Peter T. Dijkstra, Marco A. Haan, Lambert Schoonbeek
Leniency Programs and the Design of Antitrust:
Experimental Evidence with Free-Form Communication
Review of Industrial Organization, 2020, vol. xxx, pg. xxx

Online Appendix A: Theoretical background

Here we discuss a theoretical infinitely-repeated game to provide further insight in the incentives that subjects face in our experiment. In our model two firms sell a homogeneous good on a market at zero production costs in each period. Demand is inelastic and normalized to unity. The firms compete in prices. Prices are chosen from $\{1, 2, \ldots, 10\}$. With equal prices the market is shared equally. Hence, in the stage game the fully collusive (monopoly) profit of each firm is $\pi^c = 5$ and the Nash equilibrium profit is $\pi^n = 0.5$. The discount factor of each firm is δ. In the experiment we have $\delta = 0.8$, but in the theoretical analysis we allow for $\delta \in [0, 1)$. We consider a number of variants of the model which are related to, respectively, Benchmark, Antitrust, Profound and Superficial. For each variant we investigate under which conditions a cartel can be sustained in a (pure-strategy) subgame-perfect equilibrium. We then say that the cartel is stable.

In Benchmark there is no Antitrust Authority (AA), and a cartel is never detected. In Antitrust, the AA discovers a cartel with probability $q = 0.15$. Each cartel member then has to pay a fine $f = 9$. In Profound and Superficial, a leniency program is in place. The AA starts an investigation with probability α. In turn, without reports, the AA discovers a cartel with probability p; with reports it does so for sure. In Profound, $\alpha = 0.2$ and $p = 0.75$, while $\alpha = 0.75$ and $p = 0.20$ in Superficial. A cartel member that is detected and has not reported incurs the full fine f. The first cartel member to report receives a reduced fine $r_1 = 1$ if an investigation has started, and one of $r_2 = 0$ if not. If both report, fines are shared equally. For each firm, the costs of reporting are $c = 0.5$. In each period, each firm first sets its price; next, it observes the opponent's price and learns if the AA has opened an investigation; finally, it decides whether or not

to report the cartel.

This appendix consists of two parts. Section A.1 examines the case where firms employ standard grim trigger strategies, where firms re-establish a new cartel if the previous one is discovered but no firm has defected (Friedman, 1971)**?**. Section A.2 investigates a variant where firms in the first cartel agree to set the monopoly price in the future without creating a new cartel once the first cartel is discovered but no firm has defected.

A.1. Standard grim trigger strategies

Assume that firms use standard grim trigger strategies in a cartel. Hence, according to the cartel agreement, each firm sets the monopoly price as long as no one has defected. Moreover, in cases with an AA, they agree to directly establish a new cartel once the previous one is discovered by the AA but no firm has defected. After a defection firms revert to the punishment phase and set the competitive price forever (i.e. the price of the Nash equilibrium of the stage game). In addition, if a leniency program is in place, then each firm immediately reports the cartel after a defection on price.

In Benchmark, each firm's payoff of being in a cartel is given by

$$V_b = \frac{\pi^c}{1-\delta}, \tag{A.1}$$

while the payoff associated with a defection is

$$V_b^d = (2\pi^c - 1) + \frac{\delta \pi^n}{1-\delta}. \tag{A.2}$$

The indices 'b' and 'd' denote, respectively, Benchmark and defection. Defecting on price by undercutting the rival yields profit $2\pi^c - 1$, as prices are integers. After a defection, the firms will earn π^n in each subsequent period. The cartel is stable if and only if

$$V_b \geq V_b^d. \tag{A.3}$$

In a similar way and using obvious notation, each firm's cartel payoff in Antitrust

equals

$$V_a = \frac{\pi^c - qf}{1-\delta}, \tag{A.4}$$

while the defection payoff is

$$V_a^d = (2\pi^c - 1) - \frac{qf}{1-\delta(1-q)} + \frac{\delta\pi^n}{1-\delta}. \tag{A.5}$$

The second term on the RHS of (A.5) reflects that the cartel may still be discovered and punished in the future. The cartel is stable in ANTITRUST if and only if

$$V_a \geq V_a^d. \tag{A.6}$$

In PROFOUND and SUPERFICIAL we not only have to take the price, but also the reporting decision in the cartel into account (Motta and Polo, 2003)**?**. Defections from a cartel agreement can be twofold: a firm can defect on the market by undercutting the price of its rival, and it can defect from the reporting agreement. Cartels can choose to collude and reveal (CR; set the monopoly price and report if an investigation has started) or to collude and not reveal (CNR; set the monopoly price and do not report if an investigation has started).

