The effect of whistle–blowing incentives on collusion:
an experimental study of leniency programmes

Nick Feltovich∗
Department of Economics
Monash University
Clayton VIC 3800, Australia
nicholas.feltovich@monash.edu

Yasuyo Hamaguchi
Faculty of Economics
Nagoya City University
Nagoya 467–8501, Japan
yhamagu@econ.nagoya-cu.ac.jp

May 22, 2017

Abstract

Policy makers are increasingly using and encouraging whistle–blowing incentives aimed at curtailing various types of illegal or unethical behaviour. We theoretically and experimentally investigate one version of whistle–blowing incentive: leniency programmes aimed at curbing anti–competitive activities by firms, by reducing the punishment faced by a cartel member who reports the cartel’s behaviour. The theoretical model captures the two important effects of whistle–blowing incentives: the direct effect, a reduction in the stability of cartels, and the counterproductive indirect effect, an increase in the incentives to form cartels in the first place by lowering the cost of exiting them. As these point in opposite directions, the net theoretical effect is indeterminate. Our laboratory experiment compares two leniency programmes – full immunity from fines and partial immunity – against a baseline with no whistle–blowing incentives in place. We find evidence of the direct effect but not the indirect effect, and thus both programmes reduce the extent of price fixing and the damage associated with it.

Journal of Economic Literature classifications: L41, K42, D43, C73, D03.
Keywords: leniency programme; whistle–blowing; antitrust policy; oligopoly; collusion.

∗Corresponding author. Some of this research took place while Feltovich was at University of Aberdeen. Financial support from the UK Office of Fair Trading, the Japan Society for the Promotion of Science (Grants–in–Aid for Young Scientists (B), #21730231) and the Nomura Foundation for Social Science is gratefully acknowledged, though the views expressed are solely those of the authors. We thank Shuya Hayashi, Stephen King, Erika Seki and participants at several conferences and seminars for helpful suggestions and comments.
1 Background

Many forms of illegal or unethical behaviour – from cheating on taxes, to bribery of officials, to firms’ collusion – are difficult for outsiders to detect. The use of “whistle-blowers” – people with knowledge of misbehaviour, either as one of the participants (e.g., a cartel member) or as an inside observer (e.g., a public official who notices a colleague soliciting bribes) – is invaluable in curbing these kinds of activities, and perhaps prosecuting those involved. Whistle-blowing often entails substantial personal cost, through self-incrimination or fear of reprisal. In order to overcome these costs, governments are increasingly instituting (or in some cases, are pressing or mandating other organisations to institute) whistle-blowing incentives designed to reduce the cost or increase the benefit of whistle-blowing to the whistle-blower. These incentives take many forms, such as workplace protections for employees blowing the whistle on their bosses, a share of tax receipts for citizens reporting tax cheats, or reduced fines and punishments for collusive firms reporting their activity.

Obviously, the aim of these whistle-blowing incentives is to increase whistle-blowing. We refer to this as the direct effect of a whistle-blowing incentive. It is possible, however, that the social benefits arising from this direct effect can be offset by a perverse indirect effect: as whistle-blowing incentives reduce the cost of exit from a corrupt arrangement, such arrangements may be more likely to form in the first place. This latter effect may be unlikely in some settings (e.g., paying people to report tax cheats may not increase the prevalence of cheating on taxes), but in other settings it is potentially important. Experimental studies of corruption have sometimes found that the possibility of whistle-blowing leads to more, not fewer, corrupt transactions (Abbink 2006; Lambsdorff and Frank 2010), while theoretical and empirical analyses of leniency programmes show that while they can make cartels less stable, they can also make them more likely to form in the first place (Motta and Polo 2003; Chen and Harrington 2007; Marvão 2016).

In this paper, we investigate these two effects. We concentrate attention on the setting of an oligopolistic industry, where the misbehaviour in question is price fixing by the firms, and the whistle-blowing incentive is a leniency programme. This setting is worthy of our focus for several reasons, most notably because of the extent of anti-competitive behaviour by real firms, and the corresponding scale of damage to consumers and to social welfare. Additionally, the increasing prevalence of leniency programmes (see Hamaguchi et al. 2009 for some statistics) makes it important to understand their effects; in particular, as noted above, the indirect effect of such programmes is more likely to be seen here than in other settings of this kind, meaning that the overall effects are especially difficult to determine based on intuition, or even theoretical analysis.

Attempts to determine the effects of leniency programmes empirically using field data can suffer from other

---

1 More precisely, they aim to increase whistle-blowing in cases where illegal or unethical behaviour has taken place. The Economist (2015) reports that the US Securities and Exchange Commission acknowledges that its programme of giving cash bounties to successful whistle-blowers has led to “a problem with ‘serial submitters’, who file dozens of spurious claims in the hope that one will lead to a payout”.

2 Estimates of aggregate harm to consumers from anti-competitive behaviour are difficult to find, but Laitenberger and Smuda (2015) estimate that a single cartel (detergent), lasting for three years in eight countries, caused 315 million Euros in consumer damage. Baker (2003) estimates the deadweight loss from anti-competitive behaviour in the US to be at least 1 percent of GDP, which is likely a conservative estimate of harm to consumers since pure transfers from consumer to producer resulting from higher prices are not directly counted (though may be indirectly included if these transfers lead to rent seeking by would-be monopolists).
difficulties. Foremost among these is that the researcher may have good information about anti–competitive behaviour that is detected by the competition authority, but less about undetected cartels. As a result of this selection bias, identifying the entire population of collusive episodes – and hence reaching general conclusions about how leniency programmes affect collusion – can be problematic. Additionally, many of the relevant variables governing behaviour (such as firms’ costs and the shape of the demand curve) are largely unobservable to the researcher. Even when these obstacles can be surmounted, natural experiments (pairs of policy environments that are identical except for the leniency programme) are very rare, so it is difficult to accomplish direct tests of policy variables.³

Because of these difficulties, we turn to the laboratory. Lab experiments offer several advantages for the study of whistle–blowing incentives. First, the experimenter has much greater control over the preferences of decision makers than field researchers do. Second, relevant policy variables can be varied in a systematic way, so that the effect of the policy itself can be isolated. Third, an important source of selection bias is eliminated in experimental data; all decisions of all relevant decision makers are collected and available for the researcher to view. While lab experiments also have their drawbacks (chiefly, questions of external validity arising from differences between the population of experimental subjects and that of real firms’ decision makers), these drawbacks are arguably outweighed by the advantages.

