

# When Can We Expect a Corporate Leniency Program to Result in Fewer Cartels?\*

Joseph E. Harrington, Jr.  
Dept. of Bus. Econ. & Public Policy  
The Wharton School  
University of Pennsylvania  
Philadelphia, PA 19102  
harrij@wharton.upenn.edu

Myong-Hun Chang  
Department of Economics  
Cleveland State University  
Cleveland, OH 44115  
m.chang@csuohio.edu

25 November 2013

## Abstract

Leniency programs have become widespread and are generally quite active as reflected in the number of applications. What is not well-understood is how they affect the number of cartels. This paper develops and explores a theoretical framework to help understand when leniency programs are likely to be effective in reducing the presence of cartels. Plausible conditions are derived whereby a leniency program can result in more cartels. On a more positive note, we identify situations and policies that a competition authority can pursue that will make it more likely that a leniency program will have the intended effect of reducing the number of cartels.

---

\*For their valuable comments on a previous version of this paper, we thank Louis Kaplow, Kai-Uwe Kühn, Yossi Spiegel, Juuso Välimäki, participants of the 15th WZB Conference on Markets and Politics/2nd Conference of the Research Network on Innovation and Competition Policy, Helsinki Center of Economic Research Conference on Cartels and Collusion, Tercer Taller de Organización Industrial (Zapallar, Chile), Third Annual Searle Research Symposium on Antitrust Economics and Competition Policy (Northwestern University) and seminar participants at George Washington University, Lafayette College, U. of Texas at Dallas, U. of North Carolina at Chapel Hill, Northwestern University-KGSM, Universitat Pompeu Fabra, CEMFI, Universität Düsseldorf (DICE), VU Amsterdam & The Tinbergen Institute, and Harvard Law School. The first author gratefully acknowledges the financial support of the National Science Foundation (SES-0516943, SES-1148129) and Santander through a Cátedras de Excelencia. He also expresses his gratitude to the faculty of the Departamento de Economía of the Universidad Carlos III de Madrid for providing a most collegial and stimulating environment.

# 1 Introduction

The Corporate Leniency Program of the U.S. Department of Justice's Antitrust Division gives a member of a cartel the opportunity to avoid government penalties if it is the first to report the cartel and fully cooperate. Since its revision in 1993, the Program has been flush with applications. Deputy Assistant Attorney General Scott Hammond noted in 2005 that "the revised Corporate Leniency Program has resulted in a surge in amnesty applications. Under the new policy, the application rate has jumped to roughly two per month."<sup>1</sup> Furthermore, he commented: "The extraordinary success of the Division's leniency program has generated widespread interest around the world."<sup>2</sup> That is indeed the case as the steady flow of leniency applications in the U.S. led the European Commission (EC) to institute its own leniency program in 1996 and a decade later 24 out of 27 EU members had one. Globally, leniency programs are now present in more than 50 countries and jurisdictions.<sup>3</sup>

In many of these countries, leniency programs have similarly been active and viewed as successful. In South Africa, which put in place its program in 2004, the number of applicants were flowing in at a rate of about three per month by 2009-10, which even exceeded that in the U.S. A week prior to Spain's institution of its leniency program in June 2008, cartelists were literally lining up outside the doors of the Comisión Nacional de la Competencia's offices in order to be the first from their cartel to apply for leniency. In Germany, the Cartel Office noted: "The first version of the Leniency Programme was already a success. This can be seen by the number of leniency applications filed: Between 2000 and 2005 a total of 122 leniency applications were filed."<sup>4</sup>

It is clear that many leniency programs have sparked numerous applications. It is also clear that one can identify specific cases for which a leniency program was responsible for the discovery of the cartel and was instrumental in the successful prosecution of the cartel. What is far less clear, however, is whether leniency programs have been successful in the sense that these economies are populated by fewer cartels. Ultimately, success is to be measured by a smaller number of cartels, not a greater number of leniency applications.

In light of the widespread adoption and usage of leniency programs, there is a vast and growing body of scholarly work intended to examine the effect of these programs.<sup>5</sup> Starting with the pioneering paper of Motta and Polo (2003), theoretical analyses include Ellis and Wilson (2001), Spagnolo (2003), Motchenkova (2004), Aubert, Kovacic, and Rey (2006), Chen and Harrington (2007), Chen and Rey (2007), Harrington (2008), Harrington and Chang (2009), Houba, Motchenkova, and Wen (2009), Silbye

---

<sup>1</sup>Scott D. Hammond, "An Update of the Antitrust Division's Criminal Enforcement Program," ABA Section of Antitrust Law, Cartel Enforcement Roundtable, November 16, 2005; p. 10.

<sup>2</sup>Ibid, p.10.

<sup>3</sup>For a list of countries with leniency programs, see Borrell, Jiménez, and García (2012) who also estimate how leniency programs have changed the perceptions of managers.

<sup>4</sup>"Effective Cartel Prosecution: Benefits for the Economy and Consumers," Bundeskartellamt, July 2010, p. 19.

<sup>5</sup>A review of some of the early work is provided in Spagnolo (2008).

(2010), Bos and Wandschneider (2011), Choi and Gerlach (2012), Lefouili and Roux (2012), Sauvagnat (2012), Gärtner (2013), and Marshall, Marx, and Mezzetti (2013). The general conclusion of this body of work is that leniency programs make collusion more difficult.

A common feature to all of these models is the assumption that the introduction of a leniency program does not impact enforcement through non-leniency means. Non-leniency enforcement is modelled as the probability that a cartel is discovered, prosecuted, and convicted in the absence of a member having entered the leniency program. As a cartel member will apply for leniency only if it believes that doing so is better than running the risk of being caught and convicted, non-leniency enforcement is integral to inducing firms to apply for leniency. If this probability is low then few cartel members will use the leniency program, while if the probability is high then cartel members will race to apply for leniency. The impact of a leniency program is then intrinsically tied to the level of non-leniency enforcement.

Next note that it is natural to expect that the introduction of a leniency program will affect the level of non-leniency enforcement. A leniency program may cause a competition authority's scarce resources and attention to shift from non-leniency cases to leniency cases. However, this doesn't necessarily imply that non-leniency enforcement is weaker. If a leniency program is successful in resulting in fewer cartels, there will be fewer non-leniency cartel cases, in which case the authority may still have ample resources to effectively prosecute them. Furthermore, a welfare-maximizing competition authority can adjust its enforcement policy in response to what is occurring with leniency applications. Thus, while we expect non-leniency enforcement to change when a leniency program is put in place, it isn't clear whether it'll be weakened or strengthened.

The objective of the current paper is to develop and explore a theoretical framework to understand when leniency programs are likely to be effective in reducing the presence of cartels. Its primary innovation is in providing a more comprehensive assessment of how a leniency program affects the activity and efficacy of a competition authority by taking account of its impact on the entire portfolio of cases; both those generated through leniency applications and through more traditional methods. Our model is the first to examine the effect of a leniency program while endogenizing non-leniency enforcement and, as we'll see, doing so requires modelling innovations and overcoming some technical hurdles.

Contrary to existing results in the literature and the general impression of practitioners, we find that a leniency program can result in *more* cartels, and this can occur at the same time that a leniency program is generating many applications. On a more positive note, we also identify situations and policies that a competition authority can pursue that will make it more likely that a leniency program will have the intended effect of reducing the number of cartels.

Before moving on, it is useful to note that, in addition to a burgeoning theoretical literature on leniency programs, there is a growing body of experimental work. Research here includes Apestegua, Dufwenberg, and Selten (2007), Hinlopen and Soetevent (2008), Hamaguchia, Kawagoeb, and Shibatac (2009), Dijkstra, Haan, and

Schoonbeek (2011), and Bigoni, Fridolfsson, Le Coq, and Spagnolo (2012). These experimental studies generally find that a leniency program reduces cartel formation though some studies also find that prices are higher, conditional on a cartel forming, when there is a leniency program. Finally, there is an increasing number of empirical studies that measure the impact of leniency programs but are decidedly mixed and tentative in their findings. Miller (2009) examines the impact of the leniency program in the U.S., Choi (2009) considers the program in Korea, and Stephan (2008), Brenner (2009), Klein (2010), and Zhou (2011) investigate the impact of the European Commission’s leniency program.

In the next section, the model is presented. In Section 3, the conditions determining the equilibrium cartel rate are derived. The impact of a leniency program on the cartel rate when non-leniency enforcement is fixed is examined in Section 4. While those results are of intrinsic interest as a benchmark, they are primarily an intermediate step towards endogenizing non-leniency enforcement. Section 5 delivers the main contribution of the paper which is a characterization of how a leniency program impacts the cartel rate when non-leniency enforcement is allowed to adjust to a competition authority experiencing a flow of leniency applications. Section 6 concludes. Proofs are in an appendix.

## 2 Model

The modelling strategy is to construct a birth-and-death process for cartels in order to generate an average cartel rate for a population of industries, and then to assess how the introduction of a leniency program influences that cartel rate. We build on the birth-and-death process developed in Harrington and Chang (2009) by introducing a leniency program, endogenizing the intensity of non-leniency enforcement, and allowing a competition authority (CA) to decide on its caseload. In this manner, the full effect of a leniency program can be assessed.<sup>6</sup>

### 2.1 Industry Environment

Firm behavior is modelled using a modification of a Prisoners’ Dilemma formulation. Firms simultaneously decide whether to *collude* (set a high price) or *compete* (set a low price). Prior to making that choice, firms observe a stochastic realization of the market’s profitability that is summarized by the variable  $\pi \geq 0$ .<sup>7</sup> If all firms choose *collude* then each firm earns  $\pi$ , while if all choose *compete* then each earns  $\alpha\pi$  where  $\alpha \in [0, 1)$ .  $1 - \alpha$  then measures the competitiveness of the non-collusive environment.  $\pi$  has a continuously differentiable cdf  $H : [\underline{\pi}, \bar{\pi}] \rightarrow [0, 1]$  where  $0 < \underline{\pi} < \bar{\pi}$ .  $h(\cdot)$  denotes the associated density function and let  $\mu \equiv \int \pi h(\pi) d\pi$  denote its finite

---

<sup>6</sup>While a leniency program is briefly considered in Harrington and Chang (2009), that model assumes - like the remainder of the literature - that non-leniency enforcement is fixed. As hopefully this paper will convince the reader, it is a technically and economically substantive extension of the Harrington-Chang model to endogenize non-leniency enforcement.

<sup>7</sup>The informational setting is as in Rotemberg and Saloner (1986).

mean. If all other firms choose *collude*, the profit a firm earns by deviating - choosing *compete* - is  $\eta\pi$  where  $\eta > 1$ . This information is summarized in the table below.

Own action	All other firms' actions	Own profit
<i>collude</i>	<i>collude</i>	$\pi$
<i>compete</i>	<i>collude</i>	$\eta\pi$
<i>compete</i>	<i>compete</i>	$\alpha\pi$

Note that the Bertrand price game is represented by  $(\alpha, \eta) = (0, n)$  where  $n$  is the number of firms. The Cournot quantity game with linear demand and cost functions in which firms collude at the joint profit maximum is represented by  $(\alpha, \eta) = \left(\frac{4n}{(n+1)^2}, \frac{(n+1)^2}{4n}\right)$ .<sup>8</sup>

Firms interact in an infinite horizon setting where  $\delta \in (0, 1)$  is the common discount factor. It is not a repeated game because, as explained later, each industry is in one of two states: cartel and non-cartel. If *firms are a cartel* then they have the *opportunity to collude* but do so if and only if (iff) it is incentive compatible. More specifically, if firms are cartelized then they simultaneously choose between *collude* and *compete*, and, at the same time, whether or not to apply to the corporate leniency program. Details on the description of the leniency program are provided later. If it is incentive compatible for all firms to choose *collude* then each earns  $\pi$ . If instead a firm prefers *compete* when all other firms choose *collude* then collusion is not incentive compatible (that is, it is not part of the subgame perfect equilibrium for the infinite horizon game) and each firm earns  $\alpha\pi$ . In that case, collusion is not achieved. If *firms are not a cartel* then each firm earns  $\alpha\pi$  as, according to equilibrium, they all choose *compete*.

At the end of the period, there is the random event whereby the CA may pursue an investigation; this can only occur if, in the current period, the cartel was either active or shut down and no firm applied for leniency. Let  $\sigma \in [0, 1]$  denote the probability that firms are discovered, prosecuted, and convicted (below, we will endogenize  $\sigma$  though, from the perspective of an individual industry, it is exogenous). In that event, each firm incurs a penalty of  $F$ .

It is desirable to allow  $F$  to depend on the extent of collusion. Given there is only one level of collusion in the model, the "extent of collusion" necessarily refers to the number of periods that firms had colluded. A proper accounting of that effect would require that each cartel have a state variable equal to the length of time for which it has been active; such an extension would seriously complicate the analysis. As a simplifying approximation, it is instead assumed that the penalty is proportional to the average increase in profit from being cartelized (rather than the *realized* increase in profit). If  $Y$  denotes the expected per period profit from being in the "cartel state" then  $F = \gamma(Y - \alpha\mu)$  where  $\gamma > 0$  and  $\alpha\mu$  is average non-collusive profit. This specification avoids the need for state variables but still allows the penalty to

---

<sup>8</sup>We have only specified a firm's profit when all firms choose *compete*, all firms choose *collude*, and it chooses *compete* and all others firm choose *collude*. We must also assume that *compete* strictly dominates *collude* for the stage game. It is unnecessary to provide any further specification.

be sensitive to the (average) extent of collusion.<sup>9</sup> As the CA will not be presumed to manipulate  $F$ , one can suppose that penalties are already set at their maximum level.

In addition to being discovered by the CA, a cartel can be uncovered because one of its members comes forth under the corporate leniency program. Suppose a cartel is in place. If a single firm applies for leniency then all firms are convicted for sure and the firm that applied receives a penalty of  $\theta F$  where  $\theta \in [0, 1)$ , while the other cartel members each pay  $F$ . If all firms simultaneously apply for leniency then each firm pays a penalty of  $\omega F$  where  $\omega \in (\theta, 1)$ . For example, if only one firm can receive leniency and each firm has an equal probability of being first in the door then  $\omega = \frac{n-1+\theta}{n}$  when there are  $n$  cartel members. It is sufficient for the ensuing analysis that we specify the leniency program when either one firm applies or all firms apply. Also, leniency is not awarded to firms that apply after another firm has done so.

From the perspective of firms, competition policy is summarized by the four-tuple  $(\sigma, \gamma, \theta, \omega)$  which are, respectively, the probability of paying penalties through non-lenieny enforcement, the penalty multiple, the leniency parameter when only one firm applies (where  $1 - \theta$  is the proportion of fines waived), and the leniency parameter when all firms apply (where  $1 - \omega$  is the proportion of fines waived).

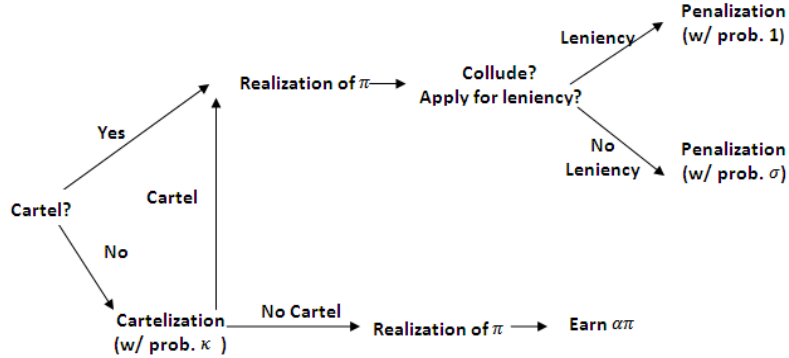
Next, let us describe how an industry's cartel status evolves. Suppose it enters the period cartelized. The industry will exit the period still being a cartel if: 1) all firms chose *collude* (which requires that collusion be incentive compatible); 2) no firm applied for leniency; and 3) the CA did not discover and convict the firms of collusion. Otherwise, the cartel collapses and firms revert to the "no cartel" state. If instead the industry entered the period in the "no cartel" state then with probability  $\kappa \in (0, 1)$  firms cartelize. For that cartel to still be around at the end of the period, conditions (1)-(3) above must be satisfied. Note that whenever a cartel is shutdown - whether due to internal collapse, applying to the leniency program, or having been successfully prosecuted - the industry may re-cartelize in the future. Specifically, it has an opportunity to do so with probability  $\kappa$  in each period that it is not currently colluding. The timing of events is summarized in the figure below.

In modelling a population of industries, it is compelling to allow industries to vary in terms of cartel stability. For this purpose, industries are assumed to differ in the parameter  $\eta$ . If one takes this assumption literally, it can be motivated by heterogeneity in the elasticity of firm demand or the number of firms (as with the Bertrand price game). Our intent is not to be literal but rather to think of this as a parsimonious way in which to encompass industry heterogeneity. Let the cdf on industry types be represented by the continuously differentiable strictly increasing function  $G : [\underline{\eta}, \bar{\eta}] \rightarrow [0, 1]$  where  $1 < \underline{\eta} < \bar{\eta}$ .  $g(\cdot)$  denotes the associated density function. The appeal of  $\eta$  is that it is a parameter which influences the frequency of collusion but does not directly affect the value of the firm's profit stream since, in

---

<sup>9</sup>A more standard assumption in the literature is to assume  $F$  is fixed which is certainly simpler but less realistic than our specification. All of our qualitative results hold when  $F$  is fixed.

equilibrium, firms do not cheat; this property simplifies the analysis.



## 2.2 Enforcement Technology

Non-leniency enforcement is represented by  $\sigma$  which is the probability that a cartel pays penalties without one of its members having entered the leniency program. Here, we explain how  $\sigma$  is determined.  $\sigma$  is the compound probability that: 1) the cartel is discovered by the CA; 2) the CA decides to investigate the cartel; and 3) the CA is successful in its investigation and penalties are levied. The initial discovery of a cartel is presumed to be exogenous and to come from customers, uninvolved employees, the accidental discovery of evidence through a proposed merger, and so forth.  $q \in [0, 1]$  denotes the probability of discovery and is a parameter throughout the paper. What the CA controls is how many cases to take on which is represented by  $r \in [0, 1]$  which is the fraction of reported cases that the CA chooses to investigate. Initially, we will derive results when  $r$  is fixed and then allow  $r$  to be endogenous. Finally, of those cases discovered and investigated, the CA is successful in a fraction  $s \in [0, 1]$  of them where  $s$  is determined by the relationship between the CA's resources and its caseload.

The CA is faced with a resource constraint: the more cases it takes on, the fewer resources are applied to each case and the lower is the probability of winning any individual case. More specifically, it is assumed  $s = p(\lambda L + R)$  where  $\lambda \in (0, 1]$ .  $L$  is the fraction of industries that are involved in leniency cases,  $R$  is the fraction of industries that are involved in non-leniency cases, and  $s$  is the proportion of  $R$  cases that result in a conviction. Leniency cases are assumed to be won for sure.  $\lambda \leq 1$  because leniency cases may take up fewer resources than those cases lacking an informant. We will refer to  $L + R$  as the number of cases and  $\lambda L + R$  as the caseload.  $p : [0, 1] \rightarrow [0, 1]$  is a continuous decreasing function so that a bigger caseload means a lower probability of winning a non-leniency case. In sum, the probability that a cartel pays penalties is  $\sigma = q \times r \times s = q \times r \times p(\lambda L + R)$ .  $\sigma$  is endogenous because  $s$  is determined by the caseload which depends on the number of cartels, and  $r$  may be chosen by the CA.<sup>10</sup>

<sup>10</sup>It should be noted that Motta and Polo (2003) do allow for optimal enforcement expenditure by

Key to the analysis is the implicit assumption that the CA faces a resource constraint in the sense that resources per case decline with the number of cases as reflected in the specification that the probability of any investigation being successful is decreasing in the caseload. In practice, an CA can move around resources to handle additional cartel activity by, for example, shifting lawyers and economists from merger cases to cartel cases. However, there is a rising opportunity cost in doing so and that ought to imply that resources per cartel case will decline with the number of cartel cases. Of course, CA officials can lobby their superiors (either higher level bureaucrats or elected officials) for a bigger budget but, at least in the U.S., the reality is that the CA's budget does not scale up with its caseload. While the budget of the Antitrust Division of the U.S. Department of Justice is increasing in GDP (Kwoka, 1999), DOJ antitrust case activity is actually countercyclical (Ghosal and Gallo, 2001).

