. if $\Delta \in [\Delta_5, \Delta_3]$ then $p_3 \geq 0$ and $p_3 \leq \frac{1}{2} - \Delta$. Hence for $p \in [0, p_3)$ (KEEP AND NOT REPORT, KEEP AND NOT REPORT) Pareto dominates (DESTROY, DESTROY), while for $p \in [p_3, \frac{1}{2} - \Delta]$ (DESTROY, DESTROY) Pareto dominates (KEEP AND NOT REPORT, KEEP AND NOT REPORT).

   iii. if $\Delta \in [\Delta_3, \frac{1}{2}]$ then $p_3 < 0$ and $p_3 \leq \frac{1}{2} - \Delta$. Hence (DESTROY, DESTROY) Pareto dominates (KEEP AND NOT REPORT, KEEP AND NOT REPORT) for all $p \in [0, \frac{1}{2} - \Delta]$.

(b) If $\frac{\pi}{F} \in [\frac{1}{2\beta}, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$ then from Lemma 25(5) $\Delta_5 > \Delta_3 > \frac{1}{2}$. Taking into consideration Lemmata 25(1), 25(2), 25(3) and 25(4) if $\Delta < \frac{1}{2}$ then $p_3 \geq 0$ and $p_3 > \frac{1}{2} - \Delta$. Hence, (KEEP AND NOT REPORT, KEEP AND NOT REPORT) Pareto dominates (DESTROY, DESTROY).

3. From Lemmata 25(7), 25(8) if $\beta \in [\frac{1}{2}, \frac{2}{3}]$ then $\frac{1}{\beta} > \frac{2-\beta}{2\beta} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)} > 1 > \frac{1}{2\beta}$. By taking into consideration Lemmata 25(2), 25(4) and 25(5) if $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$ then $\Delta_5 > \Delta_3 > \frac{1}{2}$. By Lemmata 25(1) and 25(3) if $\Delta < \frac{1}{2}$ then $p_3 \geq 0$ and $p_3 > \frac{1}{2} - \Delta$. Hence for $\Delta < \frac{1}{2}$ (KEEP AND NOT REPORT, KEEP AND NOT REPORT) Pareto dominates (DESTROY, DESTROY) for all $p \in [0, \frac{1}{2} - \Delta]$.

4. From Lemmata 25(7), 25(8) if $\beta > \frac{2}{3}$ then $\frac{1}{\beta} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)} > 1 > \frac{2-\beta}{2\beta} > \frac{1}{2\beta}$. By taking into account Lemmata 25(2), 25(4) and 25(5) $1 > \Delta_5 > \Delta_3 > \frac{1}{2}$. Hence, from Lemmata 25(1) and 25(3) if $\Delta < \frac{1}{2}$ then $p_3 \geq 0$ and $p_3 > \frac{1}{2} - \Delta$. Therefore, for $\Delta < \frac{1}{2}$ (KEEP AND NOT REPORT, KEEP AND NOT REPORT) Pareto dominates (DESTROY, DESTROY) for all $p \in [0, \frac{1}{2} - \Delta]$.

Case 2. $\frac{\pi}{F} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$ *so that* $\Delta_7 > 1$

Consider the second case now where $\Delta_7 > 1$ so that for all Δ and $p \le \frac{1}{2} - \Delta$ we have two NE.

1. From Lemmata 25(7), 25(8) and 25(9) if $\beta < 0.414$ or $\beta \in [0.414, 0.5]$ and $\delta > \frac{\beta^2}{(1-\beta)^2}$ then $\frac{1}{\beta} > \frac{2-\beta}{2\beta} > \frac{1}{2\beta} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)} > 1$. Therefore, given that $\frac{\pi}{F} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}$ and taking into account Lemmata 25(1), 25(2), 25(3), 25(4) and 25(5) we may then distinguish four cases:

   (a) if $\frac{\pi}{F} > \frac{1}{\beta}$ then both $\Delta_5 > \Delta_3 > 1$ implying that $p_3 \ge 0$ and $p_3 > \frac{1}{2} - \Delta$. Therefore for $\Delta < \frac{1}{2}$ (KEEP AND NOT REPORT, KEEP AND NOT REPORT) Pareto dominates (DESTROY, DESTROY) for all $p \in [0, \frac{1}{2} - \Delta]$.

   (b) if $\frac{\pi}{F} \in [\frac{2-\beta}{2\beta}, \frac{1}{\beta}]$ then $\Delta_5 > 1 > \Delta_3 > \frac{1}{2}$. Therefore, for $\Delta < \frac{1}{2}$ then $p_3 \ge 0$ and $p_3 > \frac{1}{2} - \Delta$ and hence (KEEP AND NOT REPORT, KEEP AND NOT REPORT) Pareto dominates (DESTROY, DESTROY) for all $p \in [0, \frac{1}{2} - \Delta]$.