There have been a few previous experimental studies over the last decade or so examining the effects of leniency programmes. We discuss this literature in some detail in Section 2.4, but for now, we mention two features that are important to recognise. First, any implementation of oligopoly with the possibility of collusion in the lab requires substantial simplification from the real world, and there is a fair amount of variation across studies in what kinds of simplifications are made. Given this, it is of little use to critique any of the settings used in these studies as being “unrealistic”: all of them are unrealistic, as is ours. Second, the previous research shows a reasonable consensus on how leniency programmes affect the frequency of reporting – the “direct effect” discussed above. Nearly always when this effect is examined, its sign is the expected one: increasing the incentive to whistle–blow leads to more whistle–blowing (though not always significantly more). However, the effects of these programmes on the overall level of collusion and on prices can be either negative or positive, suggesting that the “indirect effect” varies in strength from negligible to powerful. The range of conclusions in this literature suggests a need for additional research.

The current paper contributes to this literature, with the use of a simple but rich theoretical model designed to replicate the important features of firms’ decisions under various competition policies, and a laboratory experiment implementing this model. Subjects in the experiment play the role of firms in a repeated duopoly with communication possible before the first period. Each period comprises a simultaneous choice of prices, and in the event that both firms choose high prices, they run the risk of being punished for anti–competitive behaviour by a competition authority.⁴ In two “leniency” treatments, firms can avoid the fine – either partly or completely – by blowing the whistle on the rival firm. We compare the results of these treatments to results from a baseline

³It should be noted that these difficulties are not insurmountable. See, for example, Miller (2009) and Brenner (2009).
⁴That is, communication between firms is neither necessary nor sufficient for a finding of collusion in our setting; what matters is whether firms actually chose high prices. See Section 2.4 for an extensive discussion of this assumption; for now, we note that our definition is unrealistic, but so is every other definition of collusion that has been used in a lab experiment.
where reporting and leniency are not possible. Formulating theoretical predictions based on our model is complicated by the large number of equilibria (as was true for previous papers in this literature). However, we show that the set of equilibria differs across treatments in systematic ways. Specifically, moving from the baseline treatment to partial leniency, or from partial to full leniency, makes it easier to support equilibria involving collusion and reporting. Thus, both the direct and the indirect effects of leniency policies are seen. However, since the two effects have opposite implications for prices, the extent of collusion, and firms’ profits, understanding the net effects on these variables still requires empirical testing via our experiment.

Broadly speaking, our experimental results should inspire optimism about the usefulness of leniency programmes, and whistle–blowing incentives more broadly. Our data indicate that the indirect effect of leniency programmes is negligible, as there is no increase in cartel formation under either partial or full leniency, compared to our baseline treatment. On the other hand, there is plenty of evidence for the direct effect, as collusion is less stable in either of our leniency treatments than in the baseline, and this decrease in stability is due to firms reporting to the competition authority. The overall impact of these leniency programmes, while smaller than the incidence of whistle–blowing on its own would suggest (due to a countervailing reduction in price undercutting), includes both a decrease in the fraction of time firms spend in collusion, and decreases in the harm done to consumers (measured by either prices or firms’ excess profits).

2 The duopoly game and experimental design

Our theoretical analysis begins with a symmetric 3x3 “underlying game”, shown in Figure 1 (see Hinloopen 2006 for a more general model of collusion, competition policy and reporting). This game is not actually used in our experiment, but the stage games we do use involve only minor modifications to it. Two firms simultaneously choose prices. For simplicity, we limit the set of prices to High, Medium and Low. The Medium price signifies non–collusive competition between the firms. The High price represents an attempt to collude; this yields high profits if the other firm also chooses High (and thus (High, High) is the collusive outcome), but low profits otherwise. Indeed, High and Medium on their own form a prisoners’ dilemma. The Low price is strictly dominated in the stage game, but as we will see below, its existence allows for a richer set of repeated–game equilibria than would be possible with only the High and Medium prices. In the stage game, then, (Medium, Medium) is the unique Nash equilibrium – and indeed the unique rationalisable outcome – with resulting payoffs

5We focus on a market with two firms, as the highly competitive nature of markets with even a small additional number of sellers is well recognised by economists. In the theoretical industrial organisation literature, for example, Selten (1973) shows that firms are substantially more likely to behave competitively in markets with five or more firms than in markets with fewer firms. In the empirical IO literature, Bresnahan and Reiss (1991) look at geographically separated markets in several industries, and find that three sellers in a market are typically enough to yield essentially competitive prices. Isaac and Reynolds (2002) find a similar result in experiments: four–firm posted–offer markets give rise to a strong push toward competitive price levels, while no such tendency is observed when there are only two firms. Based on these results, our view ex ante was that duopoly markets were desirable in order to guarantee a non–negligible level of collusion in our baseline treatment. As noted in our literature review, however, both duopolies and triopolies have been used successfully in previous experimental tests of competition policy.
of 4 for each firm.\footnote{Our underlying game is meant to give the flavour of Bertrand competition, though our payoffs were not derived from any particular parameterisation of the Bertrand model. We chose the payoffs from High and Medium choices so that observed levels of the collusive outcome (High, High) would be likely to be far from either zero or one in our baseline treatment, so that we could detect a treatment effect in either direction. To do this, we set a fairly low incentive to deviate from collusion (only 1 unit is gained by switching from High to Medium if the rival chooses High) and a fairly high payoff difference (6 units) between the collusive outcome and the Nash equilibrium. The intuition behind the payoffs following a Low price choice are as follows. Choosing Low ensures capturing nearly all of the market share, but at an unprofitable price. A rival who also chooses Low will be selling at a similar unprofitable price, so that both firms earn zero. A rival choosing a higher price will earn a non–zero markup on each unit, but will sell few units at a Medium price, and fewer still at a High price.}

This stage game is infinitely repeated, with a constant discount factor $\delta \in (0, 1)$ that is identical for all firms. Repeated–game payoffs are defined in the usual way. For Firm $i$, if $\pi_i(t)$ is the one–shot–game payoff earned in round $t$, we have

$$U_i(\pi_i(1), \pi_i(2), \pi_i(3), \ldots) = \sum_{t=1}^{\infty} \delta^{t-1} \pi_i(t).$$

(1)

Since (Medium, Medium) is the Nash equilibrium of the one–shot game, there always exists a subgame perfect equilibrium of the repeated game in which both firms choose Medium in every round, irrespective of $\delta$. However, infinite repetition potentially leads to a larger set of equilibria. In particular, when $\delta$ is sufficiently high, collusion – in the sense of (High, High) stage–game outcomes – can be achieved. This is accomplished by the use of “trigger strategies”: strategies that choose a cooperative action (High, in our setting) as long as the opponent also has behaved cooperatively, but punish the opponent with a less cooperative strategy if the opponent has deviated from cooperation. In the game shown in Figure 1, there exist subgame perfect equilibria in which (High, High) is the outcome in every round as long as $\delta \geq \frac{1}{4}$; this collusion is supported by the threat of switching to Medium after a deviation, which is credible since (Medium, Medium) is a stage–game equilibrium.