The last element to specify is the determination of the fraction of cases that the CA takes on,  $r$ , which requires specifying an objective to the CA. As a benchmark, it is assumed here that the CA is welfare-maximizing in that it chooses  $r$  to minimize the cartel rate. The results in Section 5 actually are true either holding  $r$  fixed or with this objective assigned to the CA. Some numerical analysis is also conducted and there it is assumed  $r$  is selected to minimize the cartel rate because it eliminates having to specify one more parameter value. We view the assumption that the CA minimizes the cartel rate to be a useful benchmark for gaining insight into the possible implications of a leniency program though not necessarily as a good description of CA behavior. However a CA is rewarded, it is natural to assume that rewards are based on observable measures of performance. Given that the cartel rate is not observable (only discovered cartels are observable) then presumably the CA is not rewarded based on how its policies affect the cartel rate (at least not directly). Thus, it is not clear that there is an incentive scheme that will induce the CA to minimize the cartel rate. We intend to explore the modelling of the CA in future research.<sup>11</sup>

### 3 Equilibrium Conditions

In this section, we describe the conditions determining the equilibrium frequency with which industries are cartelized. Given that there are several steps in the construction of equilibrium conditions, it may prove beneficial to the reader to provide an overview.

1. Taking as given  $\sigma$  (the per period probability that a cartel pays penalties through non-lenieny enforcement), we first solve for equilibrium collusive be-

---

modelling a trade-off between monitoring and prosecution. They endow a CA with a fixed amount of resources that can be allocated between finding suspected episodes of collusion and prosecuting the cases that are found or, in the language of our model, between raising  $q$  and lowering  $s$  (assuming  $r = 1$ ). However, the model is very different from ours - for example, they do not consider a population of industries and do not solve for the steady-state frequency of cartels - and it does not address the questions we are raising here.

<sup>11</sup>In Harrington (2011), sufficient conditions are derived for the optimal policy of a CA, whose objective is to maximize the number of (observable) convictions, to coincide with the policy that minimizes the cartel rate.



havior for a type- $\eta$  industry and the maximum value for  $\pi$  whereby collusion is incentive compatible, denoted  $\phi^*(\sigma, \eta)$ .

2. With the conditions for internal collapse - which occurs when  $\pi > \phi^*(\sigma, \eta)$  - and the likelihood of non-leniency enforcement,  $\sigma$ , along with the probability of cartel formation,  $\kappa$ , a Markov process on cartel birth and death is constructed from which is solved the stationary distribution of industries in terms of their cartel status, for each industry type  $\eta$ . By aggregating over all industry types, the equilibrium cartel rate,  $C(\sigma)$ , is derived, given  $\sigma$ .
3. The next step is to solve for the equilibrium value of  $\sigma$ , denoted  $\sigma^*$ . The probability that the CA's investigation is successful,  $p(\lambda L + R)$ , depends on the mass of leniency cases,  $L$ , and the mass of non-leniency cases,  $R$ ; both  $L$  and  $R$  depend on  $\sigma$  as they depend on the cartel rate  $C(\sigma)$ .  $\sigma^*$  is then a fixed point:

$$\sigma^* = qrp(\lambda L(\sigma^*) + R(\sigma^*))$$

In other words,  $\sigma$  - the probability that firms are caught, prosecuted, and convicted - determines the cartel rate  $C(\sigma)$ , the cartel rate determines the caseload  $\lambda L(\sigma) + R(\sigma)$ , and the caseload determines the probability that they are able to get a conviction on a case and thus  $\sigma$ . Given  $\sigma^*$ , the equilibrium cartel rate is  $C(\sigma^*)$ .

4. When  $r$  is fixed, the analysis ends with step 3. When  $r$  is endogenous, the final step is to solve for the value that minimizes the cartel rate:

$$r^* \in \arg \min_{r \in [0,1]} C(\sigma^*(r)).$$

By way of comparison, the model in Harrington and Chang (2009) involved the two nested fixed point problems in steps 1 and 2 when  $\sigma$  is fixed. The current model embeds that problem in a third fixed point problem in order to endogenize  $\sigma$  (step 3). We believe this extension is both technically and economically substantive in that it introduces fundamentally new forces relevant to assessing the effect of leniency programs.

### 3.1 Cartel Formation and Collusive Value

A collusive strategy for a type- $\eta$  industry entails colluding when  $\pi$  is sufficiently low and not colluding otherwise. The logic is as in Rotemberg and Saloner (1986). When  $\pi$  is high, the incentive to deviate is strong because a firm increases current profit by  $(\eta - 1)\pi$ . At the same time, the future payoff is independent of the current realization of  $\pi$ , given that  $\pi$  is *iid*. Since the payoff to cheating is increasing in  $\pi$  while the future payoff is independent of  $\pi$ , the incentive compatibility of collusion is more problematic when  $\pi$  is higher.

Suppose firms are able to collude for at least some realizations of  $\pi$ , and let  $W^o$  and  $Y^o$  denote the payoff when the industry is not cartelized and is cartelized,

respectively. If not cartelized then, with probability  $\kappa$ , firms have an opportunity to cartelize with resulting payoff  $Y^o$ . With probability  $1 - \kappa$ , firms do not have such an opportunity and continue to compete. In that case, each firm earns current expected profit of  $\alpha\mu$  and a future value of  $W^o$ . Thus, the payoff when not colluding is defined recursively by:

$$W^o = (1 - \kappa) (\alpha\mu + \delta W^o) + \kappa Y^o. \quad (1)$$

As it'll be easier to work with re-scaled payoffs, define:

$$W \equiv (1 - \delta) W^o, \quad Y \equiv (1 - \delta) Y^o.$$

Multiplying both sides of (1) by  $1 - \delta$  and re-arranging yields:

$$W = \frac{(1 - \kappa) (1 - \delta) \alpha\mu + \kappa Y}{1 - \delta (1 - \kappa)}$$

Also note that the incremental value to being in the cartelized state is:

$$Y - W = Y - \left( \frac{(1 - \kappa) (1 - \delta) \alpha\mu + \kappa Y}{1 - \delta (1 - \kappa)} \right) = \frac{(1 - \kappa) (1 - \delta) (Y - \alpha\mu)}{1 - \delta (1 - \kappa)}. \quad (2)$$

Suppose firms are cartelized and  $\pi$  is realized. When a firm decides whether to collude or cheat, it decides at the same time whether to apply for leniency. If it decides to collude, it is clearly not optimal to apply for leniency since the cartel is going to be shut down by the authorities in which case the firm ought to maximize current profit by cheating. The more relevant issue is whether it should apply for leniency if it decides to cheat. The incentive compatibility constraint (ICC) is:

$$\begin{aligned} & (1 - \delta) \pi + \delta [(1 - \sigma) Y + \sigma W] - (1 - \delta) \sigma \gamma (Y - \alpha\mu) \\ & \geq (1 - \delta) \eta \pi + \delta W - (1 - \delta) \min \{ \sigma, \theta \} \gamma (Y - \alpha\mu). \end{aligned} \quad (3)$$

Examining the LHS expression, if it colludes then it earns current profit of  $\pi$  (given all other firms are colluding). With probability  $1 - \sigma$ , the cartel is not shut down by the CA and, given the industry is in the cartel state, the future payoff is  $Y$ . With probability  $\sigma$ , the cartel is caught and convicted by the CA - which means a one-time penalty of  $\gamma (Y - \alpha\mu)$  - and since the industry is no longer cartelized, the future payoff is  $W$ . Turning to the RHS expression, the current profit from cheating is  $\eta\pi$ . Since this defection causes the cartel to collapse, the future payoff is  $W$ . There is still a chance of being caught and convicted and a deviating firm will apply for leniency iff the penalty from doing so is less than the expected penalty from not doing so (and recall that the other firms are colluding and thus do not apply for leniency); that is, when  $\theta\gamma (Y - \alpha\mu) < \sigma\gamma (Y - \alpha\mu)$  or  $\theta < \sigma$ . Given optimal use of the leniency program, the deviating firm's expected penalty is then  $\min \{ \sigma, \theta \} \gamma (Y - \alpha\mu)$ . Re-arranging (3) and using (2), the ICC can be presented as:

$$\begin{aligned} \pi & \leq \frac{\delta (1 - \sigma) (1 - \kappa) (Y - \alpha\mu) - [1 - \delta (1 - \kappa)] [\sigma - \min \{ \sigma, \theta \}] \gamma (Y - \alpha\mu)}{(\eta - 1) [1 - \delta (1 - \kappa)]} \quad (4) \\ & \equiv \phi (Y, \sigma, \eta). \end{aligned}$$

Collusion is incentive compatible iff the current market condition is sufficiently low.<sup>12</sup>

In deriving an expression for the value to colluding, we need to discuss usage of the leniency program in equilibrium. Firms do not use it when market conditions result in the cartel being stable but may use it when the cartel collapses. As the continuation payoff is  $W$  regardless of whether leniency is used, a firm applies for leniency iff it reduces the expected penalty. First note that an equilibrium either has no firms applying for leniency or all firms doing so because if at least one firm applies then another firm can lower its expected penalty by also doing so. This has the implication that it is always an equilibrium for all firms to apply for leniency. Furthermore, it is the unique equilibrium when  $\theta < \sigma$ . To see why, suppose all firms were not to apply for leniency. A firm would then lower its penalty from  $\sigma F$  to  $\theta F$  by applying. When instead  $\sigma \leq \theta$ , there is also an equilibrium in which no firm goes for leniency as to do so would increase its expected penalty from  $\sigma F$  to  $\theta F$ . Using the selection criterion of Pareto dominance, we will assume that, upon internal collapse of the cartel, no firms apply when  $\sigma \leq \theta$  and all firms apply when  $\theta < \sigma$ .

The expected payoff to being cartelized,  $\psi(Y, \sigma, \eta)$ , is then recursively defined by:

$$\psi(Y, \sigma, \eta) = \begin{cases} \int_{\underline{\pi}}^{\phi(Y, \sigma, \eta)} \{(1 - \delta)\pi + \delta[(1 - \sigma)Y + \sigma W] - (1 - \delta)\sigma\gamma(Y - \alpha\mu)\} h(\pi) d\pi & \text{if } \sigma \leq \theta \\ + \int_{\phi(Y, \sigma, \eta)}^{\bar{\pi}} [(1 - \delta)\alpha\pi + \delta W - (1 - \delta)\sigma\gamma(Y - \alpha\mu)] h(\pi) d\pi & \\ \int_{\underline{\pi}}^{\phi(Y, \sigma, \eta)} \{(1 - \delta)\pi + \delta[(1 - \sigma)Y + \sigma W] - (1 - \delta)\sigma\gamma(Y - \alpha\mu)\} h(\pi) d\pi & \text{if } \theta < \sigma \\ + \int_{\phi(Y, \sigma, \eta)}^{\bar{\pi}} [(1 - \delta)\alpha\pi + \delta W - (1 - \delta)\omega\gamma(Y - \alpha\mu)] h(\pi) d\pi & \end{cases}$$

To understand this expression, first consider when  $\sigma \leq \theta$ , in which case leniency is not used. If  $\pi \in [\underline{\pi}, \phi(Y, \sigma, \eta)]$  then collusion is incentive compatible; each firm earns current profit of  $\pi$ , incurs an expected penalty of  $\sigma\gamma(Y - \alpha\mu)$ , and has an expected future payoff of  $(1 - \sigma)Y + \sigma W$ . If instead  $\pi \in (\phi(Y, \sigma, \eta), \bar{\pi}]$  then collusion is not incentive compatible; so each firm earns current profit of  $\alpha\pi$ , incurs an expected penalty of  $\sigma\gamma(Y - \alpha\mu)$ , and has an expected future payoff of  $W$ . The expression

---

<sup>12</sup>As specified in the ICC in (3), the penalty is slightly different from that in Harrington and Chang (2009) or HC09. In terms of rescaled payoffs, HC09 assumes the penalty is  $\gamma(Y - \alpha\mu)$ , while here it is  $(1 - \delta)\gamma(Y - \alpha\mu)$ . This means that HC09 assumes that a conviction results in an infinite stream of single-period penalties of  $\gamma(Y - \alpha\mu)$  which has a present value of  $\gamma\left(\frac{Y - \alpha\mu}{1 - \delta}\right)$ , while the current paper assumes a one-time penalty of  $(1 - \delta)\gamma(Y - \alpha\mu)$  which has a present value of  $\gamma(Y - \alpha\mu)$ . We now believe the latter specification is more sound. For the specification in HC09, every time a cartel is convicted, it has to pay a penalty of  $\gamma(Y - \alpha\mu)$  *ad infinitum*. Thus, if it has been convicted  $k$  times in the past then it is paying  $k\gamma(Y - \alpha\mu)$  in each period, while earning an average collusive profit of  $\mu$  in each period. As  $k \rightarrow \infty$ , the penalty is unbounded while the payoff from collusion is not. It can be shown that the penalty specification in HC09 implies  $\lim_{\delta \rightarrow 1} Y = \alpha\mu$  so that the penalty wipes out all gains from colluding. These properties do not seem desirable, and we believe it is better to assume the penalty is a one-time payment  $\gamma(Y - \alpha\mu)$  rather than an infinite stream of  $\gamma(Y - \alpha\mu)$ . It is important to note that this change in specification does not affect the conclusions in HC09 because of the parameter  $\gamma$ . Starting with the original specification  $\gamma(Y - \alpha\mu)$  and defining  $\tilde{\gamma} \equiv \gamma/(1 - \delta)$ , the analysis in HC09 is equivalent to when the penalty is  $(1 - \delta)\tilde{\gamma}(Y - \alpha\mu)$ . This transformation works as long as  $\delta$  is fixed. As the main results in HC09 do not involve performing comparative statics with respect to  $\delta$  or letting  $\delta \rightarrow 1$ , the conclusions in HC09 remain intact.

when  $\theta < \sigma$  differs only when collusion breaks down in which case all firms apply for leniency and the expected penalty is  $\omega\gamma(Y - \alpha\mu)$ .

A fixed point to  $\psi$  is an equilibrium value for  $Y$ . That is, given an anticipated future collusive value  $Y$ , the resulting equilibrium behavior - as represented by  $\phi(Y, \sigma, \eta)$  - results in firms colluding for market states such that the value to being in a cartel is  $Y$ . We then want to solve:  $Y^* = \psi(Y^*, \sigma, \eta)$ . As an initial step to exploring the set of fixed points, first note that  $\psi(\alpha\mu, \sigma, \eta) = \alpha\mu$ . Hence, one fixed point to  $\psi$  is the degenerate solution without collusion. If there is a fixed point with collusion - that is,  $Y > \alpha\mu$  - then we select the one with the highest value.

Given  $Y^*(\sigma, \eta)$ , define

$$\phi^*(\sigma, \eta) \equiv \max \{ \min \{ \phi(Y^*(\sigma, \eta), \sigma, \eta), \bar{\pi} \}, \underline{\pi} \}, \quad (5)$$

as the maximum profit realization such that a type- $\eta$  cartel is stable. It is a measure of cartel stability since the cartel is stable iff  $\pi \leq \phi^*(\sigma, \eta)$  and thus internally collapses with probability  $1 - H(\phi^*(\sigma, \eta))$ . Note that if  $\phi^*(\sigma, \eta) = \bar{\pi}$  then the cartel is stable for all market conditions (so it never internally collapses), and if  $\phi^*(\sigma, \eta) = \underline{\pi}$  then the cartel is unstable for all market conditions (so firms never collude).

### 3.2 Stationary Distribution of Cartels

Given  $\phi^*(\sigma, \eta)$ , the stochastic process by which cartels are born and die (either through internal collapse or being shut down by the CA) is characterized in this section. The random events driving this process are the opportunity to cartelize, market conditions, and conviction by the CA. We initially characterize the stationary distribution for type- $\eta$  industries. The stationary distribution for the entire population of industries is then derived by integrating the type specific distributions over all types.

Consider an arbitrary type- $\eta$  industry. If it is not cartelized at the end of the preceding period then, by the analysis in Section 3.1, it'll be cartelized at the end of the current period with probability  $\kappa(1 - \sigma)H(\phi^*(\sigma, \eta))$ . With probability  $\kappa$  it has the opportunity to cartelize, with probability  $H(\phi^*(\sigma, \eta))$  the realization of  $\pi$  is such that collusion is incentive compatible, and with probability  $1 - \sigma$  it is not caught and convicted by the CA. If instead the industry was cartelized at the end of the previous period, it'll still be cartelized at the end of this period with probability  $(1 - \sigma)H(\phi^*(\sigma, \eta))$ .

Let  $NC(\sigma, \eta)$  denote the proportion of type- $\eta$  industries that are not cartelized. The stationary rate of non-cartels is defined by :

$$\begin{aligned} NC(\sigma, \eta) = & NC(\sigma, \eta) [(1 - \kappa) + \kappa(1 - H(\phi^*)) + \kappa\sigma H(\phi^*)] \\ & + [1 - NC(\sigma, \eta)] [(1 - H(\phi^*)) + \sigma H(\phi^*)] \end{aligned} \quad (6)$$

Examining the RHS of (6), a fraction  $NC(\sigma, \eta)$  of type- $\eta$  industries were not cartelized in the previous period. Out of those industries, a fraction  $1 - \kappa$  will not have the opportunity to cartelize in the current period. A fraction  $\kappa(1 - H(\phi^*))$  will have the opportunity but, due to a high realization of  $\pi$ , find it is not incentive compatible to collude, while a fraction  $\kappa\sigma H(\phi^*)$  will cartelize and collude but then are discovered

by the CA. Of the industries that were colluding in the previous period, which have mass  $1 - NC(\sigma, \eta)$ , a fraction  $1 - H(\phi^*)$  will collapse for internal reasons and a fraction  $\sigma H(\phi^*)$  will instead be shut down by the authorities.

Solving (6) for  $NC(\sigma, \eta)$ :

$$NC(\sigma, \eta) = \frac{1 - (1 - \sigma) H(\phi^*(\sigma, \eta))}{1 - (1 - \kappa)(1 - \sigma) H(\phi^*(\sigma, \eta))}. \quad (7)$$

For the stationary distribution, the fraction of cartels among type- $\eta$  industries is then:

$$C(\sigma, \eta) \equiv 1 - NC(\sigma, \eta) = \frac{\kappa(1 - \sigma) H(\phi^*(\sigma, \eta))}{1 - (1 - \kappa)(1 - \sigma) H(\phi^*(\sigma, \eta))}. \quad (8)$$

Finally, the derivation of the entire population of industries is performed by integrating the type- $\eta$  distribution over  $\eta \in [\underline{\eta}, \bar{\eta}]$ . The mass of cartelized industries, which we refer to as the cartel rate  $C(\sigma)$ , is then defined by:

$$C(\sigma) \equiv \int_{\underline{\eta}}^{\bar{\eta}} C(\sigma, \eta) g(\eta) d\eta = \int_{\underline{\eta}}^{\bar{\eta}} \left[ \frac{\kappa(1 - \sigma) H(\phi^*(\sigma, \eta))}{1 - (1 - \kappa)(1 - \sigma) H(\phi^*(\sigma, \eta))} \right] g(\eta) d\eta. \quad (9)$$

### 3.3 Equilibrium Non-Leniency Enforcement

Recall that  $\sigma = qrs$  where  $q$  is the probability of a cartel being discovered,  $r$  is the probability that the CA investigates a reported case, and  $s$  is the probability of it succeeding with the investigation. We now want to derive the equilibrium value of  $s$ , where  $s = p(\lambda L + R)$ ,  $L$  is the mass of leniency cases, and  $R$  is the mass of non-leniency cases handled by the CA. As both  $L$  and  $R$  depend on the cartel rate  $C$  and the cartel rate depends on  $s$  (through  $\sigma$ ), this is a fixed point problem. We need to find a value for  $s$ , call it  $s'$ , such that, given  $\sigma = qrs'$ , the induced cartel rate  $C(qrs')$  is such that it generates  $L$  and  $R$  so that  $p(\lambda L + R) = s'$ .