With the strategy CR, each firm's payoff is given by

$$V_{cr} = \pi^c - \alpha\left[c + \frac{1}{2}(r_1 + f)\right] + \delta V_{cr}. \tag{A.7}$$

This can be seen as follows. In each period, each firm earns collusive profits π^c. With probability α an investigation starts. If so, each firm reports, incurs reporting costs c, and either pays the reduced fine r_1 or the full fine f, both with equal probability. From the next period onwards, the payoff of the game is again V_{cr}. Solving for V_{cr} yields

$$V_{cr} = \frac{\pi^c - \alpha\left[c + \frac{1}{2}(r_1 + f)\right]}{1-\delta}. \tag{A.8}$$

After a defection on price, both firms immediately report the cartel. The corresponding payoff is

$$V_{cr}^d = (2\pi^c - 1) - c - \frac{1}{2}\left[\alpha r_1 + (1-\alpha)r_2 + f\right] + \frac{\delta\pi^n}{1-\delta}. \tag{A.9}$$

Clearly, it is not optimal to defect from CR by not revealing if an investigation has started. It is also not optimal to defect by revealing if no investigation has started. We conclude that the cartel is stable with CR if and only if

$$V_{cr} \geq V_{cr}^{d}. \tag{A.10}$$

Along the same lines, with the strategy CNR, the payoff of being in the cartel is $V_{cnr} = \pi^c - \alpha pf + \delta V_{cnr}$, hence

$$V_{cnr} = \frac{\pi^c - \alpha pf}{1-\delta}. \tag{A.11}$$

Note that $V_{cnr} = V_a$, i.e. being in a cartel with a leniency program in which participants never report is equivalent to being in a cartel without a leniency program.

A defection on price in CNR yields V_{cnr}^{d}, which equals V_{cr}^{d} from (A.9). A defection from not reporting can only make sense after an investigation has opened. Sticking to the agreement yields $V_{cnr|i} = \pi^c - pf + \delta V_{cnr}$, or

$$V_{cnr|i} = \frac{\pi^c - [1-\delta(1-\alpha)]\, pf}{1-\delta}, \tag{A.12}$$

while reporting yields

$$V_{cnr|i}^{d} = \pi^c - c - r_1 + \frac{\delta \pi^n}{1-\delta}. \tag{A.13}$$

Here '$cnr|i$' indicates CNR given that an investigation has started.

Summing up, the cartel is stable with CNR if and only if both

$$V_{cnr} \geq V_{cnr}^{d} \tag{A.14}$$

and

$$V_{cnr|i} \geq V_{cnr|i}^{d}. \tag{A.15}$$

Taking $\delta = 0.8$, we obtain the payoffs of Table A.1.[35] For each case, we have also determined the minimum value of $\delta \in [0, 1)$ that sustains a stable cartel, while keeping all other parameters at their values of the experiment. These thresholds are denoted by

[35] Strictly speaking, the values reported in Table A.1 thus only hold from period 20 onwards.

	Bench	Anti	Prof	Sup
V_b	25			
V_b^d	11			
$\bar{\delta}_b$	0.47			
V_a		18.25		
V_a^d		6.78		
$\bar{\delta}_a$		0.50		
V_{cr}			19.50	4.38
V_{cr}^d			5.90	5.63
$\bar{\delta}_{cr}$			0	0.88
V_{cnr}			18.25	18.25
V_{cnr}^d			5.90	5.63
$\bar{\delta}_{cnr}$			0.07	0.00
$V_{cnr\|i}$			12.85	17.80
$V_{cnr\|i}^d$			5.50	5.50
$\bar{\delta}_{cnr\|i}$			0.63	0.09

Table A.1: Payoffs and thresholds of the discount factor in the different cases if the players use standard grim trigger strategies. Bench, Anti, Prof and Sup denote, respectively, Benchmark, Antitrust, Profound and Superficial.