Even with lower discount factors, collusion is possible if we relax the assumption of subgame perfection and consider the full set of Nash equilibria. In particular, as long as $\delta \geq \frac{1}{4}$, (High, High) in each round can be enforced using trigger strategies in which deviations are punished with play of Low forever. While non–subgame–perfect Nash equilibria are often viewed by theorists as less appealing than subgame perfect equilibria, there is a fair amount of evidence from experiments involving repeated social dilemmas (e.g., Result 6 in Dal Bó and Fréchette’s 2016 survey article, which states “...subjects use punishments to support cooperation, but punishments are not necessarily credible (not SPE)” (p. 36); see also Bayer 2014, and see Hamaguchi et al.


<table>
<thead>
<tr>
<th>Firm 2</th>
<th>High</th>
<th>Medium</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>10, 10</td>
<td>2, 11</td>
<td>2, 0</td>
</tr>
<tr>
<td>Firm 1</td>
<td>11, 2</td>
<td>4, 4</td>
<td>2, 0</td>
</tr>
<tr>
<td>Low</td>
<td>0, 2</td>
<td>0, 2</td>
<td>0, 0</td>
</tr>
</tbody>
</table>
2003 for an example using one–shot extensive–form games) that the theoretical “credibility” of a threat has little bearing on whether it will be made, and whether it can actually support cooperation. If credibility in repeated social dilemmas is less important than we generally believe, then it is correct to use Nash equilibrium – rather than subgame perfect equilibrium – as the solution concept in our setting. Even for those not convinced by the evidence in the previous literature, our use of an experiment as well as theoretical analysis means that we are not assuming credibility does not matter when formulating theoretical predictions; we are collecting data to examine whether credibility matters.

2.1 The baseline game

Our baseline game (i.e., the game we use in our baseline treatment) arises from one small modification to the underlying game in Figure 1: we introduce competition law, in the shape of a non–strategic competition authority. The competition authority discovers anti–competitive behaviour, which we define as a choice of High by both firms, with exogenous probability \( p \in [0, 1] \). (See Section 2.4 for an extensive discussion of this definition.) When such behaviour is discovered, two penalties are imposed: (i) an exogenous fine \( F > 0 \) in the current round, and (ii) a prohibition of choices of High in all future rounds of the repeated game. This latter penalty is meant to reflect the likelihood that once the competition authority has detected anti–competitive behaviour, it will have gained information about market conditions that will help it identify excessive prices more easily in future, and additionally it will keep these firms under increased scrutiny. Our assumption that this makes future collusion impossible, rather than just more difficult, is obviously a simplification of the real world (firms caught colluding are able to re–form cartels and sometimes do so), and it means that while we can study cartel formation and stability within our setting, we cannot investigate recidivism. However, this assumption is not unknown in the literature (see, e.g., Hinloopen 2006). Also, the alternative assumption typically made in this experimental literature – that the probability of a second cartel being detected is no higher than it was for the first cartel containing the same firms – is arguably also unrealistic.

Penalty (ii) reduces the stage game to the game shown in Figure 2, which we refer to as the game with punishment in effect.\(^7\) This reduced game also has (Medium, Medium) as the unique Nash equilibrium. Importantly,

\[
\begin{array}{c|cc}
\text{Firm 1} & \text{Medium} & \text{Low} \\
\hline
\text{Medium} & 4, 4 & 2, 0 \\
\text{Low} & 0, 2 & 0, 0
\end{array}
\]

Figure 2: The stage game when punishment is in effect

there is no scope for cooperation between firms when punishment is in effect, as for any \( \delta \), there is no repeated–

\(^7\)Figure 2 makes clear another benefit of including Low as an available price choice: a non–degenerate decision for each player when punishment is in effect (by contrast, Hamaguchi et al. 2009 have only one available action choice once collusion has been caught by the competition authority).
game equilibrium in which any action other than Medium is observed.

Adding competition law has no effect on payoffs as long as the firms do not collude (i.e., as long as at least one firm chooses Medium or Low); thus there continues to be a subgame perfect equilibrium of the repeated game in which both firms always choose Medium. However, as before, it might be that collusion can also be supported. To examine this possibility, we start by noting that in the first round, it is not possible that punishment is in effect (since there was no previous round for collusion to be detected by the competition authority), so it is possible for firms to choose High. We next compute the repeated–game payoff to each firm in a path of play in which both firms choose High as long as punishment is not in effect; call this payoff \( \pi^* \). Choices of (High, High) in the first round yield a stage–game payoff of 10 to each firm. With probability \( p \), the firms are caught by the competition authority, fined \( F \) in the first round, and limited to payoffs of 4 in each subsequent round. With probability \( 1 - p \), they are not caught, incur no fine, and begin the next round in the same state as the current one (leading to a continuation payoff of \( \pi^* \), discounted once). \( \pi^* \) thus satisfies the equation

\[
\pi^* = 10 + p(-F + 4\delta + 4\delta^2 + \cdots) + (1-p)\delta\pi^*; \tag{2}
\]

solving for \( \pi^* \) yields

\[
\pi^* = \frac{4p\delta + (1-\delta)(10-pF)}{(1-\delta)(1-\delta+p\delta)}. \tag{3}
\]

It follows that (High, High) is consistent with subgame perfect equilibrium if it can be supported by a credible trigger strategy (threatening to choose Medium forever if the other firm deviates from High), which is true if \( \pi^* \) is higher than the payoff from a (Medium, High) outcome in the first round and (Medium, Medium) in all rounds thereafter:

\[
\frac{4p\delta + (1-\delta)(10-pF)}{(1-\delta)(1-\delta+p\delta)} \geq 11 + 4\delta + 4\delta^2 + \cdots = 11 + 4 + \frac{\delta}{1-\delta}. \tag{4}
\]

This equilibrium condition can be rewritten as \( \left[7(1-p)\delta - (1 + pF)\right](1-\delta) \geq 0 \), and since both \( (1-\delta) \) and \( (1-p) \) are always positive, the condition simplifies further to \( \delta \geq \frac{1+pF}{7(1-p)} \). As was true in the underlying game, (High, High) in this baseline game can be supported for some lower \( \delta \) in Nash (but not subgame perfect) equilibrium.\(^8\)