With our expression for the cartel rate, we can provide expressions for  $L$  and  $R$ . The mass of cartel cases generated by the leniency program is:

$$L(\sigma) = \begin{cases} 0 & \text{if } \sigma \leq \theta \\ \int_{\underline{\eta}}^{\bar{\eta}} (1 - H(\phi^*(\sigma, \eta))) C(\sigma, \eta) g(\eta) d\eta & \text{if } \theta < \sigma \end{cases} \quad (10)$$

In (10), note that an industry does not apply for leniency when it is still effectively colluding. When collusion stops, leniency is used when the only equilibrium is that all firms apply for leniency, which is the case when  $\theta < \sigma$ . Thus, when  $\theta < \sigma$ ,  $L$  equals the mass of cartels that collapse due to a high realization of  $\pi$ . That it is dying cartels that apply for leniency is consistent with the European Commission experience.<sup>13, 14</sup>

<sup>13</sup>EC official Olivier Guersent expressed a concern that leniency applications were coming from dying cartels at the 11th Annual EU Competition Law and Policy Workshop: Enforcement of Prohibition of Cartels in Florence, Italy in June 2006. Jun Zhou expressed in a conversation that only 13 out of 110 EC cases with a leniency awardee (over 1996-2012) involved applications before the death of the cartel.

<sup>14</sup>That either all firms or no firms apply for leniency is a property of not only our analysis but

The mass of cartel cases generated without use of the leniency program is

$$R(\sigma) = \begin{cases} qrC(\sigma) & \text{if } \sigma \leq \theta \\ qr \int_{\underline{\eta}}^{\bar{\eta}} H(\phi^*(\sigma, \eta)) C(\sigma, \eta) g(\eta) d\eta & \text{if } \theta < \sigma \end{cases} \quad (11)$$

If the leniency program is never used (which is when  $\sigma \leq \theta$ ), then the mass of cases being handled by the CA is  $qrC(\sigma)$ . If instead  $\theta < \sigma$ , so that dying cartels use the leniency program, then the cartels left to be caught are those which have not collapsed in the current period which is  $\int_{\underline{\eta}}^{\bar{\eta}} H(\phi^*(\sigma, \eta)) C(\sigma, \eta) g(\eta) d\eta$ .

The equilibrium probability of a CA successfully getting a cartel to pay penalties (without use of the leniency program) is the solution to the following fixed point problem:

$$\sigma = \Psi(\sigma) \equiv \begin{cases} qrp(qrC(\sigma)) & \text{if } \sigma \leq \theta \\ qrp \left( \lambda \int_{\underline{\eta}}^{\bar{\eta}} (1 - H(\phi^*(\sigma, \eta))) C(\sigma, \eta) g(\eta) d\eta \right. \\ \left. + qr \int_{\underline{\eta}}^{\bar{\eta}} H(\phi^*(\sigma, \eta)) C(\sigma, \eta) g(\eta) d\eta \right) & \text{if } \theta < \sigma \end{cases} \quad (12)$$

where we have substituted for  $L$  using (10) and  $R$  using (11).<sup>15</sup> If there are multiple solutions to (12) then it is assumed the maximal one is selected.<sup>16</sup>

### 3.4 Optimal Competition Policy

The analytical results in Section 5 are derived taking enforcement policy - as parameterized by  $r$  which is the fraction of possible cases that the CA takes on - as fixed. However, as argued later, they extend as well to when  $r$  is endogenized by assuming the CA acts to minimize the cartel rate. Let  $\sigma^*(r)$  denote the maximal solution to (12), where its dependence on  $r$  is now made explicit. For when  $r$  is endogenized, it is assumed that  $r = r^*$  where

$$r^* \in \arg \min_{r \in [0,1]} C(\sigma^*(r)).$$

By having the prosecution policy chosen to minimize the cartel rate, the analysis delivers an upper bound on welfare.

---

all previous analyses on leniency programs with the exception of some recent work by one of the authors, Harrington (2012, 2013). In those two papers, there is private information between cartel members which can explain why only one firm would come forward to the CA.

<sup>15</sup>Note that the fixed point can be defined in terms of either  $\sigma$  or  $s$  given that  $\sigma = qrs$  and, at this point of the analysis,  $q$  and  $r$  are parameters.

<sup>16</sup>We conjecture that results hold with some other selections, such as the minimal fixed point to  $\Psi$ . What is necessary is that a shift up (down) in  $\Psi$  increases (decreases)  $\sigma^*$ .

## 4 Impact of a Leniency Program When Non-Leniency Enforcement is Exogenous

In this section, results are derived under the standard assumption in the literature that non-leniency enforcement is fixed. These results are a necessary intermediate step towards deriving the paper’s main results for when non-leniency enforcement is endogenized, but also serve as a benchmark for highlighting how an evaluation of a leniency program significantly changes when a more comprehensive analysis is performed.

To begin, consider the cartel rate function  $C(\sigma)$ ; that is, the cartel rate that results for a given level of non-leniency enforcement  $\sigma$ . Theorem 1 is a re-statement of a result in Harrington and Chang (2009) and shows that when firms assign a higher probability to the CA discovering, prosecuting, and convicting cartels then a smaller fraction of industries are cartelized. This result is derived for when the penalty multiple  $\gamma$  is not too high, which must hold if collusion is to emerge in equilibrium. In the ensuing analysis, it is assumed (without generally being stated) that  $\gamma$  is sufficiently low so that Theorem 1 holds.<sup>17</sup>

**Theorem 1**  $\exists \hat{\gamma} > 0$  such that if  $\gamma \in [0, \hat{\gamma})$  then  $C(\sigma)$  is non-increasing in  $\sigma$  and if  $C(\sigma) > 0$  then  $C(\sigma)$  is decreasing in  $\sigma$ .

Next let us consider how a leniency program affects the cartel rate function; that is, whether a leniency program results in a higher or lower cartel rate for a given value of  $\sigma$ . For this purpose, all functions will be subscripted with  $\theta$  when there is a leniency program with parameter  $\theta$ ; for example,  $C_\theta(\sigma)$  denotes the cartel rate function in that case.<sup>18</sup> Similarly, all functions will be subscripted with  $NL$  when there is no leniency program; for example,  $C_{NL}(\sigma)$  denotes the cartel rate function.

Theorem 2 shows that a leniency program does not raise the cartel rate function, and it reduces the cartel rate function if, in the absence of a leniency program, there is a positive measure of industries that cannot fully collude and a positive measure that can collude (Assumption A1).<sup>19</sup> In sum, for a given level of non-leniency enforcement, a leniency program results in fewer cartels.

**Assumption A1** There is positive measure of values for  $\eta$  such that  $\phi_{NL}^*(\sigma, \eta) < \bar{\pi}$  and a positive measure of values for  $\eta$  such that  $\phi_{NL}^*(\sigma, \eta) > \underline{\pi}$ .

---

<sup>17</sup>The analogous result in Harrington and Chang (2009) assumes  $\sigma$  is sufficiently small which we do not want to do here because  $\sigma$  will later be endogenized. Instead, results are derived for when the penalty multiple  $\gamma$  is sufficiently small. This involves a straightforward modification of the proof in Harrington and Chang (2009) and thus is not included here though is available on request.

<sup>18</sup>Recall that a firm pays a fraction  $\theta$  of the standard penalty when it receives leniency and pays, in expectation, a fraction  $\omega$  when all firms apply for leniency. To reduce notational clutter, we will suppress  $\omega$ .

<sup>19</sup>A1 ensures that the cartel rate is positive but not maximal, and rules out extreme cases in which a leniency program does not lower the cartel rate because the cartel rate is either zero without a leniency program or the environment is so conducive to collusion that the cartel rate is maximal with or without a leniency program.

**Theorem 2** *If  $\sigma \in (\theta, \omega)$  then  $C_{NL}(\sigma) \geq C_\theta(\sigma)$  and if A1 holds then  $C_{NL}(\sigma) > C_\theta(\sigma)$ .*

Prior to explaining why Theorem 2 is true, let us interpret the restriction  $\sigma \in (\theta, \omega)$ . If  $\sigma > \theta$  then a firm that contemplates deviating from a cartel would apply for leniency because it reduces the expected penalty from  $\sigma\gamma(Y - \alpha\mu)$  to  $\theta\gamma(Y - \alpha\mu)$ .  $\sigma > \theta$  also has the implication that, in response to the internal collapse of a cartel, all firms apply for leniency because it is the unique equilibrium play.<sup>20</sup> In that case, if  $\sigma < \omega$  then a firm's expected penalty rises with a leniency program from  $\sigma\gamma(Y - \alpha\mu)$  to  $\omega\gamma(Y - \alpha\mu)$ . Thus, if  $\sigma \in (\theta, \omega)$  then a firm will use the leniency program if it deviates or if the cartel collapses and, in the latter situation, expected penalties are higher compared to when there is no leniency program. Assessing the restrictiveness of  $\sigma \in (\theta, \omega)$ , if there is full leniency ( $\theta = 0$ ) then  $\sigma > \theta$  is not restrictive at all; in fact, we will focus on the case of full leniency in Section 5. If all firms applying for leniency gives each an equal chance of receiving it, then  $\omega = \frac{n-1+\theta}{n} \geq \frac{n-1}{n} \geq \frac{1}{2}$ . Thus,  $\sigma < 1/2$  is sufficient for  $\sigma < \omega$  to be satisfied. Given that  $\sigma = qrs$ , it is sufficient to assume that the probability of discovering a cartel is less than one half,  $q < 1/2$ , which is compelling. In conclusion,  $\sigma \in (\theta, \omega)$  is a weak restriction.

Under fairly general conditions, Theorem 2 shows that a leniency program reduces the frequency of cartels *when non-lenieny enforcement is fixed*. Let us summarize the forces that are the basis for that result.<sup>21</sup> A leniency program increases the payoff to cheating because now a firm can reduce its penalty by simultaneously applying for leniency. This shrinks the set of market conditions for which collusion is stable and thereby reduces expected cartel duration and the value to colluding. A leniency program also reduces the value to colluding because, upon collapse, firms race for leniency and that results in higher expected penalties. Due to the lower value to colluding, either a cartel no longer forms or it has shorter duration and this translates into a lower aggregate cartel rate.

## 5 Impact of a Leniency Program When Non-Leniency Enforcement is Endogenous

We are now prepared to provide a comprehensive assessment of how a corporate leniency program influences the cartel rate. Our plan is to address the following questions: Taking account of the effect of a leniency program on a CA's prosecutions of both leniency and non-lenieny cases, can a leniency program raise the cartel rate? If it can, under what circumstances can we be assured that a leniency program is lowering the cartel rate?

---

<sup>20</sup>If all other firms were expected not to apply then a firm's penalty from applying is  $\theta\gamma(Y - \alpha\mu)$  and from not applying is  $\sigma\gamma(Y - \alpha\mu)$ . If  $\sigma > \theta$  then a firm prefers to apply in which case all firms not applying is not an equilibrium. Given that it is optimal to apply when one or more other firms apply, equilibrium play must then involve all firms applying when  $\sigma > \theta$ .

<sup>21</sup>These forces are not new to the literature; see Motta and Polo (2003), Spagnolo (2003), and Harrington (2008).



The analysis will focus on when the leniency program provides full leniency to the first firm to come forward ( $\theta = 0$ ). This is a natural case to consider because almost all leniency programs waive all government penalties to the first firm to come forward prior to the start of an investigation.<sup>22</sup> We also focus on this case for technical reasons related to the existence of equilibrium.

To economize on notation and make it easier for the reader to follow the analysis, expressions with an  $NL$  subscript will refer to the case of "no leniency program," while those with an  $L$  subscript will refer to the case of a "full leniency program." For example,  $C_{NL}(\sigma)$  and  $C_L(\sigma)$  are, respectively, the cartel rate functions without leniency and with (full) leniency. The associated equilibrium values for non-leniency enforcement are  $\sigma_{NL}^*$  and  $\sigma_L^*$  in which case the equilibrium cartel rates are  $C_{NL}(\sigma_{NL}^*)$  and  $C_L(\sigma_L^*)$ .

With this notation, we can summarize the task before us. Theorem 2 showed that a leniency program lowers the cartel rate given a value for  $\sigma$ :  $C_L(\sigma) < C_{NL}(\sigma)$ . Whether a leniency program raises or lowers the cartel rate then comes down to its impact on non-leniency enforcement. If a leniency program strengthens non-leniency enforcement -  $\sigma_L^* > \sigma_{NL}^*$  - then clearly a leniency program lowers the cartel rate because, by Theorem 1, cartel rate functions are decreasing:  $C_L(\sigma_L^*) < C_{NL}(\sigma_{NL}^*)$ . If instead a leniency program weakens non-leniency enforcement -  $\sigma_L^* < \sigma_{NL}^*$  - then the ultimate impact on the cartel rate depends on the extent to which a leniency program reduces non-leniency enforcement.

The analysis is conducted holding fixed the CA's non-leniency enforcement instrument  $r$ , which is the proportion of discovered cases that it prosecutes. Note that if it is shown that a leniency program decreases (increases) the cartel rate for all values of  $r > 0$  then allowing  $r$  to be chosen to minimize the cartel rate will still result in a leniency program decreasing (increasing) the cartel rate. Thus, the conclusions of this section are applicable to when a CA acts in a welfare-maximizing manner by choosing its caseload to minimize the cartel rate.

In Section 5.1, it is shown that an equilibrium cartel rate exists for either when there is no leniency program or a full leniency program. In Section 5.2, it is shown that a leniency program can raise the cartel rate. In Section 5.3, conditions are derived for a leniency program to lower the cartel rate. While the analysis largely focuses on how a leniency program affects the frequency of cartels, Section 5.4 shows the differential impact of a leniency program across industries; specifically, a leniency program can make collusion more difficult in some industries but less difficult in other industries.

## 5.1 Existence of an Equilibrium Cartel Rate

The equilibrium level of non-leniency enforcement  $\sigma^*$  is a fixed point to  $\Psi$  which is defined in (12).  $\sigma^*$  has the property that if firms believe the per period probability

---

<sup>22</sup>In the U.S., a firm that receives amnesty is still liable for single customer damages so leniency is not full. Most other jurisdictions do not have customer damages in which case government fines encompass the entirety of penalties and, therefore, leniency is full.

of paying penalties (through non-leniency enforcement) is  $\sigma^*$  then the induced cartel birth and death rates generate a caseload for the CA whereby the equilibrium probability is indeed  $\sigma^*$ .

**Theorem 3** For  $\theta \in \{0, 1\}$ ,  $\exists \hat{\gamma} > 0$  such that if  $\gamma \in [0, \hat{\gamma})$  then  $\sigma^*$  exists.

While we will only be assessing the impact of providing full leniency, we do not believe the intuition behind the results is tied to leniency being full. We conjecture that, as long as  $\sigma^*$  exists, results will go through if, in equilibrium, firms use leniency in response to cartel collapse (that is,  $\theta < \sigma^*$ ) and penalties are higher as a result (that is,  $\sigma^* < \omega$ ).

## 5.2 A Leniency Program Increases the Cartel Rate

In this section, we show that a leniency program can be counter-productive. When penalties are not severe enough and the amount of resources saved by prosecuting a leniency case are not large enough then the introduction of a leniency program *raises* the cartel rate.

**Theorem 4** Assume

$$\int_{\underline{\eta}}^{\bar{\eta}} (1 - H(\phi_{NL}^*(\sigma_{NL}^*, \eta))) C_{NL}(\sigma_{NL}^*, \eta) g(\eta) d\eta > 0 \quad (13)$$

so that, without a leniency program, there are cartels that collude and internally collapse. Generically, there exists  $\hat{\lambda} < 1$  and  $\hat{\gamma} > 0$  such that if  $(\gamma, \lambda) \in [0, \hat{\gamma}] \times [\hat{\lambda}, 1]$  then the cartel rate with a leniency program strictly exceeds the cartel rate without a leniency program.

In understanding the forces that drive this result, first note that a leniency program can affect the cartel rate by *disabling* active cartels (that is, shutting them down) and by *deterring* new cartels from forming. A leniency program can have a perverse effect because, while it generally promotes deterrence, it can actually result in fewer cartels being shut down.

Prior to the introduction of a leniency program, the CA is discovering, prosecuting, and convicting cartels through non-leniency means. While some of the cartels that are convicted will just so happened to have internally collapsed, many of them will have been active in which case it is their prosecution and conviction that shuts the cartel down. When a leniency program is introduced, cartels that collapse race for leniency and these leniency applications comprise part of the caseload of the CA. Of particular note is that leniency cases are coming from *dying cartels* and thus prosecution of them is not shutting down an active cartel. However, these leniency cases are adding to the CA's caseload and thereby result in less success in prosecuting non-leniency cases which, if it had led to a conviction, would have disabled a well-functioning cartel. In essence, leniency cases - which do not shut down an active

cartel - are crowding out non-leniency cases - which generally do shut down active cartels. If leniency cases do not save much in terms of prosecutorial resources (that is,  $\lambda$  is not sufficiently below one) then this crowding out effect is significant and the end result is that many fewer cartels are shut down when there is a leniency program.

This is not the end of the story, however. Due to the leniency program, a dying cartel is now assured of paying penalties because one of its members will enter the leniency program and aid the CA is obtaining a conviction. In contrast, without a leniency program, only a fraction of those cartels would have been discovered and penalized (specifically, the fraction is  $\sigma_{NL}^*$ ). Thus, a leniency program is raising the expected penalties for a cartel in the event of its death which serves to deter some cartels from forming. However, if penalties are not large enough (that is,  $\gamma$  is not sufficiently great) then the additional cartels deterred due to the leniency program is small in comparison to the reduction in the number of cartels shut down because leniency cases crowd out non-leniency cases. As a result, on net, the cartel rate is higher. Thus, in spite of the leniency program apparently "working" in the sense of bringing forth leniency applications, it is actually counter-productive in that the latent cartel rate is higher.

Theorem 4 shows that, for any value of  $r$  (which recall is the fraction of possible non-leniency cases that the CA chooses to prosecute), the cartel rate is higher without a leniency program:<sup>23</sup>

$$C_L(qr s_L^*(r)) > C_{NL}(qr s_{NL}^*(r)), \forall r > 0. \quad (14)$$

Recall that  $\sigma = qrs$ , and in (14) we have made explicit the dependence of the conviction rate  $s$  on  $r$ . If the CA chooses its caseload in order to minimize the cartel rate, the optimal prosecution policy with and without a leniency program, respectively, is

$$r_L^* \in \arg \min_{r \in [0,1]} C_L(qr s_L^*(r)), \quad r_{NL}^* \in \arg \min_{r \in [0,1]} C_{NL}(qr s_{NL}^*(r)).$$

It then follows from (14),

$$C_L(qr_L^* s_L^*(r_L^*)) > C_{NL}(qr_{NL}^* s_{NL}^*(r_{NL}^*)).$$

This leads to the following corollary.