   (c) if $\frac{\pi}{F} \in [\frac{1}{2\beta}, \frac{2-\beta}{2\beta}]$ then $1 > \Delta_5 > \Delta_3 > \frac{1}{2}$. Therefore, for $\Delta < \frac{1}{2}$ then $p_3 \ge 0$ and $p_3 > \frac{1}{2} - \Delta$ and hence (KEEP AND NOT REPORT, KEEP AND NOT REPORT) Pareto dominates (DESTROY, DESTROY) for all $p \in [0, \frac{1}{2} - \Delta]$.

   (d) $\frac{\pi}{F} \in [\frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}, \frac{1}{2\beta}]$ then $\frac{1}{2} > \Delta_3 > \Delta_5$. Therefore:
      i. if $\Delta < \Delta_5$, then $p_3 > 0$ and $p_3 > \frac{1}{2} - \Delta$ and hence (KEEP AND NOT REPORT, KEEP AND NOT REPORT) Pareto dominates (DESTROY, DESTROY).
      ii. if $\Delta \in [\Delta_5, \Delta_3]$ then $p_3 \ge 0$ and $p_3 \le \frac{1}{2} - \Delta$ and hence for $p < p_3$ (KEEP AND NOT REPORT, KEEP AND NOT REPORT) Pareto dominates (DESTROY, DESTROY), while for $p \in [p_3, \frac{1}{2} - \Delta]$ (DESTROY, DESTROY) Pareto dominates (KEEP AND NOT REPORT, KEEP AND NOT REPORT).
      iii. if $\Delta > (\Delta_3, \frac{1}{2}]$ then $p_3 < 0$ and hence (DESTROY, DESTROY) Pareto dominates (KEEP AND NOT REPORT, KEEP AND NOT REPORT) for all $p \in [0, \frac{1}{2} - \Delta]$.

2. From Lemmata 25(7), 25(8) and 25(9) if $\beta \in [0, 414, 0, 5]$ and $\delta < \frac{\beta^2}{(1-\beta)^2}$ then $\frac{1}{\beta} > \frac{2-\beta}{2\beta} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)} > \frac{1}{2\beta} > 1$. By taking into account Lemmata 25(1), 25(2), 25(3), 25(4) and 25(5) we may then distinguish the following three cases:

178

(a) if $\frac{\pi}{F} > \frac{1}{\beta}$ then both $\Delta_5 > \Delta_3 > \frac{1}{2}$ implying that for all $\Delta < \frac{1}{2}$ $p_3 > 0$ and $p_3 > \frac{1}{2} - \Delta$. Therefore (KEEP AND NOT REPORT, KEEP AND NOT REPORT) Pareto dominates (DESTROY, DESTROY) for all $p \in [0, \frac{1}{2} - \Delta]$.

(b) if $\frac{\pi}{F} \in [\frac{2-\beta}{2\beta}, \frac{1}{\beta}]$ then $\Delta_5 > 1 > \Delta_3 > \frac{1}{2}$. Therefore, for $\Delta < \frac{1}{2}$ then $p_3 \geq 0$ and $p_3 > \frac{1}{2} - \Delta$ and hence (KEEP AND NOT REPORT, KEEP AND NOT REPORT) Pareto dominates (DESTROY, DESTROY) for all $p \in [0, \frac{1}{2} - \Delta]$.

(c) if $\frac{\pi}{F} \in [\frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}, \frac{2-\beta}{2\beta}]$ then $1 > \Delta_5 > \Delta_3 > \frac{1}{2}$. Therefore, for $\Delta < \frac{1}{2}$ then $p_3 \geq 0$ and $p_3 > \frac{1}{2} - \Delta$ and hence (KEEP AND NOT REPORT, KEEP AND NOT REPORT) Pareto dominates (DESTROY, DESTROY) for all $p \in [0, \frac{1}{2} - \Delta]$.

3. From Lemmata 25(7) and 25(8) if $\beta \in [\frac{1}{2}, \frac{2}{3}]$ then $\frac{1}{\beta} > \frac{2-\beta}{2\beta} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)} > 1 > \frac{1}{2\beta}$. By taking into consideration Lemmata 25(1), 25(2), 25(3), 25(4) and 25(5) we may then distinguish the following three cases:

(a) if $\frac{\pi}{F} > \frac{1}{\beta}$ then both $\Delta_5 > \Delta_3 > 1$ implying that for all $\Delta < \frac{1}{2}$, $p_3 > 0$ and $p_3 > \frac{1}{2} - \Delta$. Therefore, for all $\Delta < \frac{1}{2}$ (KEEP AND NOT REPORT, KEEP AND NOT REPORT) Pareto dominates (DESTROY, DESTROY) for all $p \in [0, \frac{1}{2} - \Delta]$.