### 2.2 The game with leniency

From the baseline game, we make one more modification to obtain our game with leniency: following a (High, High) pair, the competition authority allows each firm the opportunity to report their anti–competitive behaviour (see Figure 3). If either firm does so, then the competition authority detects their anti–competitive behaviour with probability one, and the restriction on future prices to Medium and Low is put into place for both firms as before. The firm reporting the anti–competitive behaviour receives a reduced fine of \( F - R \) (with \( R > 0 \) the

\(^8\)The Nash condition for \( \delta \) is \(-2p\delta + (1-\delta)[9\delta(1-p) - (1 + pF)] \geq 0 \). This quadratic in \( \delta \) holds over an interval that strictly includes the interval between \( \frac{1+pF}{7(1-p)} \) (the lowest \( \delta \) for which collusion can hold in subgame perfect equilibrium) and 1. That is, collusion supported by a threat of Low prices is possible for any discount factor that allows collusion with a threat of Medium prices, and for some lower discount factors.
“reward” for reporting), while the firm not reporting receives the same fine ($F$) as in the baseline game. If both firms choose to report, both receive the reduced fine.\(^9\)

Figure 3: Sequence of decisions in the stage game with leniency

![Sequence of decisions](image)

The analysis of the associated repeated game (with the price decision and, if applicable, the reporting decision in each stage) depends on parameter values, but as long as $R$ is no more than the amount of the fine, the set of subgame perfect equilibrium outcomes is identical to what it was in the baseline game: there are equilibria in which both firms choose Medium in all rounds, as well as equilibria in which both choose High (and do not report) in all rounds until they are caught. Notably, for these values of $R$, there are no subgame perfect equilibria in which the firms report their anti–competitive behaviour.

However, reporting can happen in non–subgame–perfect Nash equilibria. Specifically, consider a strategy according to which the firm chooses High in the first round, and if the other firm also chooses High, reports and then chooses Medium forever, but if the other firm chose Medium or Low in the first round, punishes by choosing Low forever. For a given value of $R$, both firms’ playing this strategy constitutes a Nash equilibrium as long as $\delta$ is sufficiently high:

$$\delta \geq 1 - 2/(3 + F - R).$$

In this case, the observed result would be both firms colluding in the first round, both reporting the collusion, then both choosing Medium from then on. (As we will see in the next section, the parameters used in the experiment will satisfy this condition.) The reporting is supported in equilibrium because of the relatively gentle punishment (fine reduced by leniency, followed by Medium prices) for reporting compared to the more severe consequences from under–cutting (high profit in the current round, but Low prices thence). The condition (5) additionally implies that as $R$ increases, collusion followed by reporting becomes still easier to support, either in the sense that a lower discount factor is needed for such equilibria to exist, or because a larger set of punishments for deviators is available to support collusion and reporting. That is, not only is the incentive to report greater as $R$ increases (the “direct effect” of leniency programmes mentioned in the introduction), but so is the incentive to collude in the first place (the “indirect effect”).

\(^9\)Our giving both reporting firms an equal fine reduction – rather than giving a larger reduction to the first reporting firm – is not unusual within the experimental literature: Apesteguia et al. (2007), Hamaguchi et al. (2009) and Bigoni et al. (2012) also do so, though some studies do not (e.g., Hinloopen and Soetevent 2008 give full, half, and no fine reduction to the first, second, and third reporting firm). While it may seem implausible for real competition authorities to give equal benefits to all reporting firms – and indeed Bigoni et al. (2015) conclude that deterrence will be stronger if only the first reporting firm benefits – there is some evidence that when collusion occurs in multiple jurisdictions, colluding firms coordinate by each reporting first in a different jurisdiction (see Laitenberger and Šmuda 2013, and especially their footnote 14 which details such a case).
2.3 Experimental design and theoretical implications

In the experiment, the discount factor $\delta$ is set to 0.8, which is large enough to support many kinds of equilibrium but small enough to allow multiple supergames in an experimental session. Given that collusion (both firms choosing High) occurred, exogenous detection occurs with probability $p = 0.08$, comparable to the values used in other collusion experiments (which range from zero to 0.4, as noted in Section 2.4), and the fine for being caught is set at the stage-game gain from collusion over the competitive outcome ($F = 6$). The reward for reporting anti-competitive behaviour $R$ is a treatment variable, and takes on one of two values: $R = 3$ or $R = 6$, corresponding to partial or full immunity from the fine. We will refer to these as our “partial leniency” and “full leniency” treatments, though we admit that this is a slight abuse of terminology, since even under full leniency, firms that report are still subject to the restriction on future prices. In our baseline treatment, as already noted, reporting is not possible, though detection by the competition authority is.

In all three of our treatments, given these parameter values, there exist both anti-competitive subgame perfect equilibria in which both firms choose High in all rounds (unless caught by the competition authority) and pro-competitive subgame perfect equilibria in which both firms choose Medium in all rounds. While there is no subgame perfect equilibrium in which reporting occurs, there are Nash equilibria – along the lines of those described in the previous sections – in which it does, and as noted at the beginning of Section 2, non-subgame-perfect Nash equilibria are often behaviourally relevant in repeated social dilemmas such as our setting, even though the desirable theoretical property of credibility is lacking. In both our full-lenieny and our partial-lenieny treatments, $\delta$ is sufficiently high that collusion and reporting can be supported by a threat to choose Low in every subsequent round in the event of a deviation. However, there exist punishments that will support collusion and reporting under full leniency but not under partial leniency (for example, using the threat to punish deviations by choosing Medium in the next round, then Low in the following round, and continuing alternating between Medium and Low forever).

To summarise the implications of the model, the multiplicity of equilibria in each of our treatments makes point predictions impossible, but the sets of equilibria they yield have the following properties:

1. Collusion is possible, but not certain, in all three treatments.
2. Reporting of collusive behaviour by firms is possible, but not certain, in both leniency treatments.
3. Collusion is easier to support under full than under partial leniency, and easier under partial leniency than in the baseline.
4. Reporting of collusive behaviour by firms is easier to support under full than under partial leniency.