**Corollary 5** *Under the conditions of Theorem 4, a leniency program raises the cartel rate even if the competition authority is choosing its caseload to minimize the cartel rate.*

It is worth emphasizing this last result: Even if the CA is acting in society's best interests - by choosing how many cases to prosecute in order to minimize the presence of cartels - a leniency program can still cause there to be more cartels. While, in

---

<sup>23</sup>Theorem 4 does require that for any value of  $r$ , (13) is satisfied which is not very restrictive because whether a cartel collapses is partly due to forces unrelated to the CA. For example, even in the absence of a CA, all cartels will collapse with positive probability when  $\bar{\pi}$  is sufficiently high and  $\eta > 1$ .

practice, the interests of a CA may not align with the interests of society, it does not take an agency problem for a leniency program to be counter-productive. Next note that a welfare-maximizing CA could ensure that the cartel rate is no higher with a leniency program *if* it controlled both how many leniency and non-leniency cases it takes on (because it could always decline to pursue all leniency applications and thus result in the same cartel rate as when there is not a leniency program). Consistent with established practice, we have assumed that a CA must take on all leniency cases and thus can only control how many non-leniency cases it prosecutes. But even if a CA could control how many leniency cases it takes on, we conjecture that a reasonable specification of a CA's objective would have it prosecuting all leniency cases (now, there is an agency problem). A CA wants to appear effective and taking on many cases that deliver convictions serves that end. The appeal of a leniency case is that conviction is far more likely than with a non-leniency case which would probably lead a CA to take on all leniency cases. However, as revealed by Theorem 4, such a policy is not necessarily in the best interests of society. It may be better to take on fewer leniency cases and more non-leniency cases because the latter serves to shut down more cartels and, as we will show in Section 5.4, deters some cartels from forming.

In concluding this section, let us argue that the sufficient conditions in Theorem 4 for a leniency program to raise the cartel rate could plausibly hold in some jurisdictions. A leniency program raises the cartel rate when: 1) a leniency case takes up a reasonable amount of resources so that there is a crowding out of resources for non-leniency cases; and 2) penalties are not so severe that they significantly deter cartel formation. With regards to the first condition, the Directorate General Competition (DG Comp) of the European Commission has been overwhelmed with leniency applications which could well have significantly limited the availability of resources for prosecuting other cases.<sup>24</sup>

DG Competition is now in many ways the victim of its own success; leniency applicants are flowing through the door of its Rue Joseph II offices, and as a result the small Cartel Directorate is overwhelmed with work. ... It is open to question whether a Cartel Directorate consisting of only approximately 60 staff is really sufficient for the Commission to tackle the 50 cartels now on its books.

Furthermore, the impact of a leniency program on enforcement through other means is a concern emphasized in Friederiszick and Maier-Rigaud (2008). Both authors were members of DG Comp and their paper recommended that the DG Comp step up non-leniency enforcement methods such as being active in detecting cartels. Consistent with these views, Kai-Uwe Kühn, who recently stepped down as Chief Economist of the European Commission, expressed at the Searle Research Symposium in September 2010 that leniency cases seem to be as long and involved as non-leniency cases. Based on these sources, it seems quite plausible that jurisdictions with an active leniency program could be experiencing weakened non-leniency enforcement because of the crowding out of non-leniency methods of enforcement.

---

<sup>24</sup>Riley (2007), pp. 1-2.

Turning to the second condition, many countries have set legal caps on fines that, for cartels in reasonably-sized markets, are likely to be far below the incremental profit from colluding and thus do little to deter cartel formation. For example, the maximum penalty in Chile is around \$25 million per defendant. By way of comparison, there is currently a case against a suspected cartel in the wholesale chicken market which has annual sales on the order of \$1 billion. Even more paltry are caps of around \$7 million in Mexico (at least until 2011 when it was increased) and \$5 million in Japan.<sup>25</sup>

In sum, there is at least this anecdotal evidence to suggest that - consistent with the conditions for a leniency program to raise the cartel rate - leniency cases could be negatively affecting non-lenieny enforcement - and thereby reducing the rate at which a CA shuts down active cartels - and penalties could be at levels that do not result in leniency programs substantively deterring cartel formation.

### 5.3 A Leniency Program Decreases the Cartel Rate

As just shown, a leniency program is not assured of reducing the frequency of cartels; it can raise the cartel rate. One tactic that a CA can take to avoid this outcome and ensure that a leniency program serves the cause of fighting cartels is to set up a procedure that will expeditiously handle leniency cases. Theorem 6 establishes that if leniency cases save sufficient resources - relative to non-lenieny cases - then a leniency program will lower the cartel rate.<sup>26</sup>

**Theorem 6** *If  $\sigma_{NL}^* \in (0, \omega)$  then there exists  $\hat{\lambda} > 0$  such that if  $\lambda \in [0, \hat{\lambda}]$  then the cartel rate is weakly lower with a leniency program and if A1 holds then the cartel rate is strictly lower with a leniency program.*

While this result is not particularly surprising, it is important to understand why it is true. It is not just that the crowding out of non-lenieny enforcement is reduced - and thus non-lenieny enforcement does not fall as much - but rather that non-lenieny enforcement can actually be higher.

Key to a leniency program raising the cartel rate is that leniency cases are consuming valuable CA resources which detracts from non-lenieny enforcement. However, if leniency cases can be handled with few resources then they will not crowd out many non-lenieny cases. Though there will still be some crowding out - which would suggest non-lenieny enforcement would still be harmed and thus fewer cartels are shut down by the CA - there is a mitigating effect coming from the enhanced deterrence of cartel formation due to a leniency program. By raising the penalties that a cartel can expect to pay when it collapses (and cartel members subsequently race for leniency), a leniency program results in fewer cartels forming. With fewer

---

<sup>25</sup>"Setting of Fines for Cartels in ICN Jurisdictions," Report to the 7th ICN Annual Conference, Kyoto, April 2008.

<sup>26</sup>Recall from Section 4 that, for a reasonable specification of  $\omega$ ,  $q < 1/2$  is sufficient for  $\sigma_{NL}^* \in (0, \omega)$ .

cartels forming, there will be fewer non-leniency cases. Thus, non-leniency enforcement could be stronger (that is,  $\sigma_L^* > \sigma_{NL}^*$ ) because, while there are fewer resources for non-leniency cases due to the presence of leniency cases on a CA's docket, there are also fewer non-leniency cases.

To appreciate how non-leniency enforcement can be stronger when  $\lambda$  is sufficiently small, consider the extreme case of  $\lambda = 0$  so that leniency cases require no resources. (As results are continuous in  $\lambda$ , the argument will also apply when  $\lambda$  is sufficiently low.) Holding the cartel rate fixed at the level without a leniency program, a leniency program enhances non-leniency enforcement because there is no crowding out (due to  $\lambda = 0$ ) *and* there are fewer non-leniency cases due to some of them being handled as leniency cases. Given that non-leniency enforcement is stronger and expected penalties are higher, the cartel rate is lower which means there are fewer non-leniency cases which serves to enhance non-leniency enforcement more. This feedback effect initiated by a leniency program - fewer cartels leads to fewer non-leniency cases which leads to stronger non-leniency enforcement which leads to fewer cartels - ultimately results in both a lower cartel rate and stronger non-leniency enforcement. This feedback effect suggests that there may be a large return to reducing  $\lambda$  by streamlining the handling of leniency cases.

Another circumstance in which a CA can be assured that a leniency program will lower the cartel rate is if it is introduced in a jurisdiction for which enforcement is already very weak as reflected in a low likelihood that a cartel will be even considered for prosecution. In that case, a leniency program is sure to be beneficial.

**Theorem 7** *There exists  $\hat{q} > 0$  such that if  $q \in [0, \hat{q}]$  then the cartel rate is weakly lower with a leniency program and if A1 holds then the cartel rate is strictly lower with a leniency program.*

If non-leniency enforcement is largely absent prior to the introduction of a leniency program then a leniency program cannot have much of a crowding-out effect for the simple reason that there aren't many non-leniency cases to crowd out. Hence, if a CA is not actively engaged in enforcement prior to introducing a leniency program then a leniency program is sure to be effective in reducing the frequency of cartels.

It is worth noting that Theorem 7 is the one result that does depend on there being full leniency:  $\theta = 0$ . If  $\theta > 0$  then, as  $q \rightarrow 0$  and  $\sigma (= qsr) \rightarrow 0$ , the leniency program has no effect because no firm would use it. This comment highlights the complementarity between leniency and non-leniency enforcement: If  $\sigma < \theta$  then a leniency program is irrelevant because the chances of being caught through non-leniency means are sufficiently low to make applying for leniency not to be in a firm's interests. The efficacy of a leniency program depends on cartel members believing there is a sufficient chance of them being caught and convicted by the CA.<sup>27</sup>

---

<sup>27</sup>A caveat is appropriate here because we have assumed that firms achieve the Pareto-superior equilibrium when it comes to applying for leniency; that is, if there is an equilibrium in which no firms seek leniency then that is the equilibrium upon which firms coordinate. However, experimental evidence suggests that a leniency program can be effective even when  $\sigma = 0$  (Bigoni et al, 2012). In that case, presumably a firm is applying for leniency out of concern that a rival will apply for leniency which is sensible for the rival only if it possesses a similar concern.

## 5.4 Inter-Industry Variation in the Impact of a Leniency Program

Our analysis has shown how a leniency program impacts the frequency of cartels. A leniency program can lower the cartel rate by resulting in some cartels no longer forming and reducing average cartel duration for those that do form. The focus thus far has been on the aggregate cartel rate but, as the analysis of this section will reveal, a leniency program can have qualitatively distinct effects across industries.

Recall that industries vary with respect to the parameter  $\eta$  where a higher value of  $\eta$  means a higher profit increase from cheating on the collusive arrangement. A higher value for  $\eta$  could be due, for example, to more firms (assuming Bertrand price competition) or a higher price elasticity to the firm demand function. When  $\eta$  is higher, the greater incentive to deviate means that the cartel is less stable in the sense that it will internally collapse for a wider set of market conditions (specifically,  $\phi(Y, \sigma, \eta)$  is decreasing in  $\eta$ ). This property has two implications. First, industries with sufficiently high values of  $\eta$  are unable to cartelize; and recall that  $\hat{\eta}$  denotes the highest value for  $\eta$  such that a cartel forms with positive probability. Second, when cartels are able to form ( $\eta \leq \hat{\eta}$ ), average cartel duration is lower when  $\eta$  is higher. It is straightforward to show that average duration for a cartel in a type- $\eta$  industry is

$$CD(\sigma, \eta) = \frac{1}{1 - (1 - \sigma)H(\phi^*(\sigma, \eta))},$$

and that  $\phi^*(\sigma, \eta)$  is non-increasing in  $\eta$  and is decreasing in  $\eta$  if  $\phi^*(\sigma, \eta) \in (\underline{\pi}, \bar{\pi})$ . Hence, average cartel duration is decreasing in  $\eta$ . It is useful to note that an industry type's cartel rate,  $C(\sigma, \eta)$ , and its average cartel duration,  $CD(\sigma, \eta)$ , are monotonically related:

$$C(\sigma, \eta) = \frac{\kappa(CD(\sigma, \eta) - 1)}{1 + \kappa(CD(\sigma, \eta) - 1)}.$$

Thus, in assessing how the effect of a leniency program varies across industries, we can either consider its influence on the cartel rate or on cartel duration.

Given there is inter-industry variation in the presence and duration of cartels prior to a leniency program, it is natural to examine how the impact of a leniency program varies across industries. In particular, could a leniency program make the environment *less* hospitable for collusion in some industries while making it *more* hospitable in other industries? To initially address this question, numerical analysis is conducted. Results are provided both for when the probability of conviction is linear in caseload and is a concave then convex function of caseload. Details as to parameterizations and numerical methods are in Appendix B.

Figures 1 and 2 report the change in average cartel duration due to a leniency program for each industry type  $\eta \leq \hat{\eta}$ . First note that the introduction of a leniency program reduces  $\hat{\eta}$  and thereby shrinks the range of industry types for which a cartel forms with positive probability; for example, in Figure 1(b),  $\hat{\eta}$  falls from 1.55 to 1.52. A reduction in  $\hat{\eta}$  was found for almost all parameterizations though there were a few cases in which a leniency program actually increased  $\hat{\eta}$ . Second, for those industries that do cartelize with positive probability, a leniency program has a differential effect

across industries depending on whether the industry produces relatively stable cartels ( $\eta$  is low) or unstable cartels ( $\eta$  is high). Specifically, the effect of a leniency program on average cartel duration (or the cartel rate) is decreasing in  $\eta$  so that industries that produce less stable cartels tend to experience a bigger drop in cartel duration than industries with more stable cartels. This property is apparent in Figures 1 and 2 where the change in average duration is decreasing in  $\eta$ , which holds as well for all other parameterizations considered. Even more than that, a leniency program can result in *longer* duration for the most stable cartels (that is, the change in duration is positive when  $\eta$  is low) while shutting down or shortening the duration of the least stable cartels.

**Property:** A leniency program generally reduces the range of markets that are able to form cartels. The effect of a leniency program on average cartel duration is decreasing in  $\eta$  so that markets with less stable cartels experience a bigger decline in average cartel duration. This differential effect can be so significant that a leniency program can reduce average cartel duration of relatively unstable cartels and, at the same time, increase average cartel duration of relatively stable cartels.

To understand what is driving the differential effect of a leniency program across industries, recall that it is dying cartels that use the leniency program. Once market conditions are such that collusion is no longer incentive compatible, firms stop colluding and race to apply for leniency. Of course, only one firm receives leniency with the remaining firms paying full penalties. Because the leniency program then ensures conviction when the cartel dies, expected penalties are higher with a leniency program. At the same time, the flow of leniency applications can weaken non-lenieny enforcement by reducing the likelihood of being prosecuted and convicted outside of the leniency program. In sum, expected penalties can be higher through the leniency program but can be lower outside of the leniency program. Which of these effects is more important depends on an industry's type. Firms in markets that support relatively unstable cartels know there is a significant chance the cartel will internally collapse which will induce a race for leniency. Thus, those cartels are especially harmed by the higher penalties coming from a leniency applicant and, therefore, they are worse off after the introduction of a leniency program. In contrast, firms in markets that support relatively stable cartels are less concerned with a race for leniency because cartel collapse is unlikely (and such a race only ensues in that event). The greater concern for a highly stable cartel is with non-lenieny enforcement and, if that is weaker by virtue of the crowding-out effect of a leniency program, expected penalties are actually lower and, therefore, the environment is more hospitable for collusion.

Complementing these numerical results, we can prove that the least stable cartels are harmed by a leniency program when non-lenieny enforcement is not weakened, and the most stable cartels are benefitted by a leniency program when non-lenieny enforcement is weakened. Theorem 8 shows that if a market produces sufficiently stable cartels then the impact of a leniency program is determined by how it influences



non-leniency enforcement. If the introduction of a leniency program strengthens non-leniency enforcement then the cartel rate (or average cartel duration) for highly stable cartels declines, while if it weakens non-leniency enforcement then the cartel rate (or average cartel duration) rises. These highly stable cartels are not concerned with the higher penalties coming from a race for leniency - because a race is unlikely for those cartels - and instead are concerned with whether they are more or less likely to be prosecuted and convicted outside of the leniency program.

**Theorem 8** Assume  $\underline{\eta} = 1$  and  $C_{NL}(\sigma_{NL}^*, \eta), C_L(\sigma_L^*, \eta) > 0$  then

$$\lim_{\eta \rightarrow \underline{\eta}} [C_L(\sigma_L^*, \eta) - C_{NL}(\sigma_{NL}^*, \eta)] = \frac{\kappa(\sigma_{NL}^* - \sigma_L^*)}{[1 - (1 - \kappa)(1 - \sigma_L^*)][1 - (1 - \kappa)(1 - \sigma_{NL}^*)]}.$$

Hence, for  $\eta$  close to one,  $C_L(\sigma_L^*, \eta) < (>) C_{NL}(\sigma_{NL}^*, \eta)$  if and only if  $\sigma_L^* > (<) \sigma_{NL}^*$ .

The next result shows that, unless non-leniency enforcement is weakened, a leniency program is sure to destabilize the least stable cartels. More specifically, industries for which  $\eta \in (\hat{\eta}_L(\sigma_L^*), \hat{\eta}_{NL}(\sigma_{NL}^*)]$  were able to collude in the absence of a leniency program but are not able to do so with a leniency program.

**Theorem 9** If  $\sigma_{NL}^*, \sigma_L^* \in (0, \omega)$ ,  $\sigma_L^* \geq \sigma_{NL}^*$ , and  $\hat{\eta}_{NL}(\sigma_{NL}^*) \in (\underline{\eta}, \bar{\eta})$  then  $\hat{\eta}_L(\sigma_L^*) < \hat{\eta}_{NL}(\sigma_{NL}^*)$ .

In sum, the institution of a leniency program can raise expected penalties and, as a result, shorten cartel duration (or prevents cartels from forming at all) in industries for which collusion is least stable, while it can actually lengthen cartel duration in industries for which collusion is most stable because non-leniency enforcement is weaker. Thus, this theory predicts that a leniency program can result in fewer cartels forming but those that form last a longer time.

## 6 Concluding Remarks

The main message of this paper for scholars and practitioners is that a proper assessment of the impact of a leniency program on the frequency of cartels requires taking account of how it influences enforcement through more standard non-leniency means. Holding non-leniency enforcement fixed, we find that a leniency program generally lowers the cartel rate, which is consistent with previous theoretical research. When non-leniency enforcement is endogenized, a leniency program can either lower or raise the cartel rate. Whether there are more or less cartels depends on the extent to which leniency applications shift competition authority resources away from pursuing cases without a leniency application to cases with a leniency application. In particular, introducing a leniency program into an environment in which penalties are low and leniency cases take up comparable resources to that of non-leniency cases is predicted to result in a higher frequency of cartels. Ensuring a leniency program has the desired effect of reducing the extent of collusion then requires proper setting

of complementary instruments. Specifically, penalties should be sufficiently severe and a procedure should be devised to handle leniency applications with a sufficiently small amount of resources (or the resources of the competition authority should be sufficiently increased).

While the focus in this paper has been on assessing the impact of a corporate leniency program, the framework is flexible enough to evaluate other competition policies. For example, what is the impact of screening for cartels, which can be represented as an increase in the probability that a cartel is discovered,  $q$ .<sup>28</sup> What is the impact of increasing resources for the competition authority, which can be modelled by increasing the probability of conviction function,  $p(\cdot)$ . If a leniency program is introduced, is the marginal productivity of screening and resources higher in which case these other policies should be pursued in conjunction with a leniency program. The model can also be enriched (though with some technical challenges to be surmounted) to give the competition authority more powers by, for example, allowing the penalty to depend on the actual duration of the cartel (rather than, implicitly, expected duration) or allowing the competition authority to observe an industry's type  $\eta$  and then select cases on those grounds; in particular, it would want to take cases with lower  $\eta$  because cartels tend to be more stable in those markets and thus shutting them down is more important. It could also prove insightful to endow the competition authority with some amount of resources and allow it to optimally allocate its resources across discovery (reducing  $q$ ), prosecution (increasing the probability of conviction), and penalization (increasing penalties).

Perhaps the most intriguing extension is to consider alternative assumptions to having the competition authority maximize social welfare (which, in the current model, means minimizing the cartel rate). If we think that the members of the competition authority act to maximize some observable measure of performance then, given that the cartel rate is not directly observed, it isn't clear there is an incentive scheme that will induce it to minimize the cartel rate. A more reasonable objective might be for it to maximize the number of convictions or the total amount of fines. In that case, the ultimate effect of a policy change - such as introducing a leniency program - will depend on how it is implemented by the competition authority. Better understanding the strategic behavior of a competition authority and how it impacts the efficacy of competition policy is an important avenue for future research.