$\bar{\delta}$, with a subindex indicating the relevant case. For example, in Antitrust we have $\bar{\delta}_a = 0.50$, i.e. condition (A.6) holds for all $\delta \in [0.50, 1)$. A cartel is stable with CNR in Profound for all $\delta \in [\max\{\bar{\delta}_{cnr}, \bar{\delta}_{cnr|i}\}, 1) = [0.63, 1)$. If the threshold in a given case is smaller (larger) than in another case, we say that a stable cartel is more (less) likely in the former one.

First, let us consider the cartel and defection payoffs in Benchmark and Antitrust and discuss the main differences in the incentives faced by the firms in these two cases. Table A.1 shows that $V_b = 25$, $V_b^d = 11$, $V_a = 18.25$ and $V_a^d = 6.78$ if $\delta = 0.8$. Hence, cartels are stable in both Benchmark and Antitrust in that case. We see by inspection of the components of the RHS of (A.1) and (A.4), and (A.2) and (A.5), that the reason of the lower cartel and defection payoffs in Antitrust is that cartel members have to pay a fine if they are discovered by the AA in Antitrust, whereas by definition cartels cannot be discovered in Benchmark. In particular, in Antitrust, cartel members face an expected fine in each period if they adhere to the cartel agreement

since cartels are re-established after discovery by the AA, whereas in case of a defection they only face the (smaller) discounted value of a single expected fine. Considering $\delta \in [0, 1)$, we find that $\bar{\delta}_a = 0.50$ and $\bar{\delta}_b = 0.47$. Hence, a stable cartel is less likely in ANTITRUST than in BENCHMARK.

Second, examine CR in PROFOUND and SUPERFICIAL, and compare the cartel and defection payoffs in these cases with those in ANTITRUST. To begin with, take PROFOUND and $\delta = 0.8$. We then have $V_{cr} = 19.50$ and $V_{cr}^d = 5.90$, i.e. a cartel with CR is stable. Comparing with ANTITRUST, we see that $V_{cr} > V_a$. In order to understand this, take the RHS of (A.8) and (A.4), and note that $\alpha[c + \frac{1}{2}(r_1 + f)] < qf$, i.e. the per-period expected fine in a cartel with CR in PROFOUND is smaller than the per-period expected fine of a cartel in ANTITRUST. Recall here that the probability α that an investigation is started is small in PROFOUND. We further see that $V_{cr}^d < V_a^d$. This is driven by the fact that the expected fine (including c) after a defection (on price) from CR in PROFOUND is larger than the discounted expected fine after a defection in ANTITRUST, cf. (A.9) and (A.5). To understand this, note that after a defection from CR, the fine is paid in the same period, whereas that is not necessarily the case in ANTITRUST. Considering $\delta \in [0, 1)$, we find that a stable cartel is more likely with CR in PROFOUND than in ANTITRUST, since $\bar{\delta}_{cr} = 0$ and $\bar{\delta}_a = 0.50$.

Next, take SUPERFICIAL and $\delta = 0.8$. We then have $V_{cr} = 4.38$ and $V_{cr}^d = 5.63$. Hence, there is no stable cartel with CR. The size of V_{cr}^d is quite similar in SUPERFICIAL and PROFOUND. However, V_{cr} is much smaller in SUPERFICIAL than in PROFOUND. The reason is that the per-period expected fine in a cartel with CR is much larger in SUPERFICIAL than in PROFOUND, since the probability α that an investigation is started is much larger in SUPERFICIAL than in PROFOUND. Turning to the comparison with ANTITRUST, the large value of α also implies that $\alpha[c + \frac{1}{2}(r_1 + f)] > qf$ in SUPERFICIAL, which explains why $V_{cr} < V_a$ in that case, cf. (A.8) and (A.4). Further, $V_{cr}^d < V_a^d$ for SUPERFICIAL, which can be interpreted in a similar way as has been done for PROFOUND above. Note that for SUPERFICIAL the difference between V_{cr} and V_a is much larger than the difference between V_{cr}^d and V_a^d, which explains why there is no stable cartel in this

case. Taking $\delta \in [0,1)$, we see that a stable cartel is less likely with CR in SUPERFICIAL than in ANTITRUST, since $\bar{\delta}_{cr} = 0.88$ and $\bar{\delta}_a = 0.50$.