---

10Both the probability of being caught in the absence of reporting, and the fine for being caught, are arguably lower than a regulator would choose given the power to do so. However, setting these on the low side arguably reflects the reality of competition enforcement, where these variables are generally viewed as insufficiently strong to deter collusion due to technological, legal and political constraints. As noted by Connor and Bolotova (2006, p. 1115), “punitive sanctions are the exception not the rule for illegal international price fixing” (p. 1115), leaving only a compensatory component (at most). Setting a low probability and fine for being caught serves an obvious additional purpose in our experimental setting: if collusion is rare without a leniency programme in effect, there is little means to detect a treatment effect from introducing leniency.
For these last two properties, “easier to support” in the previous sections has meant either holding true for a larger set of $\delta$ (i.e., above a lower minimum value) with the type of punishment strategy held constant, or holding true for a larger set of punishment strategies with $\delta$ held constant. Neither of these is completely satisfactory, since we do not vary $\delta$ in our experiment (though other researchers such as Hinloopen and Onderstal 2014 also appeal to this logic while not varying discount factors in their experiment), and we have little understanding of what kinds of punishment strategies subjects will choose and expect from their rivals. However, in Appendix A, we make these arguments more rigorous by adapting the technique of Bigoni et al. (2015), who extend the standard analysis by assuming players have a level of “trust” that anticipated cooperation by other players will actually happen, and showing that our notion of “easier to support” is equivalent to “not requiring such a high level of trust”.

The final important aspect of our experimental design is the opportunity for cheap talk (costless, non–binding communication) between the two firms prior to the first round. We do not attempt to model cheap talk theoretically here, but we note that experimental studies of cheap talk have tended to find that it raises the level of cooperation between players – even when such cooperation is not an equilibrium outcome. We restrict cheap talk to the first round only, so that the relevant solution concept remains subgame perfect equilibrium rather than some notion of renegotiation proofness (Farrell and Maskin, 1989).

2.4 Related literature

In this section, we discuss on the experimental literature on leniency programmes, with particular focus on the studies’ experimental designs – so that we can contrast ours at the end of this section – and on their results most relevant to our study. It should be noted that this literature parallels to some extent other strands of the literature on whistle–blowing incentives. For example, Abbink et al. (2014) investigate the effect on bribery of policies that punish bribe–receivers but not bribe–givers – in effect an asymmetric leniency programme. (See Abbink 2006, Banuri and Eckel 2012, and Serra 2012 for broader surveys of the experimental corruption literature.)

To our knowledge, Apesteguia et al. (2007, hereafter ADS) conducted the first lab experiment involving a leniency programme. Subjects played a one–shot three–firm Bertrand oligopoly game with homogeneous products. Prior to choosing prices (integers between 90 and 100), subjects could opt to communicate; non–binding communication was possible, via a computerised chat room, if all three subjects chose to communicate. ADS defined a cartel to have been formed whenever the subjects so agreed – irrespective of the subsequent price choices – and the competition authority could only detect the cartel if one of the firms reported it. In their baseline treatment, firms could report their cartel activities, but there was no benefit to doing so (no leniency programme was in effect), and there could be a cost, since all colluding firms incurred a fine of 10% of their own revenue if caught (leading to the possibility of negative earnings, though these were not enforced).

Hinloopen and Soetevent (2008, hereafter HS08) extended ADS’s design. Their stage game consisted of a three–firm Bertrand oligopoly game with homogeneous products and integer prices between 101 and 110.

\footnote{For example, Fonseca and Normann (2012) report substantially higher prices in Bertrand oligopoly when communication is possible, as long as the number of firms is small (including the case of duopoly). See also Sally (1995) and Balliet (2010) for meta–analyses of communication in social dilemmas.}
and a basic fine for anti–competitive behaviour (if caught) set to 10% of revenue. They repeated this stage game, with fixed groups playing at least 20 rounds, and with a 0.8 continuation probability thereafter. Subjects could opt for non–binding communication before each round, except in a baseline treatment where there was no communication. Like ADS, a cartel was judged to have been formed if all three subjects chose to communicate. Messages were restricted to a menu of statements of minimum and maximum prices, in contrast to the free–form communication of ADS. In an “antitrust” treatment, HS08 introduced a positive probability of the competition authority detecting collusion (though reporting by firms was not possible); this probability was set to 0.15. In a “leniency” treatment, they additionally allowed firms to report their collusive behaviour, at a nominal cost. The first firm to report received full immunity from the fine, and if a second firm also reported, its fine was reduced by half.12

Bigoni et al.’s (2012, hereafter BFLS12) subjects played a Bertrand duopoly with differentiated products, with integer prices between 0 and 12. They could opt for communication before each round; as with ADS and HS08, communication took place – and a cartel was judged to have formed – if both duopolists so opted, and similarly to HS08, communication took the form of choosing amongst pre–written messages about prices. Unlike the two previous experiments, the basic fine was a fixed amount, so it was positive even when a subject earned zero revenue in a round. Both detection by the competition authority (with probability 0.1) and reporting were possible in all treatments except for their baseline (where communication was also impossible), and reporting could take place either before or after prices were chosen and announced. Subjects played multiple supergames, with the continuation probability set to 0.85.

Hamaguchi et al. (2009, hereafter HKS) modelled a leniency programme within a degenerate two–stage game. In the pricing stage, subjects’ choices were forced: they had to choose the collusive price if collusion had not yet been caught, and the competitive price if it had. In the subsequent reporting stage (reached if collusion had not yet been caught), payoffs were set so that firms preferred to report if and only if at least one rival firm reported, but all firms were better off if none reported; if no firm reported, the competition authority detected collusion with probability 0.1. The game was indefinitely repeated with continuation probability 0.8. HKS varied the number of firms in each market (two or seven) and the consideration given to whistle–blowing firms (reduced fine, no fine or bonus).

We know of two additional experimental studies that were published several years after our experiment took place, and thus involved experiments that likely took place either concurrently with or after ours. Hinloopen and Soetevent (2014, hereafter HS14) describe a classroom duopoly experiment with a setting similar to HKS: each firm chose between a high price and a low price, and collusion was defined as both choosing the high price. The game was indefinitely repeated with continuation probability 0.8, and in an “antitrust” treatment, collusion was detected with probability 0.4 in any round where it occurred, and the fine was the total earning in that round. Two additional treatments involved leniency programmes, with reporting possible in rounds where collusion occurred. In both treatments, a sole whistle–blower received full immunity from fines, while if both reported, each got a 90–percent fine reduction in an “exploitable” treatment or a 50–percent reduction in a “non–exploitable” treatment.