---

<sup>28</sup>An argument for screening is provided in Harrington (2007).

## 7 Appendix A: Proofs

In reading the proofs, it is useful to have this summary of the solution algorithm for deriving an equilibrium.

1. Given  $\sigma$  and  $Y$  and for each  $\eta$ , solve for the maximum market condition (or threshold) for which the ICC is satisfied,  $\phi(Y, \sigma, \eta)$ .
2. Given  $\sigma$  and for each  $\eta$ , solve for the equilibrium collusive value  $Y^*(\sigma, \eta)$  which is a solution to the fixed point problem:  $Y^* = \psi(Y^*, \sigma, \eta)$ . If there are multiple fixed points, select the maximum. Given  $Y^*(\sigma, \eta)$ , define the equilibrium threshold  $\phi^*(\sigma, \eta)$ .
3. Given  $\sigma$  and  $\phi^*(\sigma, \eta)$ , derive the stationary proportion of type- $\eta$  industries that are cartels,  $C(\sigma, \eta)$ , and integrate over industry types to derive the stationary cartel rate:

$$C(\sigma) = \int_{\underline{\eta}}^{\bar{\eta}} C(\sigma, \eta) g(\eta) d\eta = \int_{\underline{\eta}}^{\bar{\eta}} \left( \frac{\kappa(1-\sigma)H(\phi^*(\sigma, \eta))}{1-(1-\kappa)(1-\sigma)H(\phi^*(\sigma, \eta))} \right) g(\eta) d\eta$$

4. Solve for the equilibrium probability of paying penalties through non-leniency enforcement  $\sigma^*$  which is a solution to the fixed point problem:  $\sigma^* = \Psi(\sigma^*)$ . If there are multiple fixed points, select the maximum. The equilibrium cartel rate is  $C(\sigma^*)$ .
5. (Optional) Solve for the optimal prosecution rate  $r^* : r^* \in \arg \min_{r \in [0,1]} C(\sigma^*(r))$ .

**Proof of Theorem 2.** The proof has three steps. First, holding  $Y$  fixed, the threshold for stable collusion is shown to be lower with a leniency program:  $\phi_{NL}(Y, \eta) > \phi_{\theta}(Y, \eta)$ . When  $\sigma > \theta$ , which holds by supposition, the deviator has lower penalties by applying for leniency and this tightens the ICC and thus raises the threshold. Second, given  $\phi_{NL}(Y, \eta) > \phi_{\theta}(Y, \eta)$  and the supposition that  $\omega > \sigma$ , it is shown that  $\psi_{NL}(Y, \sigma, \eta) \geq \psi_{\theta}(Y, \sigma, \eta)$ . That the collusive value function is weakly lower with a leniency program is due to two effects: i)  $\phi_{NL}(Y, \eta) > \phi_{\theta}(Y, \eta)$  results in weakly shorter cartel duration with a leniency program; and ii) when there is a leniency program, expected penalties upon cartel collapse are  $\omega\gamma(Y - \alpha\mu)$  rather than  $\sigma\gamma(Y - \alpha\mu)$ , and the former are higher when  $\omega > \sigma$ . Third,  $\psi_{NL}(Y, \sigma, \eta) > \psi_{\theta}(Y, \sigma, \eta)$  implies a weakly lower fixed point with a leniency program -  $Y_{NL}^*(\sigma, \eta) \geq Y_{\theta}^*(\sigma, \eta)$  - and, therefore, a weakly lower equilibrium threshold:  $\phi_{NL}^*(\sigma, \eta) \geq \phi_{\theta}^*(\sigma, \eta)$ . This proves the cartel rate is no higher with a leniency program. If, in addition, assumption A1 holds then  $\phi_{NL}^*(\sigma, \eta) > \phi_{\theta}^*(\sigma, \eta)$  for a positive measure of values for  $\eta$ . From this result, one can then conclude that, holding  $\sigma$  fixed, the cartel rate is strictly lower with a leniency program.

Holding  $Y$  fixed, the threshold function for stable collusion is lower with a leniency program:

$$\begin{aligned}
& \phi_{NL}(Y, \sigma, \eta) - \phi_{\theta}(Y, \sigma, \eta) \\
= & \frac{\delta(1-\sigma)(1-\kappa)(Y-\alpha\mu)}{(\eta-1)[1-\delta(1-\kappa)]} \\
& - \left( \frac{\delta(1-\sigma)(1-\kappa)(Y-\alpha\mu) - [1-\delta(1-\kappa)][\sigma - \min\{\sigma, \theta\}]\gamma(Y-\alpha\mu)}{(\eta-1)[1-\delta(1-\kappa)]} \right) \\
= & \frac{(\sigma - \theta)\gamma(Y-\alpha\mu)}{\eta-1} > 0
\end{aligned} \tag{15}$$

because  $\sigma > \theta$ . Using  $\phi_{NL}(Y, \eta) > \phi_{\theta}(Y, \eta)$ ,

$$\begin{aligned}
& \psi_{NL}(Y, \sigma, \eta) - \psi_{\theta}(Y, \sigma, \eta) \\
= & \int_{\underline{\pi}}^{\phi_{NL}(Y, \sigma, \eta)} \{(1-\delta)\pi + \delta[(1-\sigma)Y + \sigma W] - (1-\delta)\sigma\gamma(Y-\alpha\mu)\} h(\pi) d\pi \\
& + \int_{\phi_{NL}(Y, \sigma, \eta)}^{\bar{\pi}} [(1-\delta)\alpha\pi + \delta W - (1-\delta)\sigma\gamma(Y-\alpha\mu)] h(\pi) d\pi \\
& - \int_{\underline{\pi}}^{\phi_{\theta}(Y, \sigma, \eta)} \{(1-\delta)\pi + \delta[(1-\sigma)Y + \sigma W] - (1-\delta)\sigma\gamma(Y-\alpha\mu)\} h(\pi) d\pi \\
& - \int_{\phi_{\theta}(Y, \sigma, \eta)}^{\bar{\pi}} [(1-\delta)\alpha\pi + \delta W - \omega\gamma(Y-\alpha\mu)] h(\pi) d\pi
\end{aligned}$$

and a few manipulations yields

$$\begin{aligned}
& \psi_{NL}(Y, \sigma, \eta) - \psi_{\theta}(Y, \sigma, \eta) \\
= & \int_{\phi_{\theta}(Y, \sigma, \eta)}^{\phi_{NL}(Y, \sigma, \eta)} [(1-\delta)(1-\alpha)\pi + \delta(1-\sigma)(Y-W)] h(\pi) d\pi \\
& + \int_{\phi_{\theta}(Y, \sigma, \eta)}^{\bar{\pi}} (1-\delta)(\omega - \sigma)\gamma(Y-\alpha\mu) h(\pi) d\pi.
\end{aligned} \tag{16}$$

Given  $\omega > \sigma$ , (16) is non-negative. If  $\underline{\pi} \geq \phi_{NL}(Y, \sigma, \eta) (> \phi_{\theta}(Y, \sigma, \eta))$  or  $(\phi_{NL}(Y, \sigma, \eta) >) \phi_{\theta}(Y, \sigma, \eta) \geq \bar{\pi}$  then the first of the two terms in (16) is zero; otherwise, it is positive. If  $\phi_{\theta}(Y, \sigma, \eta) \geq \bar{\pi}$  then the second term is zero; otherwise, it is positive.

Since it has just been shown that  $\psi_{NL}(Y, \sigma, \eta) \geq \psi_{\theta}(Y, \sigma, \eta)$  then  $Y_{NL}^*(\sigma, \eta) \geq Y_{\theta}^*(\sigma, \eta)$ . Given

$$\phi^*(\sigma, \eta) \equiv \max\{\min\{\phi(Y^*(\sigma, \eta), \sigma, \eta), \bar{\pi}\}, \underline{\pi}\},$$

it follows that  $\phi_{NL}^*(\sigma, \eta) \geq \phi_{\theta}^*(\sigma, \eta)$ .

Next we want to show: if assumption A1 holds then  $\phi_{NL}^*(\sigma, \eta) > \phi_{\theta}^*(\sigma, \eta)$  for a positive measure of values for  $\eta$ . If  $\phi_{NL}^*(\sigma, \eta) < \bar{\pi}$  then either  $\phi_{NL}^*(\sigma, \eta) > \underline{\pi}$  - so that  $\phi_{NL}^*(\sigma, \eta) \in (\underline{\pi}, \bar{\pi})$  - or  $\phi_{NL}^*(\sigma, \eta) = \underline{\pi}$ ; and if  $\phi_{NL}^*(\sigma, \eta) > \underline{\pi}$  then either  $\phi_{NL}^*(\sigma, \eta) < \bar{\pi}$  - so that  $\phi_{NL}^*(\sigma, \eta) \in (\underline{\pi}, \bar{\pi})$  - or  $\phi_{NL}^*(\sigma, \eta) = \bar{\pi}$ . This results in

two mutually exclusive cases: 1) there is a positive measure of values for  $\eta$  such that  $\phi_{NL}^*(\sigma, \eta) \in (\underline{\pi}, \bar{\pi})$ ; and 2) there is not a positive measure of values for  $\eta$  such that  $\phi_{NL}^*(\sigma, \eta) \in (\underline{\pi}, \bar{\pi})$  in which case there is a positive measure of values for  $\eta$  such that  $\phi_{NL}^*(\sigma, \eta) = \underline{\pi}$  and a positive measure of values for  $\eta$  such that  $\phi_{NL}^*(\sigma, \eta) = \bar{\pi}$ .

In considering case (1), first note that

$$\phi_{NL}^*(\sigma, \eta) = \phi(Y_{NL}^*(\sigma, \eta), \sigma, \eta) > \phi_\theta(Y_{NL}^*(\sigma, \eta), \sigma, \eta) \geq \phi_\theta(Y_\theta^*(\sigma, \eta), \sigma, \eta), \quad (17)$$

where the equality follows from  $\phi_{NL}^*(\sigma, \eta) \in (\underline{\pi}, \bar{\pi})$ , the strict inequality follows from  $\phi_{NL}(Y, \sigma, \eta) > \phi_\theta(Y, \sigma, \eta)$ , and the weak inequality follows from  $Y_{NL}^*(\sigma, \eta) \geq Y_\theta^*(\sigma, \eta)$ . (17) implies  $\bar{\pi} > \phi_\theta(Y_\theta^*(\sigma, \eta), \sigma, \eta)$  and, therefore,

$$\phi_\theta^*(\sigma, \eta) = \max\{\phi_\theta(Y_\theta^*(\sigma, \eta), \sigma, \eta), \underline{\pi}\}. \quad (18)$$

(17)-(18) allow us to conclude:  $\phi_{NL}^*(\sigma, \eta) > \phi_\theta^*(\sigma, \eta)$ . Hence, for case (1), there is a positive measure of values for  $\eta$  for which  $\phi_{NL}^*(\sigma, \eta) > \phi_\theta^*(\sigma, \eta)$ . Under case (2), that  $\phi_{NL}^*(\sigma, \eta)$  is weakly decreasing in (see Lemma 2 in the Supplemental Appendix) implies  $\exists \hat{\eta}_{NL} \in (\underline{\eta}, \bar{\eta})$  such that

$$\phi_{NL}^*(\sigma, \eta) = \begin{cases} \bar{\pi} & \text{if } \eta \in [\underline{\eta}, \hat{\eta}_{NL}] \\ \underline{\pi} & \text{if } \eta \in (\hat{\eta}_{NL}, \bar{\eta}] \end{cases}. \quad (19)$$

Note that, at the critical value  $\hat{\eta}_{NL}$ ,

$$\psi_{NL}(Y, \sigma, \hat{\eta}_{NL}) \leq Y, \quad \forall Y \in [\alpha\mu, \mu], \quad (20)$$

for suppose not. Then  $\exists Y' \in (\alpha\mu, \mu)$  such that  $\psi_{NL}(Y', \sigma, \hat{\eta}_{NL}) > Y'$ . By the continuity of  $\psi_{NL}$  in  $\eta$ ,  $\exists \xi > 0$  such that  $\psi_{NL}(Y', \sigma, \hat{\eta}_{NL} + \xi) > Y'$  which implies  $Y_{NL}^*(\sigma, \hat{\eta}_{NL} + \xi) > \alpha\mu$  and  $\phi_{NL}^*(\sigma, \hat{\eta}_{NL} + \xi) > \underline{\pi}$ , but that contradicts (19). With (20) and  $\psi_{NL}(Y, \sigma, \eta) > \psi_\theta(Y, \sigma, \eta)$ , it follows  $\exists \chi > 0$  such that  $\psi_\theta(Y, \sigma, \hat{\eta}_{NL}) < Y - \chi \quad \forall Y \in [\alpha\mu, \mu]$  which implies, by the continuity of  $\psi_\theta$  in  $\eta$ ,  $\exists \hat{\eta}_\theta < \hat{\eta}_{NL}$  such that  $\phi_\theta^*(\sigma, \eta) = \underline{\pi}$  iff  $\eta > \hat{\eta}_\theta$ . We then have that there is a positive measure of values of  $\eta$  - specifically,  $\eta \in [\hat{\eta}_{NL}, \hat{\eta}_\theta)$  - for which

$$\phi_{NL}^*(\sigma, \eta) = \bar{\pi} > \underline{\pi} = \phi_\theta^*(\sigma, \eta).$$

This concludes the proof that: if assumption A1 holds then  $\phi_{NL}^*(\sigma, \eta) > \phi_\theta^*(\sigma, \eta)$  for positive measure of values for  $\eta$ .

Whether with or without a leniency program, if the threshold for a type- $\eta$  industry is  $\tilde{\phi}(\sigma, \eta)$  then the cartel rate is

$$\int_{\underline{\eta}}^{\bar{\eta}} \left[ \frac{\kappa(1-\sigma)H(\tilde{\phi}(\sigma, \eta))}{1 - (1-\kappa)(1-\sigma)H(\tilde{\phi}(\sigma, \eta))} \right] g(\eta) d\eta. \quad (21)$$

Note that the cartel rate is increasing in  $\tilde{\phi}(\sigma, \eta)$ . Given it has been shown  $\phi_{NL}^*(\sigma, \eta) \geq \phi_\theta^*(\sigma, \eta) \quad \forall \eta$ , (21) implies  $C_{NL}(\sigma) \geq C_\theta(\sigma)$ . It has also been shown that: if there is

a positive measure of values of  $\eta$  such that  $\phi_{NL}^*(\sigma, \eta) < \bar{\pi}$  and a positive measure of values for  $\eta$  such that  $\phi_{NL}^*(\sigma, \eta) > \underline{\pi}$  then there is a positive measure of values of  $\eta$  such that  $\phi_{NL}^*(\sigma, \eta) > \phi_\theta^*(\sigma, \eta)$  and, therefore,

$$\frac{\kappa(1-\sigma)H(\phi_{NL}^*(\sigma, \eta))}{1-(1-\kappa)(1-\sigma)H(\phi_{NL}^*(\sigma, \eta))} > \frac{\kappa(1-\sigma)H(\phi_\theta^*(\sigma, \eta))}{1-(1-\kappa)(1-\sigma)H(\phi_\theta^*(\sigma, \eta))}. \quad (22)$$

As (22) holds for a positive measure of values of  $\eta$ , (21) implies  $C_{NL}(\sigma) > C_\theta(\sigma)$ . ■

The existence of a fixed point to  $\Psi : [0, 1] \rightarrow [0, 1]$  is not immediate because there are two possible sources of discontinuity. Recall that  $\phi^*(\sigma, \eta)$  depends on  $Y^*(\sigma, \eta)$  which is the maximal fixed point to:  $Y = \psi(Y, \sigma, \eta)$ . Because of multiple fixed points to  $\psi(Y, \sigma, \eta)$ ,  $Y^*(\sigma, \eta)$  need not be continuous in  $\sigma$  and if  $Y^*(\sigma, \eta)$  is discontinuous then  $\phi^*(\sigma, \eta)$  is discontinuous which implies  $H(\phi^*(\sigma, \eta))$  and  $C(\sigma, \eta)$  in (12) are discontinuous. However, it is proven in Theorem 3 that these possible discontinuities in the integrand of  $\Psi$  do not create discontinuities in  $\Psi$ . The second possible source of discontinuity in  $\Psi$  is due to a discontinuity in expected penalties at  $\sigma = \theta$ . That discontinuity is present as long as  $\theta \in (0, 1)$  and, as a result, existence is established only when there is no leniency ( $\theta = 1$ ) and full leniency ( $\theta = 0$ ).

**Proof of Theorem 3.** When  $\theta = 1$  then

$$\Psi(\sigma) = qrp \left( qr \int_{\underline{\eta}}^{\bar{\eta}} C(\sigma, \eta) g(\eta) d\eta \right), \quad (23)$$

and when  $\theta = 0$  then

$$\begin{aligned} \Psi(\sigma) = & qrp \left( \lambda \int_{\underline{\eta}}^{\bar{\eta}} (1 - H(\phi^*(\sigma, \eta))) C(\sigma, \eta) g(\eta) d\eta \right. \\ & \left. + qr \int_{\underline{\eta}}^{\bar{\eta}} H(\phi^*(\sigma, \eta)) C(\sigma, \eta) g(\eta) d\eta \right). \end{aligned} \quad (24)$$

To show that a fixed point exists for (23) and for (24), the proof strategy has two steps: 1) show that, for any value of  $\sigma$ , the integrand in these equations is continuous in  $\sigma$  except for a countable set of values of  $\eta$ ; and 2) show that it follows from step 1 that  $\Psi$  is continuous. The proof will focus exclusively on proving that (24) has a fixed point as the method of proof is immediately applicable to the case of (23).<sup>29</sup>

Considering the integrand in (24), a discontinuity in

$$H(\phi^*(\sigma, \eta)) C(\sigma, \eta) g(\eta) = H(\phi^*(\sigma, \eta)) \left( \frac{\kappa(1-\sigma)H(\phi^*(\sigma, \eta))}{1-(1-\kappa)(1-\sigma)H(\phi^*(\sigma, \eta))} \right) g(\eta)$$

with respect to  $\sigma$  (or  $\eta$ ) comes from  $\phi^*(\sigma, \eta)$  being discontinuous, which comes from  $Y^*(\sigma, \eta)$  being discontinuous. Let  $\Delta(\sigma') \subseteq [\underline{\eta}, \bar{\eta}]$  be the set of  $\eta$  for which  $Y^*(\sigma, \eta)$  is discontinuous at  $\sigma = \sigma'$ . We will show that  $\Delta(\sigma)$  is countable.

<sup>29</sup>When  $\theta = 1$ , existence of a fixed point can also be established by showing that  $\Psi(\sigma)$  is non-decreasing in  $\sigma$  and appealing to Tarski's Fixed Point Theorem. However, when  $\theta < 1$ , it is generally not true that  $\Psi(\sigma)$  is non-decreasing in  $\sigma \forall \sigma$ .