(b) if $\frac{\pi}{F} \in [\frac{2-\beta}{2\beta}, \frac{1}{\beta}]$ then $\Delta_5 > 1 > \Delta_3 > \frac{1}{2}$. Therefore, for $\Delta < \frac{1}{2}$ then $p_3 \geq 0$ and $p_3 > \frac{1}{2} - \Delta$ and hence (KEEP AND NOT REPORT, KEEP AND NOT REPORT) Pareto dominates (DESTROY, DESTROY) for all $p \in [0, \frac{1}{2} - \Delta]$.

(c) if $\frac{\pi}{F} \in [\frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}, \frac{2-\beta}{2\beta}]$ then $1 > \Delta_5 > \Delta_3 > \frac{1}{2}$. Therefore, for $\Delta < \frac{1}{2}$ then $p_3 \geq 0$ and $p_3 > \frac{1}{2} - \Delta$ and hence (KEEP AND NOT REPORT, KEEP AND NOT REPORT) Pareto dominates (DESTROY, DESTROY) for all $p \in [0, \frac{1}{2} - \Delta]$.

4. From Lemmata 25(7) and 25(8) if $\beta > \frac{2}{3}$ then $\frac{1}{\beta} > \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)} > 1 > \frac{2-\beta}{2\beta} > \frac{1}{2\beta}$. By taking into consideration Lemmata 25(1), 25(2), 25(3), 25(4) and 25(5) we may then distinguish the following two cases:

(a) $\frac{\pi}{F} > \frac{1}{\beta}$ then both $\Delta_5 > \Delta_3 > \frac{1}{2}$ implying that for all $\Delta < \frac{1}{2}$, $p_3 \geq 0$ and $p_3 > \frac{1}{2} - \Delta$. Therefore (KEEP AND NOT REPORT, KEEP AND NOT REPORT) Pareto dominates (DESTROY, DESTROY) for all $p \in [0, \frac{1}{2} - \Delta]$.

(b) $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$ then $\Delta_5 > 1 > \Delta_3 > \frac{1}{2}$. Therefore, for $\Delta < \frac{1}{2}$ then $p_3 \geq 0$ and $p_3 > \frac{1}{2} - \Delta$ and hence (KEEP AND NOT REPORT, KEEP AND NOT REPORT) Pareto dominates (DESTROY, DESTROY) for all $p \in [0, \frac{1}{2} - \Delta]$.

We can now summarize the results of the above analysis to the following lemma.

Lemma 26. *The SPE of the game is (*KEEP AND NOT REPORT, KEEP AND NOT REPORT*) if :*

1. $\beta < 0.414$

   (a) $\frac{\pi}{F} \in [\frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}, \frac{1}{2\beta}]$ and $\Delta_7 > 1$ OR $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$ and $\Delta_7 < 1$

      i. $\Delta < \Delta_5$ and $p \in [0, \frac{1}{2} - \Delta]$
      ii. $\Delta \in [\Delta_5, \Delta_3]$ and $p \in [0, p_3)$

   (b) $\frac{\pi}{F} \in [\frac{1}{2\beta}, \frac{1}{\beta}]$, $\Delta < \frac{1}{2}(< \Delta_3 < 1 < \Delta_7)$ and $p \in [0, \frac{1}{2} - \Delta]$

   (c) $\frac{\pi}{F} > \frac{1}{\beta}$, $\Delta < \frac{1}{2} < 1 < \Delta_7$ and $p \in [0, \frac{1}{2} - \Delta]$

2. $\beta \in [0.414, 0.5]$

   (a) $\delta < \frac{\beta^2}{(1-\beta)^2}$

      i. $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta < \Delta_5(< \frac{1}{2} < \Delta_7 < 1)$ and $p \in [0, \frac{1}{2} - \Delta]$
      ii. $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta \in [\Delta_5, \Delta_3](< \frac{1}{2} < \Delta_7 < 1)$ and $p \in [0, p_3)$
      iii. $\frac{\pi}{F} \in [\frac{1}{2\beta}, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$, $\Delta < \frac{1}{2}(< \Delta_3 < \Delta_7 < 1)$ and $p \in [0, \frac{1}{2} - \Delta]$
      iv. $\frac{\pi}{F} \in [\frac{1}{2\beta}, \frac{1}{\beta}]$, $\Delta < \frac{1}{2}(< \Delta_3 < 1 < \Delta_7)$ and $p \in [0, \frac{1}{2} - \Delta]$
      v. $\frac{\pi}{F} > \frac{1}{\beta}$, $\Delta < \frac{1}{2}$ and $p \in [0, \frac{1}{2} - \Delta]$