12BFLS12 note that HS08 use abstract framing, in contrast to the firm/cartel context used in our paper and the other papers discussed in this section. It is possible, though far from certain, that framing would impact behaviour in settings like these.
Bigoni et al. (2015, hereafter BFLS15) report a careful study of (mainly) the effects of varying fines and the likelihood of detection. Much of the setting was like BFLS12. In a “laisser–faire” treatment, collusion was not punished, while in a “no report” treatment, any past or current collusion could be exogenously detected (with probability 0.1) but not reported; the fine of 200 was a bit more than the difference between the collusive and competitive prices in a round. The remaining six treatments included a leniency programme, with reporting either before or after prices were announced for the round. These treatments varied the extent of leniency (either no effect, or full immunity from the fine for a single whistle–blower or 50–percent reduction if both report) and the combination of fine and exogenous detection probability (200 and 0.1, 1000 and 0.02, or 1000 and nil).

Finally, we mention two studies of leniency programmes involving collusion in auctions as the setting, rather than collusion in oligopolies. There are important differences between these two settings, such as the incidence of the harm done by collusion, which is typically concentrated on one or a small number of sellers in auctions, as opposed to spread over a large group of consumers in oligopoly. However, the use of the same terminology, and the fact that researchers often work in both areas, suggests that these results may be more relevant to our experiment than those in other related areas (such as bribery, mentioned earlier). Hamaguchi et al. (2007, hereafter HIIKT) consider symmetric 5–player first–price sealed–bid auctions, with subjects not informed about how many rounds would be played. Sellers were automated. In a communication treatment, subject bidders independently chose whether to enter a chat room with unrestricted communication possible; “collusion” meant that a given bidder was one of at least two who chose to enter the chat room. (Presumably, if all bidders colluded, they would agree for each to bid the seller’s reserve price, but communication was non–binding.) In an antitrust treatment, collusion in the current round was detected with probability 0.15, i.i.d. for each subject, and the fine was 10 percent of profits earned in that round and the previous two (though a subject could not be fined twice for the same round). In a leniency treatment, reporting (in a round where there was collusion) earned a 100–percent, 50–percent and 30–percent reduction in fine for the first, second and third whistle–blower.

Hinloopen and Onderstal (2014, hereafter HO) used a similar setting to HIIKT, but with groups of three bidders, with a cartel formed if all three so choose it, and varying between first–price and English auctions. If bidders form a cartel, a designated (though non–binding) winner is randomly chosen by the computer program. In an antitrust treatment, current cartels are detected with probability 0.15, and the fine of 10 is equal to the maximum possible profit. In a leniency treatment, reporting was possible if there was a cartel in the current period, and sole whistle–blowers got full fine reduction, while multiple whistle–blowers got either full or partial reduction, depending on a random draw.

A summary of the results of these experiments is presented in Table 1. For each study and treatment, and for each of the three variables listed, the table shows the effect of moving from no leniency programme to partial leniency, and from partial leniency to either full leniency or a bonus.13

---

13When aggregating multiple studies, some simplifications are necessary. For example, we define partial leniency as anything not guaranteeing full immunity from fines; this includes programmes that give full leniency to a single whistle–blower but give less than full leniency to at least one when there are multiple whistle–blowers. Also, to allow easy comparison between the two auction experiments and the others, “price” in the former is actually the corresponding bidder surplus (which is negatively related to the bid price, the variable actually chosen by subjects). Finally, we leave out degenerate results, such as the increase in reporting associated with moving from a no–leniency treatment with reporting impossible to a leniency treatment with reporting allowed.
Table 1: Selected results from previous leniency–programme experiments

<table>
<thead>
<tr>
<th>Study</th>
<th>Treatment</th>
<th>(N, p)</th>
<th>Collusion definition</th>
<th>Change in degree of leniency</th>
<th>Average price</th>
<th>Collusion level</th>
<th>Reporting frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADS</td>
<td>(3, 0)</td>
<td>All choose to communicate</td>
<td>none → part.</td>
<td>– –</td>
<td>–</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>HS08</td>
<td>(3, 0.15)</td>
<td>All choose to communicate</td>
<td>none → part.</td>
<td>– –</td>
<td>–</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>HKS</td>
<td>duopoly (2, 0.1)</td>
<td>All choose to communicate</td>
<td>none → part.</td>
<td>N/A</td>
<td>– –</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7–firm (7, 0.1)</td>
<td>high price</td>
<td>part. → full бонус</td>
<td>N/A</td>
<td>+ (full), – (bonus)</td>
<td>– (full), + (bonus)</td>
<td>N/A</td>
</tr>
<tr>
<td>BFLS12</td>
<td>(2, 0.1)</td>
<td>All choose to communicate</td>
<td>none → part.</td>
<td>– –</td>
<td>–</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>HS14</td>
<td>(2, 0.4)</td>
<td>All choose to communicate</td>
<td>none → part.</td>
<td>+</td>
<td>+</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>High fine</td>
<td>(2, 0.02)</td>
<td>All choose to communicate</td>
<td>– –</td>
<td>–</td>
<td>–</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>BFLS15</td>
<td>Low fine (2, 0.1)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>No detection</td>
<td>(2, 0)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>HIIKT</td>
<td>FP auct. (5, 0.15)</td>
<td>≥ 2 choose to communicate</td>
<td>none → part.</td>
<td>+</td>
<td>–</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>HO</td>
<td>Eng. auct.</td>
<td>(3, 0.15)</td>
<td>All choose to collude</td>
<td>none → part.</td>
<td>–</td>
<td>–</td>
<td>N/A</td>
</tr>
<tr>
<td>FP auct.</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Key: N: number of firms in each market. p: exogenous probability of detection of collusion. + (–): positive (negative) but insignificant or significance not assessed. + + (– –): positive (negative) and significant. N/A: not applicable (treatment or variable not part of design, variable not measured or not reported in a treatment, etc.).

As the table shows, there are few general results. There is substantial, though not universal, agreement that increasing the level of leniency leads to increased reporting (our direct effect). The impact on prices or collusion (our indirect effect) is less clear, though the preponderance of the evidence suggests a non–monotonic effect: moving from no leniency to partial leniency tends to decrease both prices and collusion, while moving from partial to full leniency or a bonus tends to increase them. So, even though several experimental studies of leniency exist, the lack of consensus in their results suggests a role for additional studies such as ours.

Our experimental design, described in Section 2.3 and with some additional details in Section 2.5 below, differs from the typical paper in this literature in several ways. Here, we describe the main differences, and attempt to give some justification for the choices we made.