Finally, we investigate CNR in PROFOUND and SUPERFICIAL, and compare these cases with ANTITRUST. First, take PROFOUND and $\delta = 0.8$. We then have $V_{cnr} = 18.25$, $V^d_{cnr} = 5.90$, $V_{cnr|i} = 12.85$ and $V^d_{cnr|i} = 5.50$. Hence, we have a stable cartel with CNR. We see that $V_{cnr|i}$ is much smaller than V_{cnr}. The reason is that if firms adhere to CNR and an investigation has started, then the probability p that the cartel is discovered and fined is large in PROFOUND. The magnitudes of V^d_{cnr} and $V^d_{cnr|i}$ are almost the same. We further observe that $V_{cnr} < V_{cr}$ in PROFOUND with $\delta = 0.8$. Hence, from a theoretical point of view, we expect firms to play CR rather than CNR in this case. Yet, it is still interesting to compare CNR in PROFOUND with ANTITRUST. Doing this, we leave out $V_{cnr|i}$ and $V^d_{cnr|i}$ from our analysis, since (a defection from) an agreement about reporting is not relevant in ANTITRUST. We recall that $V_{cnr} = V_a$, i.e. being in a cartel where firms never report under a leniency program is equivalent to being in a cartel in ANTITRUST. Next, $V^d_{cnr} < V^d_a$. Recalling that $V^d_{cnr} = V^d_{cr}$, this inequality can be interpreted in the same way as $V^d_{cr} < V^d_a$ in our discussion of CR in PROFOUND, see above. Considering $\delta \in [0,1)$, we see that a defection on price is less likely for CNR in PROFOUND than in ANTITRUST, since $\bar{\delta}_{cnr} = 0.07$ and $\bar{\delta}_a = 0.50$. However, with CNR we also have to take into account the incentive to defect by reporting if an investigation has started (recall that p is large now). As a result, a stable cartel is less likely with CNR in PROFOUND since $\bar{\delta}_{cnr|i} = 0.63$, which is larger than $\bar{\delta}_a$.

Next, take SUPERFICIAL and $\delta = 0.8$. We then have $V_{cnr} = 18.25$, $V^d_{cnr} = 5.63$, $V_{cnr|i} = 17.80$ and $V^d_{cnr|i} = 5.50$, i.e. a cartel with CNR is stable. Comparing CNR in SUPERFICIAL and PROFOUND, we see that the main difference is that in SUPERFICIAL the size of $V_{cnr|i}$ is (much) larger than in PROFOUND (and close to V_{cnr}). The reason is that if firms adhere to CNR and an investigation is started, then the probability p that the cartel is discovered and fined is (much) smaller in SUPERFICIAL than in PROFOUND. Next, comparing CNR in SUPERFICIAL with ANTITRUST, we note again that $V_{cnr} = V_a$, while $V^d_{cnr} < V^d_a$ can be interpreted in the same way as $V^d_{cr} < V^d_a$ in our discussion of

CR in SUPERFICIAL above. Considering $\delta \in [0, 1)$, we see that a defection on price from CNR is less likely in SUPERFICIAL than in ANTITRUST, since $\bar{\delta}_{cnr} = 0$ and $\bar{\delta}_a = 0.50$. A defection with a report when an investigation has started is also not very likely, since $\bar{\delta}_{cnr|i} = 0.09$ (p is small now). In sum, a stable cartel is more likely with CNR in SUPERFICIAL than in ANTITRUST. With $\delta = 0.8$, as we have in the main text, there is a stable cartel in any treatment we consider.