   (b) $\delta > \frac{\beta^2}{(1-\beta)^2}$

      i. $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$, $\Delta < \Delta_5(< \Delta_7 < 1)$ and $p \in [0, \frac{1}{2} - \Delta]$
      ii. $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$, $\Delta \in [\Delta_5, \Delta_3](< \frac{1}{2} < \Delta_7 < 1)$ and $p \in [0, p_3)$
      iii. $\frac{\pi}{F} \in [\frac{1}{2\beta}, \frac{1}{\beta}]$, $\Delta < \frac{1}{2}(< \Delta_3 < 1 < \Delta_7)$ and $p \in [0, \frac{1}{2} - \Delta]$
      iv. $\frac{\pi}{F} \in [\frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}, \frac{1}{2\beta}]$, $\Delta < \Delta_5(< 1 < \Delta_7)$ and $p \in [0, \frac{1}{2} - \Delta]$
      v. $\frac{\pi}{F} \in [\frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}, \frac{1}{2\beta}]$, $\Delta \in [\Delta_5, \Delta_3](< \frac{1}{2} < 1 < \Delta_7)$ and $p \in [0, p_3)$
      vi. $\frac{\pi}{F} > \frac{1}{\beta}$, $\Delta < \frac{1}{2}(< 1 < \Delta_7)$ and $p \in [0, \frac{1}{2} - \Delta]$

3. $\beta > 0.5$

(a) $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$, $\Delta < \frac{1}{2}(< \Delta_3 < \Delta_7 < 1)$ and $p \in [0, \frac{1}{2} - \Delta]$

(b) $\frac{\pi}{F} \in [\frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}, \frac{1}{\beta}]$, $\Delta < \frac{1}{2}(< \Delta_3 < 1 < \Delta_7)$ and $p \in [0, \frac{1}{2} - \Delta]$

(c) $\frac{\pi}{F} > \frac{1}{\beta}$, $\forall \Delta < \frac{1}{2}(< 1 < \Delta_7)$ and $p \in [0, \frac{1}{2} - \Delta]$

Lemma 26(2(a)iii) and 26(2(a)iv) imply the following result.

Corollary 6. *(*Keep and not Report, Keep and Not Report*) is the SPE if $\beta \in [0.414, 0.5]$, $\delta < \frac{\beta^2}{(1-\beta)^2}$, $\frac{\pi}{F} \in [\frac{1}{2\beta}, \frac{1}{\beta}]$, $\Delta < \frac{1}{2}$ and $p \in [0, \frac{1}{2} - \Delta]$.*

From Lemma 26(2(b)i) and 26(2(b)iv) we can get the corollary below.

Corollary 7. *(*Keep and not Report, Keep and Not Report*) is the SPE if $\beta \in [0.414, 0.5]$, $\delta > \frac{\beta^2}{(1-\beta)^2}$, $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta < \Delta_5$ and $p \in [0, \frac{1}{2} - \Delta]$.*

Lemma 26(2(b)ii) and 26(2(b)v) imply the following.

Corollary 8. *(*Keep and not Report, Keep and Not Report*) is the SPE if $\beta \in [0.414, 0.5]$, $\delta > \frac{\beta^2}{(1-\beta)^2}$, $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta \in [\Delta_5, \Delta_3]$ and $p \in [0, p_3)$.*

Lemma 26(3a) and 26(3b) imply the following.

Corollary 9. *(*Keep and not Report, Keep and Not Report*) is the SPE if $\frac{\pi}{F} \in [1, \frac{1}{\beta}]$, $\Delta < \frac{1}{2}$ and $p \in [0, \frac{1}{2} - \Delta]$.*

From Lemma 26 and Corollaries 6, 7, 8 and 9 we get the following result, which identifies the parameters' values for which (Keep and Not Report, Keep and Not Report) is the Pareto dominant SPE for $p \leq \frac{1}{2} - \Delta$.

Similarly, by summarizing the results of our previous analysis we find the regions for the parameters' values for which (Destroy, Destroy) form a Pareto dominant SPE for $p \leq \frac{1}{2} - \Delta$.