Suppose  $Y^*(\sigma, \eta)$  is discontinuous in  $\sigma$  at  $(\sigma, \eta) = (\sigma', \eta')$ . Given  $\psi(Y, \sigma, \eta)$  is continuous and  $Y^*(\sigma, \eta)$  is the maximal fixed point to  $\psi(Y, \sigma, \eta)$  then

$$\psi(Y, \sigma', \eta') < Y, \forall Y \in (Y^*(\sigma', \eta'), \mu]. \quad (25)$$

If, in addition,  $\exists \xi > 0$  such that

$$\psi(Y, \sigma', \eta') > Y, \forall Y \in [Y^*(\sigma', \eta') - \xi, Y^*(\sigma', \eta')]$$

then, by the continuity of  $\psi(Y, \sigma, \eta)$  in  $\sigma$ ,  $Y^*(\sigma, \eta)$  is continuous at  $(\sigma, \eta) = (\sigma', \eta')$ , contrary to our supposition. Hence, it must be the case that  $\exists \xi > 0$  such that

$$\psi(Y, \sigma', \eta') \leq Y, \forall Y \in [Y^*(\sigma', \eta') - \xi, Y^*(\sigma', \eta')]. \quad (26)$$

Given that  $\psi(Y, \sigma, \eta)$  is continuous and decreasing in  $\eta$  (see the proof of Lemma 2 in the Supplemental Appendix) then (25) and (26) imply

$$\psi(Y, \sigma', \eta') < Y, \forall Y \in [Y^*(\sigma', \eta') - \xi, \mu], \quad \forall \eta > \eta'. \quad (27)$$

It follows from (27) that,  $\forall \eta > \eta'$ , all fixed points to  $\psi$  are bounded above by  $Y^*(\sigma', \eta') - \xi$ :

$$Y^*(\sigma', \eta) < Y^*(\sigma', \eta') - \xi, \quad \forall \eta > \eta'.$$

Next define:

$$\varepsilon(\sigma', \eta') \equiv Y^*(\sigma', \eta') - \lim_{\eta \downarrow \eta'} Y^*(\sigma', \eta)$$

where  $\varepsilon(\sigma', \eta')$  measures the size of the discontinuity in  $Y^*(\sigma', \eta)$  with respect to  $\eta$  at  $\eta = \eta'$ ; see Figure A1.

For each  $\eta \in \Delta(\sigma')$ , there has then been associated an interval of length  $\varepsilon(\sigma', \eta)$ . Note that these intervals have a null intersection because  $Y^*(\sigma, \eta)$  is non-increasing in  $\eta$ . Hence,

$$\sum_{\eta \in \Delta(\sigma')} \varepsilon(\sigma', \eta) \leq (1 - \alpha) \mu.$$

Given that a sum can only be finite if the number of elements which are positive is countable, it follows that  $\Delta(\sigma')$  is countable. Hence, the set of values for  $\eta$  for which  $Y^*(\sigma', \eta)$  is discontinuous in  $\sigma$  at  $\sigma = \sigma'$  is countable. This completes the first step.

By Jeffrey (1925), given that  $H(\phi^*(\sigma, \eta))C(\sigma, \eta)g(\eta)$  and  $(1 - H(\phi^*(\sigma, \eta)))C(\sigma, \eta)g(\eta)$  are bounded in  $(\sigma, \eta)$  on  $[0, 1] \times [\underline{\eta}, \bar{\eta}]$  and are continuous at  $\sigma = \sigma'$  for all  $\eta \in [\underline{\eta}, \bar{\eta}]$  except for a countable set then

$$\int_{\underline{\eta}}^{\bar{\eta}} H(\phi^*(\sigma, \eta))C(\sigma, \eta)g(\eta) d\eta$$

and

$$\int_{\underline{\eta}}^{\bar{\eta}} (1 - H(\phi^*(\sigma, \eta)))C(\sigma, \eta)g(\eta) d\eta$$

are continuous at  $\sigma = \sigma'$ . Given that  $p$  is a continuous function, it follows that

$$p \left( \lambda \int_{\underline{\eta}}^{\bar{\eta}} (1 - H(\phi^*(\sigma, \eta))) C(\sigma, \eta) g(\eta) d\eta + qr \int_{\underline{\eta}}^{\bar{\eta}} H(\phi^*(\sigma, \eta)) C(\sigma, \eta) g(\eta) d\eta \right)$$

is continuous in  $\sigma$ . Hence,  $\Psi$  in (24) is continuous in  $\sigma$  and maps  $[0, 1]$  into itself; therefore, a fixed point exists. The same method of proof can be used to show that a fixed point to (23) exists. ■

**Proof of Theorem 4.** The first step is to show that, as the penalty multiple  $\gamma$  goes to zero, the cartel rate function is the same with and without a leniency program:

$$\lim_{\gamma \rightarrow 0} C_{NL}(\sigma) = \lim_{\gamma \rightarrow 0} C_L(\sigma), \quad \forall \sigma.$$

The second step is to show that, as  $\gamma \rightarrow 0$ , non-leniency enforcement is weaker with a leniency program:

$$\lim_{\gamma \rightarrow 0} \sigma_{NL}^* > \lim_{\gamma \rightarrow 0} \sigma_L^*.$$

These two results together imply that the equilibrium cartel rate with a leniency program is higher than without a leniency program when  $\gamma \simeq 0$ .

For the first step, let us begin by considering the thresholds for stable collusion. Without a leniency program,

$$\phi_{NL}(Y, \sigma, \eta) = \frac{\delta(1 - \sigma)(1 - \kappa)(Y - \alpha\mu)}{(\eta - 1)[1 - \delta(1 - \kappa)]}$$

and, trivially,<sup>30</sup>

$$\lim_{\gamma \rightarrow 0} \phi_{NL}(Y, \sigma, \eta) = \frac{\delta(1 - \sigma)(1 - \kappa)(Y - \alpha\mu)}{(\eta - 1)[1 - \delta(1 - \kappa)]}.$$

With a full leniency program,

$$\phi_L(Y, \sigma, \eta) = \frac{\delta(1 - \sigma)(1 - \kappa)(Y - \alpha\mu) - [1 - \delta(1 - \kappa)]\sigma\gamma(Y - \alpha\mu)}{(\eta - 1)[1 - \delta(1 - \kappa)]}$$

and

$$\lim_{\gamma \rightarrow 0} \phi_L(Y, \sigma, \eta) = \frac{\delta(1 - \sigma)(1 - \kappa)(Y - \alpha\mu)}{(\eta - 1)[1 - \delta(1 - \kappa)]}.$$

Hence,

$$\lim_{\gamma \rightarrow 0} \phi_{NL}(Y, \sigma, \eta) = \lim_{\gamma \rightarrow 0} \phi_L(Y, \sigma, \eta). \quad (28)$$

---

<sup>30</sup>Recall that  $\phi_{NL}(Y, \sigma, \eta)$  comes out of the ICC and is the market condition that makes a firm indifferent between colluding and cheating. That  $\gamma$  does not matter is because the expected penalty is the same whether a firm sets the collusive price or cheats and undercuts the collusive price set by the other firms.



Turning to the collusive value functions, we have without a leniency program:

$$\begin{aligned}\psi_{NL}(Y, \sigma, \eta) &= \int_{\underline{\pi}}^{\phi_{NL}(Y, \sigma, \eta)} [(1 - \delta) \pi + \delta (1 - \sigma) Y + \delta \sigma W] h(\pi) d\pi \\ &\quad + \int_{\phi_{NL}(Y, \sigma, \eta)}^{\bar{\pi}} [(1 - \delta) \alpha \pi + \delta W] h(\pi) d\pi - (1 - \delta) \sigma \gamma (Y - \alpha \mu),\end{aligned}$$

and with a full leniency program:

$$\begin{aligned}\psi_L(Y, \sigma, \eta) &= \int_{\underline{\pi}}^{\phi_L(Y, \sigma, \eta)} [(1 - \delta) \pi + \delta (1 - \sigma) Y + \delta \sigma W - (1 - \delta) \sigma \gamma (Y - \alpha \mu)] h(\pi) d\pi \\ &\quad + \int_{\phi_L(Y, \sigma, \eta)}^{\bar{\pi}} [(1 - \delta) \alpha \pi + \delta W - (1 - \delta) \omega \gamma (Y - \alpha \mu)] h(\pi) d\pi.\end{aligned}$$

Using (28),

$$\begin{aligned}\lim_{\gamma \rightarrow 0} \psi_{NL}(Y, \sigma, \eta) &= \lim_{\gamma \rightarrow 0} \psi_L(Y, \sigma, \eta) \tag{29} \\ &= \int_{\underline{\pi}}^{\phi_{NL}(Y, \sigma, \eta)} [(1 - \delta) \pi + \delta (1 - \sigma) Y + \delta \sigma W] h(\pi) d\pi \\ &\quad + \int_{\phi_{NL}(Y, \sigma, \eta)}^{\bar{\pi}} [(1 - \delta) \alpha \pi + \delta W] h(\pi) d\pi.\end{aligned}$$

Generically, (29) implies

$$\lim_{\gamma \rightarrow 0} Y_{NL}^*(\sigma, \eta) = \lim_{\gamma \rightarrow 0} Y_L^*(\sigma, \eta). \tag{30}$$

(It is only generic because it requires that, in an  $\varepsilon$ -ball around  $\gamma = 0$ ,  $Y_{NL}^*(\sigma, \eta)$  and  $Y_L^*(\sigma, \eta)$  are continuous in  $\gamma$ .) It follows from (28) and (29) that:

$$\lim_{\gamma \rightarrow 0} \phi_{NL}^*(\sigma, \eta) = \lim_{\gamma \rightarrow 0} \phi_L^*(\sigma, \eta). \tag{31}$$

Given  $\sigma$ , the cartel rate without and with a leniency program, respectively, is:

$$\begin{aligned}C_{NL}(\sigma) &= \int_{\underline{\eta}}^{\bar{\eta}} C_{NL}(\sigma, \eta) g(\eta) d\eta = \int_{\underline{\eta}}^{\bar{\eta}} \left[ \frac{\kappa (1 - \sigma) H(\phi_{NL}^*(\sigma, \eta))}{1 - (1 - \kappa)(1 - \sigma) H(\phi_{NL}^*(\sigma, \eta))} \right] g(\eta) d\eta \\ C_L(\sigma) &= \int_{\underline{\eta}}^{\bar{\eta}} C_L(\sigma, \eta) g(\eta) d\eta = \int_{\underline{\eta}}^{\bar{\eta}} \left[ \frac{\kappa (1 - \sigma) H(\phi_L^*(\sigma, \eta))}{1 - (1 - \kappa)(1 - \sigma) H(\phi_L^*(\sigma, \eta))} \right] g(\eta) d\eta.\end{aligned}$$

Using (31),

$$\lim_{\gamma \rightarrow 0} C_{NL}(\sigma) = \lim_{\gamma \rightarrow 0} C_L(\sigma). \tag{32}$$

To prove the second step, we want to first show that, when  $\lambda \simeq 1$  and  $\gamma \simeq 0$ ,

$$\begin{aligned}& p \left( qr \int_{\underline{\eta}}^{\bar{\eta}} C_{NL}(\sigma, \eta) g(\eta) d\eta \right) \tag{33} \\ & > p \left( \lambda \int_{\underline{\eta}}^{\bar{\eta}} (1 - H(\phi_L^*(\sigma, \eta))) C_L(\sigma, \eta) g(\eta) d\eta + qr \int_{\underline{\eta}}^{\bar{\eta}} H(\phi_L^*(\sigma, \eta)) C_L(\sigma, \eta) g(\eta) d\eta \right).\end{aligned}$$

Given  $p$  is strictly decreasing, (33) holds iff

$$\begin{aligned} & \lambda \int_{\underline{\eta}}^{\bar{\eta}} (1 - H(\phi_L^*(\sigma, \eta))) C_L(\sigma, \eta) g(\eta) d\eta + qr \int_{\underline{\eta}}^{\bar{\eta}} H(\phi_L^*(\sigma, \eta)) C_L(\sigma, \eta) g(\eta) d\eta \\ & > qr \int_{\underline{\eta}}^{\bar{\eta}} C_{NL}(\sigma, \eta) g(\eta) d\eta \end{aligned}$$

or, equivalently,

$$\begin{aligned} & \int_{\underline{\eta}}^{\bar{\eta}} (1 - H(\phi_L^*(\sigma, \eta))) [\lambda C_L(\sigma, \eta) - qr C_{NL}(\sigma, \eta)] g(\eta) d\eta \quad (34) \\ & > qr \int_{\underline{\eta}}^{\bar{\eta}} H(\phi_L^*(\sigma, \eta)) [C_{NL}(\sigma, \eta) - C_L(\sigma, \eta)] g(\eta) d\eta. \end{aligned}$$

Given (32), (34) holds as  $\gamma \rightarrow 0$  iff

$$(\lambda - qr) \int_{\underline{\eta}}^{\bar{\eta}} (1 - H(\phi_{NL}^*(\sigma, \eta))) C_{NL}(\sigma, \eta) g(\eta) d\eta > 0. \quad (35)$$

By (13),

$$\int_{\underline{\eta}}^{\bar{\eta}} (1 - H(\phi_{NL}^*(\sigma, \eta))) C_{NL}(\sigma, \eta) g(\eta) d\eta > 0 \quad (36)$$

holds for  $\sigma = \sigma_{NL}^*$ . Given  $qr < 1$  then  $\lambda \simeq 1$  implies  $\lambda > qr$  and (35) holds. We have then shown that there exists  $\hat{\lambda} < 1$  and  $\hat{\gamma} > 0$  such that if  $(\gamma, \lambda) \in [0, \hat{\gamma}] \times [\hat{\lambda}, 1]$  then (33) holds, generically, in a small neighborhood of  $\sigma = \sigma_{NL}^*$ .

For when there is no leniency program,  $\sigma_{NL}^*$  is defined by:

$$\sigma_{NL}^* = qrp \left( qr \int_{\underline{\eta}}^{\bar{\eta}} C_{NL}(\sigma_{NL}^*, \eta) g(\eta) d\eta \right).$$

As it is the maximal fixed point then:

$$\sigma - qrp \left( qr \int_{\underline{\eta}}^{\bar{\eta}} C_{NL}(\sigma, \eta) g(\eta) d\eta \right) \geq 0 \text{ as } \sigma \geq \sigma_{NL}^*. \quad (37)$$

Hence, using (33), it follows from (37) that there exists  $\hat{\lambda} < 1$  and  $\hat{\gamma} > 0$  such that if  $(\gamma, \lambda) \in [0, \hat{\gamma}] \times [\hat{\lambda}, 1]$  then  $\exists \varepsilon > 0$  such that

$$\begin{aligned} & \sigma - qrp \left( \lambda \int_{\underline{\eta}}^{\bar{\eta}} (1 - H(\phi_L^*(\sigma, \eta))) C_L(\sigma, \eta) g(\eta) d\eta \right. \quad (38) \\ & \quad \left. + qr \int_{\underline{\eta}}^{\bar{\eta}} H(\phi_L^*(\sigma, \eta)) C_L(\sigma, \eta) g(\eta) d\eta \right) \\ & > 0, \forall \sigma \geq \sigma_{NL}^* - \varepsilon. \end{aligned}$$

Given the continuity of

$$p \left( \lambda \int_{\underline{\eta}}^{\bar{\eta}} (1 - H(\phi_L^*(\sigma, \eta))) C_L(\sigma, \eta) g(\eta) d\eta + qr \int_{\underline{\eta}}^{\bar{\eta}} H(\phi_L^*(\sigma, \eta)) C_L(\sigma, \eta) g(\eta) d\eta \right)$$

in  $\sigma$  (see the proof of Theorem 3), (38) implies the maximal fixed point  $\sigma_L^*$  is less than  $\sigma_{NL}^* - \varepsilon$ . Given (32) and having just shown

$$\lim_{\gamma \rightarrow 0} \sigma_{NL}^* > \lim_{\gamma \rightarrow 0} \sigma_L^*,$$

it follows that

$$\lim_{\gamma \rightarrow 0} C_L(\sigma_L^*) > \lim_{\gamma \rightarrow 0} C_{NL}(\sigma_{NL}^*).$$

■

**Proof of Theorem 6.** Given  $\sigma^* \in (0, \omega)$  and  $\theta = 0$ , by Theorem 2 we have that  $C_{NL}(\sigma) \geq C_L(\sigma)$  and, when there is positive measure of values for  $\eta$  such that  $\phi_{NL}^*(\sigma, \eta) < \bar{\pi}$  and a positive measure of values for  $\eta$  such that  $\phi_{NL}^*(\sigma, \eta) > \underline{\pi}$ ,  $C_{NL}(\sigma) > C_L(\sigma)$ . To prove this theorem, it is then sufficient to show  $\sigma_L^* > \sigma_{NL}^*$ .

$\sigma_{NL}^*$  and  $\sigma_L^*$  are defined by:

$$\sigma_{NL}^* = qrp \left( qr \int_{\underline{\eta}}^{\bar{\eta}} C_{NL}(\sigma_{NL}^*, \eta) g(\eta) d\eta \right)$$

where

$$C_{NL}(\sigma, \eta) = \frac{\kappa(1 - \sigma) H(\phi_{NL}^*(\sigma, \eta))}{1 - (1 - \kappa)(1 - \sigma) H(\phi_{NL}^*(\sigma, \eta))},$$

and

$$\begin{aligned} \sigma_L^* &= p \left( \lambda \int_{\underline{\eta}}^{\bar{\eta}} (1 - H(\phi_L^*(\sigma_L^*, \eta))) C_L(\sigma_L^*, \eta) g(\eta) d\eta \right. \\ &\quad \left. + qr \int_{\underline{\eta}}^{\bar{\eta}} H(\phi_L^*(\sigma_L^*, \eta)) C_L(\sigma_L^*, \eta) g(\eta) d\eta \right) \end{aligned}$$

where

$$C_L(\sigma, \eta) = \frac{\kappa(1 - \sigma) H(\phi_L^*(\sigma, \eta))}{1 - (1 - \kappa)(1 - \sigma) H(\phi_L^*(\sigma, \eta))}.$$

If  $\phi_{NL}^*(\sigma, \eta) > (=) \phi_L^*(\sigma, \eta)$  then

$$H(\phi_{NL}^*(\sigma, \eta)) > (=) H(\phi_L^*(\sigma, \eta))$$

and

$$C_{NL}(\sigma, \eta) > (=) C_L(\sigma, \eta),$$

in which case,

$$H(\phi_{NL}^*(\sigma, \eta)) C_{NL}(\sigma, \eta) > (=) H(\phi_L^*(\sigma, \eta)) C_L(\sigma, \eta).$$

It is immediate that if

$$H(\phi_{NL}^*(\sigma, \eta)) C_{NL}(\sigma, \eta) \geq H(\phi_L^*(\sigma, \eta)) C_L(\sigma, \eta), \forall \eta \quad (39)$$

and

$$H(\phi_{NL}^*(\sigma, \eta)) C_{NL}(\sigma, \eta) > H(\phi_L^*(\sigma, \eta)) C_L(\sigma, \eta), \text{ for positive measure of } \eta \quad (40)$$

then

$$\int_{\underline{\eta}}^{\bar{\eta}} H(\phi_{NL}^*(\sigma, \eta)) C_{NL}(\sigma, \eta) g(\eta) d\eta > \int_{\underline{\eta}}^{\bar{\eta}} H(\phi_L^*(\sigma, \eta)) C_L(\sigma, \eta) g(\eta) d\eta. \quad (41)$$

(39) is always true and (40) is true when there is positive measure of values for  $\eta$  such that  $\phi_{NL}^*(\sigma, \eta) < \bar{\pi}$  and a positive measure of values for  $\eta$  such that  $\phi_{NL}^*(\sigma, \eta) > \underline{\pi}$ .