A.2. Grim trigger strategies with planned tacit collusion

In this section we investigate the case where firms use grim trigger strategies like in Section A.1 but with one difference: firms now agree in the first cartel to set the monopoly price in the future without establishing a new cartel once their cartel will be discovered by the AA but no firm has defected. In this way, cartel members can be fined only once. We call the phase that starts after the discovery of the (first and only) cartel, planned tacit collusion. We now denote the payoff of a firm by W rather than V; indices indicate again the relevant case similar to Section A.1.

In BENCHMARK, a firm's payoff of being in a cartel is

$$W_b = \frac{\pi^c}{1-\delta}, \tag{A.16}$$

while the payoff corresponding to a defection is

$$W_b^d = (2\pi^c - 1) + \frac{\delta \pi^n}{1-\delta}. \tag{A.17}$$

The cartel is stable if and only if

$$W_b \geq W_b^d. \tag{A.18}$$

Obviously, incentives in BENCHMARK are identical in Sections A.1 and A.2.

In ANTITRUST, the payoff of adhering to the cartel agreement is

$$W_a = \frac{\pi^c}{1-\delta} - \frac{qf}{1-\delta(1-q)}, \tag{A.19}$$

where the second term of the RHS denotes the discounted value of the single expected fine of the cartel. A defection yields the payoff

$$W_a^d = (2\pi^c - 1) - \frac{qf}{1-\delta(1-q)} + \frac{\delta \pi^n}{1-\delta}. \tag{A.20}$$

The cartel is stable if and only if

$$W_a \geq W_a^d. \tag{A.21}$$

Note that (A.21) can be rewritten as (A.18), since in ANTITRUST the discounted expected fines in (A.19) and (A.20) are the same.

Next, suppose that a leniency program is in place (cf. PROFOUND and SUPERFICIAL). We then focus again on strategies CR and CNR, but now the cartel agreement involves a phase with planned tacit collusion. Recall that in CR, both firms report the cartel if an investigation has started. Using that strategy, a firm's payoff equals

$$W_{cr} = \frac{\pi^c}{1-\delta} - \frac{\alpha[c + \frac{1}{2}(r_1 + f)]}{1 - \delta(1-\alpha)}. \tag{A.22}$$

The payoff associated with a defection on price is

$$W_{cr}^d = (2\pi^c - 1) - c - \frac{1}{2}[\alpha r_1 + (1-\alpha)r_2 + f] + \frac{\delta \pi^n}{1-\delta}. \tag{A.23}$$

The payoff of CR in a period of the phase with planned tacit collusion is $W_{cr|t}$, while the payoff of a defection in such a period is $W_{cr|t}^d$. Here '$cr|t$' denotes CR given that the firms have reached the phase with planned tacit collusion. Note that $W_{cr|t} = W_b$ and $W_{cr|t}^d = W_b^d$, see (A.16) and (A.17). It can be verified that there is no incentive to defect from CR by not revealing if an investigation has started. There is also no incentive to defect by revealing if no investigation has started.

We conclude that a cartel with CR is stable if and only if both

$$W_{cr} \geq W_{cr}^d \tag{A.24}$$

and

$$W_{cr|t} \geq W_{cr|t}^d. \tag{A.25}$$

Note that (A.25) is equivalent to (A.18) and (A.21).

Finally, examine strategy CNR involving planned tacit collusion. Recall that in this case a firm does not report its cartel if an investigation has started. A firm's payoff of adhering to CNR is

$$W_{cnr} = \frac{\pi^c}{1-\delta} - \frac{\alpha p f}{1 - \delta(1 - \alpha p)}. \tag{A.26}$$

The payoff associated with a defection on price is W^d_{cnr}, which equals W^d_{cr} from (A.23). A defection from not reporting can only be optimal if an investigation has started. The payoff of adhering to CNR is then given by

$$\begin{aligned} W_{cnr|i} &= \pi^c + p\left[\frac{\delta\pi^c}{1-\delta} - f\right] + (1-p)\delta W_{cnr} \\ &= \frac{\pi^c}{1-\delta} - pf - (1-p)\left(\frac{\delta\alpha p f}{1-\delta(1-\alpha p)}\right), \end{aligned}$$