Lemma 27. *For $p \in [0, \frac{1}{2} - \Delta]$ (*Destroy, Destroy*) is the Pareto dominant SPE if:*

1. $\beta < 0.414$

(a) $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$, $\Delta \in [\Delta_5, \Delta_3](< \Delta_7 < 1)$ and $p \in [p_3, \frac{1}{2} - \Delta]$

(b) $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$, $\Delta \in [\Delta_3, \frac{1}{2}](< \Delta_7 < 1)$ and $p \in [0, \frac{1}{2} - \Delta]$

(c) $\frac{\pi}{F} \in [\frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}, \frac{1}{2\beta}]$, $\Delta \in [\Delta_3, \frac{1}{2}](< 1 < \Delta_7)$ and $p \in [0, \frac{1}{2} - \Delta]$

(d) $\frac{\pi}{F} \in [\frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}, \frac{1}{2\beta}]$, $[\Delta_5, \Delta_3](< \frac{1}{2} < 1 < \Delta_7)$ and $p \in [p_3, \frac{1}{2} - \Delta]$

2. $\beta \in [0.414, 0.5]$

(a) $\delta < \frac{\beta^2}{(1-\beta)^2}$

 i. $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta \in [\Delta_5, \Delta_3](< \Delta_7 < 1)$ and $p \in [p_3, \frac{1}{2} - \Delta]$

 ii. $\frac{\pi}{F} \in [\frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}, \frac{1}{2\beta}]$, $\Delta \in [\Delta_3, \frac{1}{2}](< 1 < \Delta_7)$ and $p \in [0, \frac{1}{2} - \Delta]$

(b) $\delta > \frac{\beta^2}{(1-\beta)^2}$

 i. $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$, $\Delta \in [\Delta_5, \Delta_3](< \Delta_7 < 1)$ and $p \in [p_3, \frac{1}{2} - \Delta]$

 ii. $\frac{\pi}{F} \in [1, \frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}]$, $\Delta \in [\Delta_3, \frac{1}{2}](< \Delta_7 < 1)$ and $p \in [0, \frac{1}{2} - \Delta]$

 iii. $\frac{\pi}{F} \in [\frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}, \frac{1}{2\beta}]$, $\Delta \in [\Delta_3, \frac{1}{2}](< 1 < \Delta_7)$ and $p \in [0, \frac{1}{2} - \Delta]$

 iv. $\frac{\pi}{F} \in [\frac{\delta+\beta-\beta\delta+1}{2(\delta+\beta-\beta\delta)}, \frac{1}{2\beta}]$, $[\Delta_5, \Delta_3](< \frac{1}{2} < 1 < \Delta_7)$ and $p \in [p_3, \frac{1}{2} - \Delta]$

Taking into consideration Corollary 5 we get the following result, which identifies the parameters' values for which (DESTROY, DESTROY) is the Pareto dominant SPE for $p \in [0, \frac{1}{2} - \Delta]$.

From Lemmata 26, 27 and 22 we obtain the following proposition.

Proposition 2. *For given policy and other parameters $(p, \Delta, F, \beta, \pi)$ the Pareto dominant SPE of the extended game with LP is:*

1. (DESTROY, DESTROY) *iff $\beta < \frac{1}{2}, \frac{\pi}{F} \in [1, \frac{1}{2\beta}]$ and*

(a) $\Delta \in [\Delta_5, \Delta_3]$ *and* $p \in [p_3, \frac{1}{2} - \Delta]$ *or*

(b) $[\Delta_3, \frac{1}{2})$ *and* $p \in [0, \frac{1}{2} - \Delta]$ *or*

(c) $\Delta \in [\Delta_5, \Delta_6]$ *and* $p \in [\max\{0, \frac{1}{2} - \Delta\}, p_4]$ *or*

(d) $\Delta > \Delta_6$ *and* $p \in [0, 1 - \Delta]$

2. (KEEP AND NOT REPORT, KEEP AND NOT REPORT) *iff:*

(a) $\beta < \frac{1}{2}$ *and*

 i. $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta < \Delta_5$ *and* $p \in [0, \frac{1}{2} - \Delta]$ *or*

 ii. $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta \in [\Delta_5, \Delta_3]$ *and* $p \in [0, p_3)$ *or*

 iii. $\frac{\pi}{F} > \frac{1}{2\beta}$, $\Delta < \frac{1}{2}$ *and* $p \in [0, \frac{1}{2} - \Delta]$ *or*

(b) $\beta > \frac{1}{2}, \frac{\pi}{F} \geq 1$, $\Delta < \frac{1}{2}$ *and* $p \in [0, \frac{1}{2} - \Delta]$

3. *(*Keep and Report, Keep and Report*)* *iff:*

(a) $\beta < \frac{1}{2}$ *and*

    i. $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta < \Delta_5$ *and* $p \in [\frac{1}{2} - \Delta, 1 - \Delta]$ *or*

    ii. $\frac{\pi}{F} \in [1, \frac{1}{2\beta}]$, $\Delta \in [\Delta_5, \Delta_6]$ *and* $p \in (p_4, 1 - \Delta]$ *or*

    iii. $\frac{\pi}{F} > \frac{1}{2\beta}$ *and* $p \in [\max\{0, \frac{1}{2} - \Delta\}, 1 - \Delta]$ *or*