The most important difference from much of this literature is in how we define collusion, both within the game setting (e.g., when firms risk being punished by the competition authority, or when leniency is available) and in our analysis of the data (e.g., as a statistic that might vary according to our treatment variables): in both
cases, it is simply *choice of the high price by both firms*.\(^{14}\) HKS and HS14 use the same definition as ours, while most of the other studies in this literature define collusion as *mutual agreement to communicate*, and in particular, collusion in these other studies did not depend on either the content of this communication (whether there was any agreement to set high prices) or the firms’ subsequent pricing.

Our “implementation–based” definition of collusion, like the “agreement to communicate” definition, has the advantage of being simple and easily verifiable within the experiment. (Imagine, by contrast, the logistical issues involved in assessing whether an explicit agreement to collude was made by subjects – in real time during the experimental session – and the consequences of even infrequent errors in assessment.) Unlike the “agreement to communicate” definition, our definition aligns culpability with damage caused; firms risk punishment if and only if their behaviour has harmed consumers.\(^{15}\)

But is our definition realistic? Clearly it is no less realistic than “agreement to communicate”, under which a mutual choice of high prices is neither necessary nor sufficient to establish collusion, while agreeing to communicate with no actual subsequent communication is treated the same way as making a pact about prices.\(^{16}\) However, under our definition, collusion can be established irrespective of whether any explicit agreement was made between the firms, or indeed whether the opportunity to communicate (always available in the experiment) was even taken. Though the former of these would be quite difficult to assess in real time in a lab experiment, the latter could easily be automated.\(^{17}\)

The question of whether such “tacit collusion” (in legal terminology, “tacit coordination” or “coordinated effects”) constitutes collusion is far from settled. Posner (2001) argues forcefully that from an economic standpoint, it is indeed collusion:

> If the economic evidence presented in a case warrants an inference of collusive pricing, there is neither legal nor practical justification for requiring evidence that will support the further inference that the collusion was explicit rather than tacit. From an economic standpoint it is a detail whether the collusive pricing scheme was organized and implemented in such a way as to generate evidence

\(^{14}\)Our use of a simplified price choice set – only 3 possible prices rather than 10–13 as in much of the rest of this literature – has the advantage of making it possible to identify collusive pricing unambiguously, as only one price (High) is consistent with collusion. A disadvantage, though perhaps a minor one, is that it makes our experiment unsuitable for studying differences in the severity of collusion (e.g., if prices from 101 to 110 are possible, collusion at 110 is worse for consumers than collusion at 102). HS14 had only two prices, giving them the same advantage and disadvantage as us. HKS also had only two prices, though as noted above, prices in their experiment were imposed on the firms rather than chosen.

\(^{15}\)This is a step in the direction of assigning liability based on harm to the victim (Polinsky and Shavell, 1994), as compared to proscribing all communication amongst firms, or even explicit verbal agreements to collude.

\(^{16}\)Despite Adam Smith’s famous assertion that “[p]eople of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the public, or in some contrivance to raise prices” (Smith, 1776, Book I, Chapter 10, Paragraph 82), which taken literally would mean that all communication amongst firms leads to collusion, real competition policy does not proscribe communication per se (as HS08 also acknowledge; their footnote 4 states “[s]trictly speaking, having a discussion about prices is not illegal”). Indeed Smith’s famous quote’s less famous continuation makes clear that he does not advocate outlawing communication per se: “It is impossible indeed to prevent such meetings, by any law which either could be executed, or would be consistent with liberty and justice” (Smith, Paragraph 82).

\(^{17}\)In Appendix B, we analyse the communication between firms in our experiment, including factors affecting the likelihood of verbal agreement, as well as the connection between such verbal agreements and collusive behaviour under our “implementation–based” definition and under the “agreement to communicate” definition used by other authors.
of actual communications. It is not a detail sanctified by the language of section 1 of the Sherman Act... (p. 94).

Posner goes on to list examples of US Supreme Court decisions in which evidence of explicit communication was not required to prove collusion (e.g., “American Tobacco Co v. United States actually says that ‘a tacit meeting of the minds’ satisfies the requirement of proving unlawful concerted action’. ...”, p. 95). He also identifies the main barrier to prosecuting tacit collusion: “the difficulty of proving collusive pricing by economic evidence, given the complex, technical, and often inconclusive, or even equivocal, character of such evidence” (p. 98). That is, it is not that there is any ambiguity about whether tacit collusion is illegal, but from a practical standpoint it will often be difficult to know, let alone prove, that a given profile of prices or price changes reflects collusion as opposed to more benign factors such as input costs or consumer demand. Evidence of a collusive agreement, therefore, is the “smoking gun” that in practice is needed to show that illegal behaviour has occurred; it is not the illegal behaviour itself.

Other countries’ competition authorities – in particular, the two in which this experiment was conducted – have also stated explicitly that tacit collusion constitutes illegal collusive behaviour. Japan’s submission to the 2006 OECD Global Forum on Competition states, “The Japan Fair Trade Commission...bases its approach on the theory that explicit agreement among the entrepreneurs is not necessary to prove a cartel agreement; i.e., ‘liaison of intention,’ and a tacit agreement suffices. In the Toshiba Chemical case, which involved a cartel without direct evidence, the Tokyo High Court recognised this theory” (p. 127). In the UK, the then–Office of Fair Trading (now part of the Competition and Markets Authority) noted that “the nature of the market may mean that undertakings might adopt the same pricing policy on the market without ever explicitly agreeing on price” (OFT, 2004, p. 8).

Returning to the comparison between our experimental design and others’, a second important difference is our inclusion of the Low price, giving subjects the possibility of pricing below the competitive level, and just as importantly, holding the rival firm to a profit below the competitive profit. Besides obviously reflecting reality, this feature provides subjects with punishment strategies that can be used to support many interesting and empirically relevant kinds of behaviour in equilibrium, even if the Low price itself is rarely observed. As a notable example, we have already seen (Section 2.3) that in our leniency treatments, reporting by firms is consistent with equilibrium, even along the equilibrium path itself. BFLS12 and BFLS15 also allow low prices that can serve as punishment devices, though their focus on subgame perfect equilibrium means they do not consider this possibility. The other oligopoly experiments in this literature do not allow sub–competitive pricing.

A third difference is that we only allow reporting in a round in which anti–competitive behaviour actually occurred. This is also true of HKS, HS14 and the two auction experiments, while HS08, BFLS12 and BFLS15

---

18Similarly, Mezzanotte (2009) argues that while tacit collusion is illegal under EU law under Article 82, enforcement is essentially impossible due to this attribution problem. (See also Kaplow (2011), pp. 449–450 and Kovacic et al. (2011), p. 396.) However, it should be clear that this problem does not arise in our theoretical and experimental settings, given the small price choice set and the stationarity of costs and demand.