Evaluate (41) at  $\sigma = \sigma_{NL}^*$ :

$$\int_{\underline{\eta}}^{\bar{\eta}} H(\phi_{NL}^*(\sigma_{NL}^*, \eta)) C_{NL}(\sigma_{NL}^*, \eta) g(\eta) d\eta > \int_{\underline{\eta}}^{\bar{\eta}} H(\phi_L^*(\sigma_{NL}^*, \eta)) C_L(\sigma_{NL}^*, \eta) g(\eta) d\eta. \quad (42)$$

Noting that  $\sigma_{NL}^*$  does not depend on  $\lambda$ , if  $\lambda$  is sufficiently small then it follows from (42):

$$\begin{aligned} & qr \int_{\underline{\eta}}^{\bar{\eta}} H(\phi_{NL}^*(\sigma_{NL}^*, \eta)) C_{NL}(\sigma_{NL}^*, \eta) g(\eta) d\eta \\ & > \lambda \int_{\underline{\eta}}^{\bar{\eta}} (1 - H(\phi_L^*(\sigma_{NL}^*, \eta))) C_L(\sigma_{NL}^*, \eta) g(\eta) d\eta + \\ & qr \int_{\underline{\eta}}^{\bar{\eta}} H(\phi_L^*(\sigma_{NL}^*, \eta)) C_L(\sigma_{NL}^*, \eta) g(\eta) d\eta \end{aligned} \quad (43)$$

Given that  $p$  is decreasing then (43) implies (when  $\lambda$  is sufficiently small):

$$\begin{aligned} & qrp \left( \lambda \int_{\underline{\eta}}^{\bar{\eta}} (1 - H(\phi_L^*(\sigma_{NL}^*, \eta))) C_L(\sigma_{NL}^*, \eta) g(\eta) d\eta \right. \\ & \left. + qr \int_{\underline{\eta}}^{\bar{\eta}} H(\phi_L^*(\sigma_{NL}^*, \eta)) C_L(\sigma_{NL}^*, \eta) g(\eta) d\eta \right) \\ & > qrp \left( qr \int_{\underline{\eta}}^{\bar{\eta}} H(\phi_{NL}^*(\sigma_{NL}^*, \eta)) C_{NL}(\sigma_{NL}^*, \eta) g(\eta) d\eta \right) \\ & > qrp \left( qr \int_{\underline{\eta}}^{\bar{\eta}} C_{NL}(\sigma_{NL}^*, \eta) g(\eta) d\eta \right) = \sigma_{NL}^*. \end{aligned}$$

Hence,

$$\begin{aligned}
& qrp \left( \lambda \int_{\underline{\eta}}^{\bar{\eta}} (1 - H(\phi_L^*(\sigma_{NL}^*, \eta))) C_L(\sigma_{NL}^*, \eta) g(\eta) d\eta \right. \\
& \quad \left. + qr \int_{\underline{\eta}}^{\bar{\eta}} H(\phi_L^*(\sigma_{NL}^*, \eta)) C_L(\sigma_{NL}^*, \eta) g(\eta) d\eta \right) \\
& > \sigma_{NL}^*
\end{aligned}$$

and thus  $\sigma_L^* > \sigma_{NL}^*$ .

In proving  $\sigma_L^* > \sigma_{NL}^*$ , the preceding analysis presumed  $\omega > \sigma$ . If, contrary to that presumption,  $\sigma_L^* \geq \omega$  then the supposition that  $\omega > \sigma_{NL}^*$  would again imply  $\sigma_L^* > \sigma_{NL}^*$ . ■

**Proof of Theorem 7.** Given that  $\sigma = qrs$  and  $r, s \in [0, 1]$  (hence, are bounded), it is immediate that

$$\lim_{q \rightarrow 0} \sigma_{NL}^* = 0, \lim_{q \rightarrow 0} \sigma_L^* = 0,$$

which implies

$$\lim_{q \rightarrow 0} C_{NL}(\sigma_{NL}^*) = \lim_{\sigma \rightarrow 0} C_{NL}(\sigma), \quad \lim_{q \rightarrow 0} C_L(\sigma_L^*) = \lim_{\sigma \rightarrow 0} C_L(\sigma).$$

To show the equilibrium cartel rate is lower with a leniency program, it is then sufficient to prove:

$$\lim_{\sigma \rightarrow 0} C_{NL}(\sigma) > \lim_{\sigma \rightarrow 0} C_L(\sigma). \quad (44)$$

Given  $\theta = 0 < \omega$  then  $\sigma \in (\theta, \omega)$  holds as  $\sigma \rightarrow 0$  in which case Theorem 6 proves (44). ■

**Proof of Theorem 8.** If  $C(\sigma) > 0$  then  $\hat{\eta}(\sigma) > \underline{\eta}$  and  $Y^*(\sigma, \eta) > \alpha\mu \forall \eta \in (1, \hat{\eta}(\sigma)]$ . Furthermore, since  $Y^*(\sigma, \hat{\eta}(\sigma)) > \alpha\mu$  and  $Y^*(\sigma, \eta)$  is non-increasing in  $\eta$  (see Lemma 2 in the Supplemental Appendix) then

$$\lim_{\eta \rightarrow 1} Y^*(\sigma, \eta) > \alpha\mu.$$

Recall

$$\begin{aligned}
\phi(Y, \sigma, \eta) &= \frac{\delta(1-\sigma)(1-\kappa)(Y-\alpha\mu) - [1-\delta(1-\kappa)][\sigma - \min\{\sigma, \theta\}]\gamma(Y-\alpha\mu)}{(\eta-1)[1-\delta(1-\kappa)]} \\
&= \frac{\{\delta(1-\sigma)(1-\kappa) - [1-\delta(1-\kappa)][\sigma - \min\{\sigma, \theta\}]\gamma\}(Y-\alpha\mu)}{(\eta-1)[1-\delta(1-\kappa)]}
\end{aligned}$$

and

$$\phi^*(\sigma, \eta) \equiv \max\{\min\{\phi(Y^*(\sigma, \eta), \sigma, \eta), \bar{\pi}\}, \underline{\pi}\}$$

where this encompasses both the case of a full leniency program ( $\theta = 0$ ) and no leniency program ( $\theta = 1$ ). Given that  $Y^*(\sigma, \eta)$  is bounded above  $\alpha\mu$  as  $\eta \rightarrow 1$  then

$$\begin{aligned}
\lim_{\eta \rightarrow 1} \phi(Y^*(\sigma, \eta), \sigma, \eta) &= \lim_{\eta \rightarrow 1} \frac{\{\delta(1-\sigma)(1-\kappa) - [1-\delta(1-\kappa)][\sigma - \min\{\sigma, \theta\}]\gamma\}(Y^*(\sigma, \eta) - \alpha\mu)}{(\eta-1)[1-\delta(1-\kappa)]} \\
&= +\infty
\end{aligned}$$

and, therefore,

$$\lim_{\eta \rightarrow 1} H(\phi^*(\sigma, \eta)) = \lim_{\eta \rightarrow 1} H(\max\{\min\{\phi(Y^*(\sigma, \eta), \sigma, \eta), \bar{\pi}\}, \underline{\pi}\}) = 1.$$

Thus, when  $\eta$  is close to one, if a stable cartel forms (that is,  $\phi^*(\sigma, \eta) < \bar{\pi}$ ) then it is fully stable (that is,  $\phi^*(\sigma, \eta) = \bar{\pi}$ ).

Next note that

$$C(\sigma, \eta) = \frac{\kappa(1-\sigma)H(\phi^*(\sigma, \eta))}{1-(1-\kappa)(1-\sigma)H(\phi^*(\sigma, \eta))}$$

and, therefore,

$$\lim_{\eta \rightarrow 1} C(\sigma, \eta) = \lim_{\eta \rightarrow 1} \frac{\kappa(1-\sigma)H(\phi^*(\sigma, \eta))}{1-(1-\kappa)(1-\sigma)H(\phi^*(\sigma, \eta))} = \frac{\kappa(1-\sigma)}{1-(1-\kappa)(1-\sigma)}.$$

We then have:

$$\begin{aligned} & \lim_{\eta \rightarrow 1} [C_L(\sigma_L^*, \eta) - C_{NL}(\sigma_{NL}^*, \eta)] \\ &= \frac{\kappa(1-\sigma_L^*)}{1-(1-\kappa)(1-\sigma_L^*)} - \frac{\kappa(1-\sigma_{NL}^*)}{1-(1-\kappa)(1-\sigma_{NL}^*)} \\ &= \kappa \left[ \frac{(1-\sigma_L^*)[1-(1-\kappa)(1-\sigma_{NL}^*)] - (1-\sigma_{NL}^*)[1-(1-\kappa)(1-\sigma_L^*)]}{[1-(1-\kappa)(1-\sigma_L^*)][1-(1-\kappa)(1-\sigma_{NL}^*)]} \right] \\ &= \frac{\kappa(\sigma_{NL}^* - \sigma_L^*)}{[1-(1-\kappa)(1-\sigma_L^*)][1-(1-\kappa)(1-\sigma_{NL}^*)]}. \end{aligned}$$

■

**Proof of Theorem 9.** Let us first show: if  $\sigma \in (0, \omega)$  and  $\hat{\eta}_{NL}(\sigma) \in (\underline{\eta}, \bar{\eta})$  then  $\hat{\eta}_L(\sigma) < \hat{\eta}_{NL}(\sigma)$ . By the definition of  $\hat{\eta}$ , if  $\hat{\eta} \in (\underline{\eta}, \bar{\eta})$  then there is a (maximal) fixed point in  $Y$  of  $\psi(Y, \sigma, \eta)$  such that  $Y > \alpha\mu$  for  $\eta = \hat{\eta}$  but not for  $\eta > \hat{\eta}$ :

$$\begin{aligned} \exists Y^*(\sigma, \hat{\eta}) &\in (\alpha\mu, \mu] \text{ such that } Y \leq \psi(Y, \sigma, \hat{\eta}) \text{ as } Y \geq Y^*(\sigma, \hat{\eta}) & (45) \\ \nexists Y &\in (\alpha\mu, \mu] \text{ such that } Y = \psi(Y, \sigma, \eta), \forall \eta \in (\hat{\eta}, \bar{\eta}]. \end{aligned}$$

Recall that

$$\psi(Y, \sigma, \eta) = \begin{cases} \int_{\underline{\pi}}^{\phi(Y, \sigma, \eta)} \{(1-\delta)\pi + \delta[(1-\sigma)Y + \sigma W] - (1-\delta)\sigma\gamma(Y - \alpha\mu)\} h(\pi) d\pi & \text{if } \sigma \leq \theta \\ + \int_{\phi(Y, \sigma, \eta)}^{\bar{\pi}} [(1-\delta)\alpha\pi + \delta W - (1-\delta)\sigma\gamma(Y - \alpha\mu)] h(\pi) d\pi & \\ \int_{\underline{\pi}}^{\phi(Y, \sigma, \eta)} \{(1-\delta)\pi + \delta[(1-\sigma)Y + \sigma W] - (1-\delta)\sigma\gamma(Y - \alpha\mu)\} h(\pi) d\pi & \text{if } \theta < \sigma \\ + \int_{\phi(Y, \sigma, \eta)}^{\bar{\pi}} [(1-\delta)\alpha\pi + \delta W - (1-\delta)\omega\gamma(Y - \alpha\mu)] h(\pi) d\pi & \end{cases}$$

and

$$\phi(Y, \sigma, \eta) = \frac{\delta(1-\sigma)(1-\kappa)(Y - \alpha\mu) - [1 - \delta(1-\kappa)][\sigma - \min\{\sigma, \theta\}]\gamma(Y - \alpha\mu)}{(\eta - 1)[1 - \delta(1-\kappa)]}.$$

Let us next argue that

$$\phi(Y^*(\sigma, \hat{\eta}), \sigma, \hat{\eta}) \in (\underline{\pi}, \bar{\pi}]. \quad (46)$$

Obviously,  $Y^*(\sigma, \hat{\eta}) > \alpha\mu$  implies  $\phi(Y^*(\sigma, \hat{\eta}), \sigma, \eta) > \underline{\pi}$ . If  $\phi(Y^*(\sigma, \hat{\eta}), \sigma, \eta) > \bar{\pi}$  then, by the continuity of  $\phi(Y, \sigma, \eta)$  in  $\eta$ , it follows that  $\exists \varepsilon > 0$  such that  $\phi(Y^*(\sigma, \hat{\eta}), \sigma, \eta) > \bar{\pi} \forall \eta \in (\hat{\eta}, \hat{\eta} + \varepsilon)$ . Given that  $\eta$  affects  $\psi(Y, \sigma, \eta)$  only through  $\phi(Y, \sigma, \eta)$  - and recalling that  $H(\bar{\pi}) = 1$  - then

$$\psi(Y^*(\sigma, \hat{\eta}), \sigma, \hat{\eta}) = \psi(Y^*(\sigma, \hat{\eta}), \sigma, \eta) \forall \eta \in (\hat{\eta}, \hat{\eta} + \varepsilon)$$

which implies  $Y^*(\sigma, \hat{\eta})$  is a fixed point to  $\psi(Y, \sigma, \eta) \forall \eta \in (\hat{\eta}, \hat{\eta} + \varepsilon)$  which contradicts (45). We then conclude  $\phi(Y^*(\sigma, \hat{\eta}), \sigma, \eta) \leq \bar{\pi}$  and (46) is true.

Using (46) for when there is no leniency program,  $\hat{\eta}_{NL}(\sigma) \in (\underline{\eta}, \bar{\eta})$  implies  $\phi_{NL}(Y_{NL}^*(\sigma, \hat{\eta}_{NL}), \sigma, \hat{\eta}_{NL}) \in (\underline{\pi}, \bar{\pi}]$ . Since  $\phi$  is increasing in  $Y$ , it then follows:

$$\bar{\pi} > \phi_{NL}(Y, \sigma, \hat{\eta}_{NL}), \forall Y \in [\alpha\mu, Y_{NL}^*(\sigma, \hat{\eta}_{NL})]. \quad (47)$$

In the proof of Theorem 2 it was shown: if  $\sigma > 0$  then  $\phi_{NL}(Y, \sigma, \eta) > \phi_L(Y, \sigma, \eta)$ . Given it is assumed  $\sigma > 0$ , (47) then implies

$$\bar{\pi} \geq \phi_{NL}(Y, \sigma, \hat{\eta}_{NL}) > \phi_L(Y, \sigma, \hat{\eta}_{NL}), \forall Y \in [\alpha\mu, Y_{NL}^*(\sigma, \hat{\eta}_{NL})] \quad (48)$$

Now consider:

$$\begin{aligned} & \psi_{NL}(Y, \sigma, \eta) - \psi_L(Y, \sigma, \eta) \\ = & \int_{\underline{\pi}}^{\phi_{NL}(Y, \sigma, \eta)} \{(1 - \delta)\pi + \delta[(1 - \sigma)Y + \sigma W] - (1 - \delta)\sigma\gamma(Y - \alpha\mu)\} h(\pi) d\pi \\ & + \int_{\phi_{NL}(Y, \sigma, \eta)}^{\bar{\pi}} [(1 - \delta)\alpha\pi + \delta W - (1 - \delta)\sigma\gamma(Y - \alpha\mu)] h(\pi) d\pi \\ & - \int_{\underline{\pi}}^{\phi_L(Y, \sigma, \eta)} \{(1 - \delta)\pi + \delta[(1 - \sigma)Y + \sigma W] - (1 - \delta)\sigma\gamma(Y - \alpha\mu)\} h(\pi) d\pi \\ & - \int_{\phi_L(Y, \sigma, \eta)}^{\bar{\pi}} [(1 - \delta)\alpha\pi + \delta W - (1 - \delta)\omega\gamma(Y - \alpha\mu)] h(\pi) d\pi \end{aligned}$$

After some simplifying steps:

$$\begin{aligned} & \psi_{NL}(Y, \sigma, \eta) - \psi_L(Y, \sigma, \eta) \\ = & \int_{\phi_L(Y, \sigma, \eta)}^{\phi_{NL}(Y, \sigma, \eta)} \{(1 - \delta)(1 - \alpha)\pi + \delta(1 - \sigma)(Y - W) - (1 - \delta)(\sigma - \omega)\gamma(Y - \alpha\mu)\} h(\pi) d\pi \\ & + \int_{\phi_{NL}(Y, \sigma, \eta)}^{\bar{\pi}} [(1 - \delta)\alpha\pi + \delta W - (1 - \delta)(\sigma - \omega)\gamma(Y - \alpha\mu)] h(\pi) d\pi. \end{aligned}$$

Given  $\sigma \in (0, \omega)$  and using (48), we have:

$$\psi_{NL}(Y, \sigma, \hat{\eta}_{NL}) - \psi_L(Y, \sigma, \hat{\eta}_{NL}) > 0. \quad (49)$$

Next note that it follows from  $\hat{\eta}_{NL} \in (\underline{\eta}, \bar{\eta})$  that:

$$\begin{aligned}\psi_{NL}(Y, \sigma, \hat{\eta}_{NL}) &\leq Y \quad \forall Y \in [\alpha\mu, Y_{NL}^*(\sigma, \hat{\eta}_{NL})] \\ \psi_{NL}(Y, \sigma, \hat{\eta}_{NL}) &< Y \quad \forall Y \in (Y_{NL}^*(\sigma, \hat{\eta}_{NL}), \mu].\end{aligned}$$

Using (49), this implies

$$\begin{aligned}\psi_L(Y, \sigma, \hat{\eta}_{NL}) &< Y \quad \forall Y \in [\alpha\mu, Y^*(\sigma, \hat{\eta}_{NL})] \\ \psi_L(Y, \sigma, \hat{\eta}_{NL}) &< Y \quad \forall Y \in (Y^*(\sigma, \hat{\eta}_{NL}), \mu],\end{aligned}$$

and, therefore,  $\hat{\eta}_L(\sigma) < \hat{\eta}_{NL}(\sigma)$ .

We have thus far shown:  $\hat{\eta}_L(\sigma_{NL}^*) < \hat{\eta}_{NL}(\sigma_{NL}^*)$ . Given that  $\hat{\eta}(\sigma)$  is non-increasing in  $\sigma$  (see Lemma 3 in the Supplemental Appendix), if  $\sigma_L^* \geq \sigma_{NL}^*$  then  $\hat{\eta}_L(\sigma_L^*) \leq \hat{\eta}_L(\sigma_{NL}^*)$  which then implies  $\hat{\eta}_L(\sigma_L^*) < \hat{\eta}_{NL}(\sigma_{NL}^*)$ . ■

## 8 Appendix B: Numerical Methods<sup>31</sup>

There are 9 parameters in the general model:  $n, \alpha, \omega, \theta, \kappa, \delta, q, \gamma$ , and  $\lambda$ . The baseline simulation assumes:  $(n, \alpha, \omega, \theta, \kappa, \delta, q, \gamma, \lambda) = (4, 0, .75, 0 \text{ or } 1, .05, .85, .2, .5, 1)$ , where  $\theta = 0$  with leniency program and  $\theta = 1$  without leniency program.

For the probability of conviction function, we consider two functional forms:

$$p(\lambda L + R) = \begin{cases} \max\{c - m(\lambda L + R), 0.05\}, & \text{where } c < 1, \\ \frac{\tau}{\xi + v(\lambda L + R)^\rho}, & \text{where } v > 0, \rho \geq 1, \tau \in (0, 1], \xi \geq \tau \end{cases}$$

For the first specification, the probability decreases linearly with the caseload until it reaches its minimum value of 0.05. The second specification assumes a concave then convex relationship between caseload and the probability of success. The baseline simulation assumes  $(c, m) = (.8, 40)$  for the linear specification and  $(\tau, \xi, v, \rho) = (1, 1, 1000, 1.4)$  for the non-linear specification.

We assume a log-normal distribution,  $LN(\mu, \sigma^2)$ , for the two distributions,  $H(\pi)$  and  $G(\eta)$ , where  $(\mu, \sigma) = (0, 1.5)$  for  $H(\pi)$  and  $(\mu, \sigma) = (1, 1.5)$  for  $G(\eta)$ . The lower and upper bounds for the distributions are:  $(\underline{\pi}, \bar{\pi}) = (1, \infty)$ , and  $(\underline{\eta}, \bar{\eta}) = (1.1, \infty)$ .