(b) $\beta > \frac{1}{2}$, $\frac{\pi}{F} \geq 1$ *and* $p \in [\max\{0, \frac{1}{2} - \Delta\}, 1 - \Delta]$

where $p_3 = \frac{\pi}{F} - \frac{\Delta}{\beta}$, $p_4 = \frac{F - 2\beta\Pi_M}{2(1-\beta)F}$, $\Delta_3 = \beta\frac{\pi}{F}$, $\Delta_5 = \frac{\beta}{1-\beta}\frac{\pi}{F} - \frac{\beta}{2(1-\beta)}$ *and* $\Delta_6 = \frac{1-2\beta}{2(1-\beta)} + \frac{\beta}{1-\beta}\frac{\pi}{F}$.

Appendix C

$\frac{\partial \tilde{p}}{\partial e} = \frac{1-p}{(1+e)^2} > 0$

$\frac{\partial^2 \tilde{p}}{\partial e^2} = -\frac{2(1-p)}{(1+e)^3} < 0$

$\frac{\partial \tilde{p}}{\partial p} = \frac{1}{1+e} > 0$

$\frac{\partial^2 \tilde{p}}{\partial e \partial p} = -\frac{1}{(1+e)^2} < 0$

$\frac{\partial e_{0,G}}{\partial L} = \frac{1}{2} \frac{\delta(1-p)}{\sqrt{\frac{\delta(1-p)L}{1-\delta}}(1-\delta)} > 0$

$\frac{\partial e_{0,G}}{\partial p} = -\frac{1}{2} \frac{\delta L}{\sqrt{\frac{\delta(1-p)L}{1-\delta}}(1-\delta)} > 0$

$\frac{\partial \rho_1}{\partial L} = -\frac{(1-p)^2[\delta(1-a)+a](1-\delta)}{a[\delta(1-p)(1-\delta)L]^{\frac{3}{2}}} < 0$

$\frac{\partial \rho_1}{\partial p} = -\frac{\delta L(1-p)(1-\delta)[\delta(1-a)+a]}{a[\delta(1-p)(1-\delta)L]^{\frac{3}{2}}} < 0$

$\frac{\partial \rho_1}{\partial a} = -\frac{2(1-p)\delta}{a^2\sqrt{\delta(1-p)(1-\delta)L}} < 0$

$\frac{\partial \rho_2}{\partial L} = -\frac{(1-p)^2\delta(1-\delta)\left[\delta^2(1-a)-a\right])}{a(1-\delta)[\delta(1-p)(1-\delta)L]^{\frac{3}{2}}} < 0$

$\frac{\partial \rho_2}{\partial p} = -\frac{(1-p)(1-\delta)\delta L\left[\delta^2(1-a)-a\right])}{a(1-\delta)[\delta(1-p)(1-\delta)L]^{\frac{3}{2}}} < 0$

$\frac{\partial \rho_2}{\partial a} = -\frac{2(1-p)\delta^2}{a^2(1+\delta)\sqrt{\delta(1-p)(1-\delta)L}} < 0$

$\rho_1 > \rho_2$ since $\frac{-2(1-p)(\delta+2a)}{a(1+\delta)\sqrt{\delta(1-p)(1-\delta)L}} < 0$

$\frac{\partial \bar{\psi}}{\partial \rho} = -\frac{1}{2} \frac{(1-\delta)a\sqrt{\delta(1-\delta)(1-p)L}}{[\delta(1-a)+a]\left[\sqrt{\delta(1-\delta)(1-p)L}-(1-p)(1-\delta)\right]} < 0$

$\frac{\partial \bar{\psi}}{\partial a} = -\frac{1}{2} \frac{(1-\delta)\rho\delta\sqrt{\delta(1-\delta)(1-p)L}}{\sqrt{\delta(1-\delta)(1-p)L}[\delta(1-a)+a]^2(1-p)(1-\delta)} < 0$

$\frac{\partial \bar{\psi}}{\partial p} = -\frac{1}{4} \frac{\delta(1-p)[2\delta+(1-\delta)(2-\rho)a](1-\delta)^2 L}{\sqrt{\delta(1-\delta)(1-p)L}\left[(1-p)(1-\delta)-\sqrt{\delta(1-\delta)(1-p)L}\right]^2[\delta(1-a)+a]} < 0$

$\frac{\partial \bar{\psi}}{\partial L} = -\frac{1}{4} \frac{\delta(1-p)^2[2\delta+(1-\delta)(2-\rho)a](1-\delta)^2}{\sqrt{\delta(1-\delta)(1-p)L}\left[(1-p)(1-\delta)-\sqrt{\delta(1-\delta)(1-p)L}\right]^2[\delta(1-a)+a]} < 0$