while the payoff of a defection is

$$W^d_{cnr|i} = \pi^c - c - r_1 + \frac{\delta\pi^n}{1-\delta}. \tag{A.27}$$

Combining, we see that a cartel with CNR is stable if and only if both

$$W_{cnr} \geq W^d_{cnr} \tag{A.28}$$

and

$$W_{cnr|i} \geq W^d_{cnr|i}. \tag{A.29}$$

Taking $\delta = 0.8$, we obtain Table A.2 as the counterpart of Table A.1. The table also gives the minimum values of $\delta \in [0, 1)$ sustaining a stable cartel, while keeping the other parameters at their values of the experiment. The thresholds are now denoted by $\hat{\delta}$, with an index indicating the relevant case.

Using the payoff values derived with $\delta = 0.8$, we conclude from Table A.2 that a stable cartel exists in all cases. Because $W_{cr} > W_{cnr}$ in Profound, and $W_{cr} < W_{cnr}$ in Superficial, theory predicts that firms play CR in Profound, and CNR in Superficial. These findings are similar to those derived from Table A.1. We can further discuss Table A.2 in the same way as Table A.1. In many cases the corresponding entries of Table A.2 and Table A.1 have about the same size. For brevity, we therefore focus on a three interesting differences between the tables.

First, taking Benchmark and Antitrust, we see that $\hat{\delta}_b = \hat{\delta}_a$. Hence, a stable cartel is equally likely in these two cases if firms employ grim trigger strategies involving planned tacit collusion. This follows directly from our earlier observation that (A.18)

and (A.21) are equivalent. Recall that a stable cartel is more likely in BENCHMARK than in ANTITRUST if firms use standard grim trigger strategies, i.e. $\bar{\delta}_b < \bar{\delta}_a$ in Table A.1. In that case, if firms adhere to a cartel agreement, they are fined more often in ANTITRUST than in BENCHMARK, since the cartel is then re-established (and thus can be fined again) if the previous one is discovered by the AA. With grim strategies involving planned tacit collusion, a cartel is created only once in ANTITRUST (and thus can only be fined once in that case).

Second, examine CR in SUPERFICIAL. The payoffs W_{cr} and W_{cr}^d, and threshold $\hat{\delta}_{cr}$, in Table A.2 correspond to the initial phase of the cartel (when it is not yet discovered by the AA). They correspond to, respectively, V_{cr}, V_{cr}^d and $\bar{\delta}_{cr}$ in Table A.1. Notice now that $W_{cr} = 19.84$, while $V_{cr} = 4.38$. The larger value of W_{cr} arises since only one cartel is established (and thus might be discovered and fined) if firms use CR involving planned tacit collusion, whereas cartel members might be fined multiple times if they use CR without planned tacit collusion. Note from Table A.2 that a cartel with CR involving planned tacit collusion is stable, whereas a cartel with CR without planned tacit collusion is not stable, see Table A.1.

Third, $W_{cr|t}$, $W_{cr|t}^d$ and $\hat{\delta}_{cr|t}$ in Table A.2 correspond to the phase with planned tacit collusion. Obviously, these numbers do not have counterparts in Table A.1, since this phase is not relevant in that case.

	Bench	Anti	Prof	Sup
W_b	25			
W_b^d	11			
$\hat{\delta}_b$	0.47			
W_a		20.78		
W_a^d		6.78		
$\hat{\delta}_a$		0.47		
W_{cr}			21.94	19.84
W_{cr}^d			5.90	5.63
$\hat{\delta}_{cr}$			0	0.41
$W_{cr\|t}$			25	25
$W_{cr\|t}^d$			11	11
$\hat{\delta}_{cr\|t}$			0.47	0.47
W_{cnr}			20.78	20.78
W_{cnr}^d			5.90	5.63
$\hat{\delta}_{cnr}$			0.07	0
$W_{cnr\|i}$			14.88	19.83
$W_{cnr\|i}^d$			5.50	5.50
$\hat{\delta}_{cnr\|i}$			0.61	0.09

Table A.2: Payoffs and thresholds of the discount factor in the different cases if the players use the grim trigger strategies involving tacit collusion. Bench, Anti, Prof and Sup denote, respectively, Benchmark, Antitrust, Profound and Superficial.