The numerical problem has a nested structure. Given a value of  $r$ , the underlying problem is to find a fixed point,  $\sigma^*(r)$ , to  $\sigma = q \times r \times p(\lambda L(\sigma) + R(\sigma))$ , where  $L(\sigma)$  is the mass of cartel cases generated by the leniency program and  $R(\sigma)$  is the mass of non-leniency cartel cases.

The procedure for finding  $\sigma^*(r)$  begins by specifying an initial value for  $\sigma$ . For each  $\eta$ , we need to solve for a fixed point determining the collusive value:  $Y^*(\sigma, \eta) = \psi(Y^*(\sigma, \eta), \sigma, \eta)$ . As there may be multiple fixed points, the Pareto criterion is used which selects the largest fixed point. Since  $\psi(Y, \sigma, \eta)$  is increasing and  $\psi(\mu, \sigma, \eta) < \mu$  then, by setting  $Y^0 = \mu$  and iterating on  $Y^{t+1} = \psi(Y^t, \sigma, \eta)$ , this process converges to the largest fixed point.

---

<sup>31</sup>The Mathematica code that generates the equilibrium cartel rates for the baseline case is available at: <http://academic.csuohio.edu/changm/main/research/papers/CLPcodeA.pdf>.



In computing the stationary distribution of cartels, we need to take the step of computationally searching for  $\hat{\eta}(\sigma)$  which is the smallest industry type for which collusion is not incentive compatible for any market condition.  $\hat{\eta}(\sigma)$  is defined by:  $Y^*(\sigma, \eta) > \alpha\mu$  for  $\underline{\eta} < \eta \leq \hat{\eta}(\sigma)$  and  $Y^*(\sigma, \eta) = \alpha\mu$  for  $\eta > \hat{\eta}(\sigma)$ . To perform this step, we set  $\underline{\eta} = 1.1$  and  $\bar{\eta} = 10$  and use a 1,000 element finite grid of values for  $\eta$ , denoted  $\Gamma(\underline{\eta}, \bar{\eta})$ .  $\hat{\eta}(\sigma)$  is located by applying the iterative bisection method on  $\Gamma(\underline{\eta}, \bar{\eta})$ . As part of the bisection method,  $\bar{\eta}$  needs to be set at a sufficiently high value so that  $Y^*(\sigma, \bar{\eta}) = \alpha\mu$ . Once having identified  $\hat{\eta}(\sigma)$  and using  $Y^*(\sigma, \eta)$ ,  $\phi^*(\sigma, \eta)$  is calculated for a finite grid over  $[\underline{\eta}, \hat{\eta}(\sigma)]$ . These values are then used in computing  $L(\sigma)$  and  $R(\sigma)$ . The integration uses the Newton-Cotes quadrature method with the trapezoid rule (see Miranda and Fackler, 2002).

Choosing an initial value for  $\sigma$  and using our derived expressions for  $L(\sigma)$  and  $R(\sigma)$ , we then compute:  $\hat{\sigma}(r) = q \times r \times p(\lambda L(\sigma) + R(\sigma))$ . After specifying a tolerance level  $\epsilon$ , if  $|\sigma - \hat{\sigma}(r)| > \epsilon$  then a new value for  $\sigma$  is selected using the iterative bisection method. Note that once a new value for  $\sigma$  is specified, the entire preceding procedure must be repeated. This procedure is repeated until the process converges to the fixed point value of  $\sigma^*(r)$  such that  $|\sigma^*(r) - \hat{\sigma}(r)| \leq \epsilon$ .  $\epsilon$  is set at .0002.

Given the equilibrium probability of paying penalties,  $\sigma^*(r)$ , we can calculate the equilibrium cartel rate,  $C(\sigma^*(r))$ , mass of leniency cases,  $L(\sigma^*(r))$ , and mass of non-lenieny cases,  $R(\sigma^*(r))$ .

Note that the optimal competition policy,  $r^*$ , is the value of  $r$  that minimizes the equilibrium rate of cartels,  $C(\sigma^*(r))$ . To numerically derive  $r^*$ , we allow  $r \in \{0, .1, \dots, 1\}$  and perform the procedures described above for each of these values and identify the one that generates the minimum cartel rate.

In addition to the baseline parameter values, we considered a wide variety of parameter values off of the baseline in order to check for the robustness of the main properties identified in the paper. Specifically, for both the linear and non-linear  $p(\lambda L + R)$ , we considered  $\gamma \in \{0.7, 0.8, 0.9\}$ . Further robustness checks were performed for the non-linear  $p(\lambda L + R)$  for the following parameter values off of the baseline:  $\rho \in \{1.2, 1.4, 1.6\}$ ,  $\gamma \in \{0.3, 0.7, 2.0\}$ ,  $\lambda \in \{0.6, 0.8\}$ ,  $v \in \{100, 500\}$ ,  $n = 2$  (and, hence,  $\omega = 0.5$  or 1),<sup>32</sup>  $\alpha \in \{0.2, 0.5\}$ ,  $\kappa = 0.1$ , and  $\delta \in \{0.75, 0.95\}$ . For all these parameter values, the numerical results are consistent with the properties stated in the paper: i) a leniency program can lower or raise the cartel rate; ii) the change in average cartel duration from a leniency program is decreasing in the industry type,  $\eta$ ; and iii)  $\hat{\eta}$  is (generally) lower when there is a leniency program:  $\hat{\eta}_L < \hat{\eta}_{NL}$ .

---

<sup>32</sup>Note that  $\omega = \frac{n-1+\theta}{n}$ , where  $\theta = 0$  with a leniency program and  $\theta = 1$  without a leniency program. Hence, for  $n = 2$ ,  $\omega = 0.5$  with a leniency program and  $\omega = 1$  without a leniency program.

## References

- [1] Apestequía, Jose, Martin Dufwenberg, and Reinhard Selten, “Blowing the Whistle,” *Economic Theory*, 31 (2007), 143-166.
- [2] Aubert, Cecile, Patrick Rey, and William Kovacic, 2003, “The Impact of Leniency and Whistle-Blowing Programs on Cartels,” *International Journal of Industrial Organization*, 24 (2006), 1241-1266.
- [3] Bigoni, Maria, Sven-Olof Fridolfsson, Chloé Le Coq, and Giancarlo Spagnolo, “Fines, Leniency, and Rewards in Antitrust,” *RAND Journal of Economics*, 43 (2012), 368-390.
- [4] Borrell, Joan-Ramon, Juan Luis Jiménez, and Carmen García, “Evaluating Antitrust Leniency Programs,” Universitat de Barcelona, January 2012 (*Journal of Competition Law and Economics*, forthcoming).
- [5] Bos, Iwan and Frederick Wandschneider, “Cartel Ringleaders and the Corporate Leniency Program,” University of East Anglia (UEA) - Centre for Competition Policy, CCP Working Paper 11-13, August 2011.
- [6] Brenner, Steffen, “An Empirical Study of the European Corporate Leniency Program,” *International Journal of Industrial Organization*, 27 (2009), 639-645.
- [7] Chen, Joe and Joseph E. Harrington, Jr., “The Impact of the Corporate Leniency Program on Cartel Formation and the Cartel Price Path,” in *The Political Economy of Antitrust*, V. Ghosal and J. Stennek, eds., Elsevier, 2007.
- [8] Chen, Zhijun and Patrick Rey, “On the Design of Leniency Programs,” University of Toulouse, 2007.
- [9] Choi, Yun Jeong, “An Empirical Analysis of the Corporate Leniency Program in Korea: Its Amendments in 2005,” *Korean Journal of Economics*, 18 (2011), 255-281.
- [10] Choi, Jay Pil and Heiko Gerlach, “Global Cartels, Leniency Programs and International Antitrust Cooperation,” *International Journal of Industrial Organization*, 30 (2012), 528-540.
- [11] Dijkstra, Peter T., Marco A. Haan, and Lambert Schoonbeek, “Leniency Programs and the Design of Antitrust: Experimental Evidence with Unrestricted Communication,” University of Groningen, 2011.
- [12] Ellis, Christopher J. and Wesley W. Wilson, “Cartels, Price-Fixing, and Corporate Leniency Policy: What Doesn’t Kill Us Makes Us Stronger,” University of Oregon, 2001.
- [13] Friederiszick, Hans W. and Frank P. Maier-Rigaud, “Triggering Inspections Ex Officio: Moving Beyond a Passive EU Cartel Policy,” *Journal of Competition Law and Economics*, 4 (2008), 89-113.

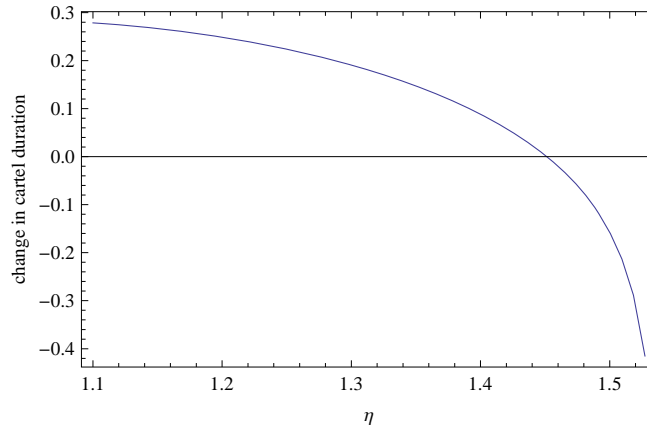
- [14] Gärtner, Dennis, “Corporate Leniency in a Dynamic World: The Preemptive Push of an Uncertain Future,” University of Bonn, October 2013.
- [15] Ghosal, Vivek and Joseph Gallo, “The Cyclical Behavior of the Department of Justice’s Antitrust Enforcement Activity,” *International Journal of Industrial Organization*, 19 (2001), 27-54.
- [16] Hamaguchia, Yasuyo, Toshiji Kawagoeb, and Aiko Shibatac, “Group Size Effects on Cartel Formation and the Enforcement Power of Leniency Programs,” *International Journal of Industrial Organization*, 27 (2009), 145–165.
- [17] Harrington, Joseph E., Jr., “Behavioral Screening and the Detection of Cartels,” in *European Competition Law Annual 2006: Enforcement of Prohibition of Cartels*, Claus-Dieter Ehlermann and Isabela Atanasiu, eds., Oxford: Hart Publishing, 2007.
- [18] Harrington, Joseph E., Jr., “Optimal Corporate Leniency Programs,” *Journal of Industrial Economics*, 56 (2008), 215-246.
- [19] Harrington, Joseph E., Jr., “When is an Antitrust Authority not Aggressive Enough in Fighting Cartels?,” *International Journal of Economic Theory*, 7 (2011), 39-50.
- [20] Harrington, Joseph E., Jr., “Corporate Leniency with Private Information: An Exploratory Example,” in *Recent Advances in the Analysis of Competition Policy and Regulation*, J. Harrington and Y. Katsoulakos, eds., Cheltenham, UK: Edward Elgar, 2012.
- [21] Harrington, Joseph E., Jr., “Corporate Leniency Programs when Firms have Private Information: The Push of Prosecution and the Pull of Pre-emption,” *Journal of Industrial Economics*, 61 (2013), 1-27.
- [22] Harrington, Joseph E., Jr. and Myong-Hun Chang, “Modelling the Birth and Death of Cartels with an Application to Evaluating Antitrust Policy,” *Journal of the European Economic Association*, 7 (2009), 1400-1435.
- [23] Hinloopen, Jeroen and Adriaan R. Soetevent, “Laboratory Evidence on the Effectiveness of Corporate Leniency Programs,” *RAND Journal of Economics*, 39 (2008), 607-616.
- [24] Houba, Harold, Evgenia Motchenkova, and Quan Wen, “The Effects of Leniency on Maximal Cartel Pricing,” TILEC Discussion Paper No. 2009-038, October 2009.
- [25] Jeffrey, R. L., "The Continuity of a Function Defined by a Definite Integral," *The American Mathematical Monthly*, 32 (1925), 297-299.
- [26] Klein, Gordon, “Cartel Destabilization and Leniency Programs - Empirical Evidence,” ZEW Discussion Paper No. 10-107, 2010.

- [27] Kwoka, John, “Commitment to Competition: An Assessment of Antitrust Agency Budgets since 1970,” *Review of Industrial Organization*, 14 (1999), 295-302.
- [28] Lefouili, Yassine and Catherine Roux, “Leniency Programs for Multimarket Firms: The Effect of Amnesty Plus on Cartel Formation,” *International Journal of Industrial Organization*, 30 (2012), 624-640.
- [29] Marshall, Robert C., Leslie M. Marx, and Claudio Mezzetti, “Antitrust Leniency with Multi-Product Colluders,” Penn State University, May 2013.
- [30] Miller, Nathan, “Strategic Leniency and Cartel Enforcement,” *American Economic Review*, 99 (2009), 750-768.
- [31] Miranda, Mario J. and Paul L. Fackler, *Applied Computational Economics and Finance*, Cambridge, Mass.: The MIT Press, 2002.
- [32] Motchenkova, Evguenia, “Effects of Leniency Programs on Cartel Stability,” Center for Economic Research Discussion Paper No. 2004-98, Tilburg University, 2004.
- [33] Motta, Massimo and Michele Polo, “Leniency Programs and Cartel Prosecution,” *International Journal of Industrial Organization*, 21 (2003), 347-379.
- [34] Riley, Alan, “Developing Criminal Cartel Law: Dealing with the Growing Pains,” *Competition Law Review*, 4 (2007), 1-6.
- [35] Rotemberg, Julio J. and Garth Saloner, “A Supergame-Theoretic Model of Price Wars During Booms,” *American Economic Review*, 76 (1986), 390-407.
- [36] Sauvagnat, Julien, “Prosecution and Leniency Programs: The Role of Bluffing in Opening Investigations,” Toulouse School of Economics, 2012.
- [37] Silbye, Frederik, *Topics in Competition Policy: Cartels, Leniency, and Price Discrimination*, PhD Thesis, University of Copenhagen, August 2010.
- [38] Spagnolo, Giancarlo, “*Divide et Impera*: Optimal Deterrence Mechanisms Against Cartels and Organized Crime,” University of Mannheim, 2003.
- [39] Spagnolo, Giancarlo, “Leniency and Whistleblowers in Antitrust,” in *Handbook of Antitrust Economics*, Paolo Buccirossi, ed., Cambridge, Mass.: The MIT Press, 2008.
- [40] Stephan, Andreas, “An Empirical Assessment of the European Leniency Notice,” *Journal of Competition Law and Economics*, 5 (2008), 537-561.
- [41] Zhou, Jun, “Evaluating Leniency and Modeling Cartel Durations: Time-Varying Policy Impacts and Sample Selection,” Bonn University, September 2011.

Figure 1: Change in average cartel duration conditional on  $\eta$  for linear  $p(\lambda L+R)$  with  $\lambda = 1.0$  and  $\gamma \in \{.7, .8, .9\}$

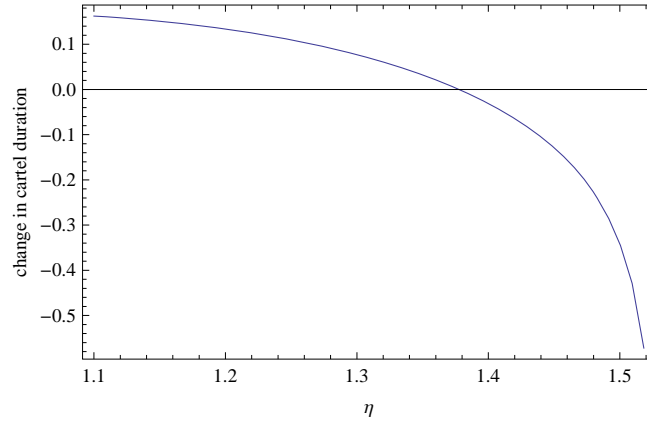
(a)  $\gamma = .7$  and  $\lambda = 1.0$

$$\hat{\eta}_L = 1.5272, \hat{\eta}_{NL} = 1.5539$$



(b)  $\gamma = .8$  and  $\lambda = 1.0$

$$\hat{\eta}_L = 1.5183, \hat{\eta}_{NL} = 1.5539$$



(c)  $\gamma = .9$  and  $\lambda = 1.0$

$$\hat{\eta}_L = 1.5094, \hat{\eta}_{NL} = 1.5450$$

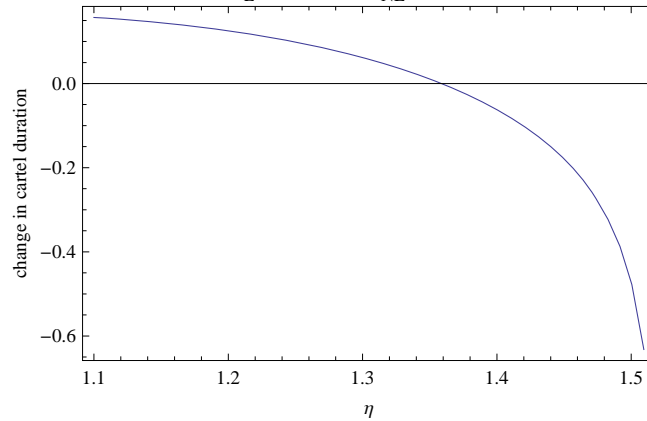
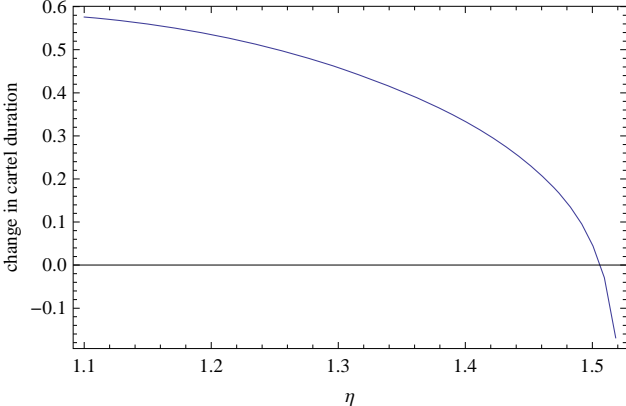


Figure 2: Change in average cartel duration conditional on  $\eta$  for non-linear  $p(\lambda L+R)$  with  $\lambda = 1.0$  and  $\gamma \in \{.7, .8, .9\}$

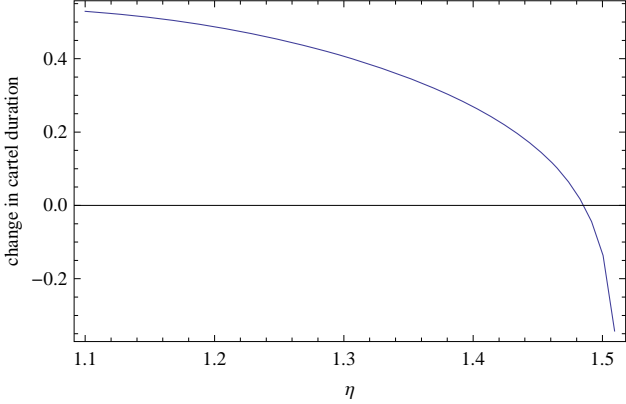
(a)  $\gamma = .7$  and  $\lambda = 1.0$

$$\hat{\eta}_L = 1.5183, \hat{\eta}_{NL} = 1.5361$$



(b)  $\gamma = .8$  and  $\lambda = 1.0$

$$\hat{\eta}_L = 1.5094, \hat{\eta}_{NL} = 1.5272$$



(c)  $\gamma = .9$  and  $\lambda = 1.0$

$$\hat{\eta}_L = 1.4916, \hat{\eta}_{NL} = 1.5272$$