$\frac{\partial \bar{\psi}}{\partial \delta} = \frac{1}{2} \frac{\delta La\rho(1-p)(1-\delta)^{\frac{3}{2}}+\frac{1}{2}[(1-\delta)^2(2-\rho)a^2+(1-\delta)\delta(\rho+4)a+2\delta^2](1-\delta)(1-p)^2 L}{\sqrt{\delta(1-\delta)(1-p)L}\left[(1-p)(1-\delta)-\sqrt{\delta(1-\delta)(1-p)L}\right]^2[(1-a)\delta+a]^2} > 0$

Appendix D
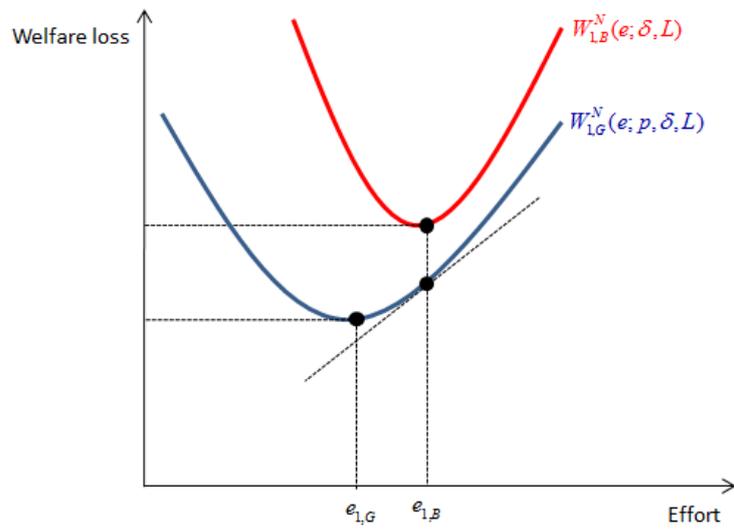
Figure 14: Optmal effort levels

Appendix E

- $\frac{\partial a^*}{\partial k} = -\frac{1}{3k} < 0$

- $\frac{\partial^2 a^*}{\partial k^2} = \frac{1}{3k^2} > 0$

- $\frac{\partial a^*}{\partial F} = \frac{1}{3F} > 0$

- $\frac{\partial^2 a^*}{\partial F^2} = -\frac{1}{3F^2} < 0$

- $\frac{\partial a^*}{\partial p} = \frac{1}{3p} > 0$

- $\frac{\partial^2 a^*}{\partial p^2} = -\frac{1}{3p^2} < 0$

- $\frac{\partial \Pi_N}{\partial p} = \frac{2}{3} \frac{\delta k^{\frac{2}{3}} F^{\frac{1}{3}}}{p^{\frac{2}{3}} (2\delta-1)} > 0$

- $\frac{\partial \Pi_N^2}{\partial p^2} = -\frac{4}{9} \frac{\delta k^{\frac{2}{3}} F^{\frac{1}{3}}}{p^{\frac{5}{3}} (2\delta-1)} < 0$

- $\frac{\partial \Pi_N}{\partial F} = \frac{2}{3} \frac{\delta k^{\frac{2}{3}} p^{\frac{1}{3}}}{F^{\frac{2}{3}} (2\delta-1)} > 0$

- $\frac{\partial \Pi_N^2}{\partial F^2} = -\frac{4}{9} \frac{\delta k^{\frac{2}{3}} p^{\frac{1}{3}}}{F^{\frac{5}{3}} (2\delta-1)} < 0$

- $\frac{\partial \Pi_N}{\partial k} = \frac{4}{3} \frac{\delta p^{\frac{1}{3}} F^{\frac{1}{3}}}{k^{\frac{1}{3}} (2\delta-1)} > 0$

- $\frac{\partial \Pi_N}{\partial k^2} = -\frac{4}{9} \frac{\delta p^{\frac{1}{3}} F^{\frac{1}{3}}}{k^{\frac{4}{3}} (2\delta-1)} < 0$

- $\frac{\partial \theta}{\partial p} = -\frac{2}{3} \left(\frac{k}{pF} \right)^{\frac{2}{3}} < 0$

- $\frac{\partial^2 \theta}{\partial p^2} = \frac{4}{9p} \left(\frac{k}{pF} \right)^{\frac{2}{3}} > 0$

- $\frac{\partial \theta}{\partial k} = -\frac{4}{3F^{\frac{1}{3}}} \left(\frac{p}{kF} \right)^{\frac{1}{3}} < 0$

- $\frac{\partial^2 \theta}{\partial k^2} = \frac{4}{9} \frac{p^{\frac{1}{3}}}{k^{\frac{4}{3}} F^{\frac{2}{3}}} > 0$

- $\frac{\partial \theta}{\partial F} = \frac{4}{3} \frac{p^{\frac{1}{3}} k^{\frac{2}{3}}}{F^{\frac{5}{3}}} > 0$

- $\frac{\partial^2 \theta}{\partial F^2} = -\frac{20}{9} \frac{p^{\frac{1}{3}} k^{\frac{2}{3}}}{F^{\frac{8}{3}}} < 0$

186

- $\frac{\partial \hat{\theta}}{\partial k} = -\frac{8}{3} \frac{p^{\frac{1}{3}}}{(2-\delta)k^{\frac{1}{3}}F^{\frac{2}{3}}} < 0$

- $\frac{\partial^2 \hat{\theta}}{\partial k^2} = \frac{8}{9} \frac{p^{\frac{1}{3}}}{(2-\delta)k^{\frac{4}{3}}F^{\frac{2}{3}}} > 0$

- $\frac{\partial \hat{\theta}}{\partial F} = \frac{8}{3} \frac{p^{\frac{1}{3}}k^{\frac{2}{3}}}{(2-\delta)F^{\frac{5}{3}}} > 0$

- $\frac{\partial^2 \hat{\theta}}{\partial F^2} = -\frac{40}{9} \frac{p^{\frac{1}{3}}k^{\frac{2}{3}}}{(2-\delta)F^{\frac{8}{3}}} < 0$

- $\frac{\partial \hat{\theta}}{\partial p} = -\frac{4}{3} \frac{k^{\frac{2}{3}}}{(2-\delta)p^{\frac{2}{3}}F^{\frac{2}{3}}} < 0$

- $\frac{\partial^2 \hat{\theta}}{\partial p^2} = \frac{8}{9} \frac{k^{\frac{2}{3}}}{(2-\delta)p^{\frac{5}{3}}F^{\frac{2}{3}}} > 0$

- $\bar{\theta} > \hat{\theta}$

$$\begin{aligned}
\bar{\theta} &> \hat{\theta} \\
2\left(1 - \frac{2p^{\frac{1}{3}}k^{\frac{2}{3}}}{F^{\frac{2}{3}}}\right) &> \frac{2\left(F - 2p^{\frac{1}{3}}F^{\frac{1}{3}}k^{\frac{2}{3}}\right)}{(2-\delta)F} \\
\frac{2(1-\delta)(p^{\frac{2}{3}}F^{\frac{5}{3}}k^{\frac{1}{3}} - 2pkF)}{p^{\frac{2}{3}}F^{\frac{2}{3}}k^{\frac{1}{3}}(2-\delta)} &> 0 \text{ if } F > 2\sqrt{2pk}
\end{aligned}$$

- $\Pi_N < \hat{\Pi}(\hat{\theta})$

$$\begin{aligned}
\hat{\Pi}(\hat{\theta}) &> \Pi_N \\
\frac{\delta\left[F(1-\delta) + 2p^{\frac{1}{3}}F^{\frac{1}{3}}k^{\frac{2}{3}}\right]}{5\delta - 2(1+\delta^2)} &> \frac{2\delta p^{\frac{1}{3}}F^{\frac{1}{3}}k^{\frac{2}{3}}}{2\delta - 1} \\
\frac{\delta(1-\delta)(F - 2p^{\frac{1}{3}}F^{\frac{1}{3}}k^{\frac{2}{3}})}{2(2-\delta)(\delta - \frac{1}{2})} &> 0 \text{ if } F > 2\sqrt{2pk}
\end{aligned}$$

- $V_N > V_R(\hat{\theta})$

$$\begin{aligned}
V_N &> V_R(\hat{\theta}) \\
\frac{\Pi - 2p^{\frac{1}{3}}F^{\frac{1}{3}}k^{\frac{2}{3}}}{1-\delta} &> \frac{2\Pi - F - (\Pi - F)\delta - 2k^{\frac{2}{3}}p^{\frac{1}{3}}F^{\frac{1}{3}}}{(2-\delta)(1-\delta)} \\
\frac{F^{\frac{1}{3}}(p^{\frac{2}{3}}F^{\frac{2}{3}}k^{\frac{1}{3}} - 2pk)}{(2-\delta)k^{\frac{1}{3}}p^{\frac{2}{3}}} &> 0 \text{ if } F > 2\sqrt{2pk}
\end{aligned}$$

- $\hat{\theta} > \underline{\theta}$

$$\begin{aligned}
\hat{\theta} &> \underline{\theta} \\
\frac{2\left(F - 2p^{\frac{1}{3}}F^{\frac{1}{3}}k^{\frac{2}{3}}\right)}{(2-\delta)F} &> 1 - 3\sqrt{\frac{pk}{2F}} \\
F &> \frac{2\sqrt{2}k}{p}
\end{aligned}$$