19Some countries have gone the other way, however. Australia’s competition law defines collusion as explicit collusion only, making tacit collusion legal. It is unclear whether this is due to a genuine belief that high prices without communication are socially desirable, or to capture.
allow reporting in any round after communication took place. Our design choice removes the possibility that reporting is used (or threatened) as retribution for deviating from collusion, so that reporting serves the purpose that the competition authority intends: breaking collusive arrangements currently in place.\textsuperscript{20}

A fourth difference between our setting and the previous experiments’ is that we only allow communication in the first round of a supergame – unlike HS08, BFLS12, BFLS15 and HIIKT who allow it in all rounds, and also unlike HKS, HO14 and HS14 who do not allow it at all. As noted in Section 2.3, our choice to have communication only at the beginning maintains standard equilibrium concepts (rather than alternatives such as renegotiation–proofness) as the source of theoretical predictions. Fifth, unlike all of the above papers except for HKS, our punishment for anti–competitive behaviour includes not only fines, but also a restriction on conduct (prices) in future rounds. This is meant to reflect the likelihood that when such behaviour has been caught in the past, the competition authority pays closer attention to the industry in the future.\textsuperscript{21} We believe that this captures some aspects of reality – in particular, that even under full leniency, a reporting firm does not get off completely scot–free – though the downside is that our experiment cannot be used to study recidivism, whereas several of the other studies can and do.

Summing up, it is clear that there are many differences between our experiment and others from the literature. We have presented some justification for our particular design choices here, and hope to have convinced some readers that our choices are reasonable. Obviously, this does not preclude that the choices made in other studies are also reasonable. It is not our view that there is only one “correct” choice out of any set of possible experimental design choices; arguing in favour of our choices is not a claim that others were wrong. As noted in the introduction, all of the experimental designs in this literature sacrifice substantial realism, as does ours. Given that experiments cannot be completely realistic, it becomes important to verify that observed results are not due to the particular unrealistic assumptions that were introduced. Thus there is always a role for new experiments using somewhat–different settings to understand previously–examined questions; this is part of the contribution of the current paper (along with simply adding to the weight of evidence regarding the effects of leniency programmes, given the incomplete agreement in the literature thus far, as noted in the discussion surrounding Table 1).

\subsection*{2.5 Experimental procedures}

We implemented infinite repetition with discounting by means of a fixed termination probability (equal to $\delta$, the discount factor) at the end of each round. This is a standard technique in experimental economics, and is methodologically sound as long as individuals are expected–utility maximisers with additively–separable (supergame) utility functions, as the corresponding objective functions are identical.

\textsuperscript{20}Hinloopen and Soetevent (2008) emphasise a distinction between “exploitable” and “non–exploitable” leniency programmes (see also Spagnolo 2004). A leniency programme is exploitable if the joint profits from forming a cartel and then reporting are higher than those from either maintaining a cartel or from under–cutting. Both of the leniency programmes used in our experiment are non–exploitable. Hinloopen and Soetevent (2014) compare behaviour under exploitable and non–exploitable leniency programmes.

\textsuperscript{21}Other justifications for this design feature are (1) an enforceable consent decree is put in place, under which firms must lower their prices; or (2) after a finding of collusion, the competition authority gains understanding of firms’ cost structures, and hence can more easily identify when prices are supra–competitive.
The experimental sessions were conducted in 2009 and 2010, and took place at the the University of Aberdeen’s Scottish Experimental Economics Laboratory (SEEL) and Kyoto Sangyo University’s Kyoto Experimental Economics Laboratory (KEEL); ethics approval was given by both universities. There were 16 sessions at SEEL, each with between 8 and 16 subjects, while the seven sessions at KEEL had 22–28 subjects each. Subjects in both locations were primarily undergraduate students, recruited using ORSEE (Greiner 2015) from databases of people expressing interest in participating in economics experiments. No one took part more than once; there were no other exclusion conditions.

Each session comprised at least five supergames; a few sessions that finished the first five especially quickly (due to the realisations of the draws for random termination) played a sixth supergame. In the first three supergames, there was never a leniency programme; we call this “Part 1” of a session – and the remainder of the session “Part 2”. Having competition policy but no leniency programme in Part 1 of all sessions gave subjects an opportunity to become familiar with what is a fairly complex setting; it will also allow a measure of control for unobservable subject characteristics in our later regression analysis (see Section 3.2).

In five of the sessions, Part 2 also had no leniency programme; these sessions make up our baseline treatment. (Some information is displayed in Table 2 at the beginning of the results section.) In the remaining eighteen sessions, Part 2 had a leniency programme in which both firms could receive a reduced fine by reporting to the competition authority: eight sessions with partial leniency \((R = 3 < F)\) and ten with full leniency \((R = 6 = F)\).

At the beginning of a session, subjects were seated in a single room and, after signing consent forms, were given written instructions for Part 1.\(^{22}\) The instructions stated that the experiment would comprise two parts, but details of the second part would not be announced until after the first part had ended. The instructions were also read aloud to the subjects, in an attempt to make the rules of the game common knowledge. Then, the first round of play began (there was no instructions quiz). Because each supergame involved indefinite repetition, the total number of rounds played varied across sessions. After the third supergame was completed, each subject was given instructions for Part 2. This was done even in the baseline treatment, where the game did not change. (In this case, the new instructions were quite short – stating essentially just that the game was not changing.) These new instructions were also read aloud, before the fourth supergame was played.

The experiment took place on networked computer terminals, using z–Tree (Fischbacher 2007). Subjects were asked not to communicate with other subjects except via the computer program. Subjects were randomly matched at the beginning of the first round of a supergame, but pairings were then fixed until the supergame ended. No identifying information was given to subjects about their opponents (in an attempt to minimise incentives for reputation building and other supergame effects). Rather than using potentially biasing terms like “opponent” or “partner” for the other player, we used neutral though somewhat more cumbersome terms such as “other firm”, “firm matched to you” and similar phrases.

Each round began with subjects being prompted to choose a price (High, Medium or Low). In the first round of each supergame, the screen also contained a “chat room” so that subjects could send and receive cheap–talk messages (see Figure 4 for the version used in the UK). Subjects could send as many or as few messages as

\(^{22}\)The instructions from the partial–leniency treatments conducted in the UK are in Appendix C. Instructions from the remaining treatments (including those conducted in Japan), and other experimental materials, are available from the corresponding author upon request.
they wished; they were instructed not to send messages containing (a) personal or identifying information or (b) physical threats, but messages were otherwise unrestricted. Messages were visible until a price was chosen or until the time available for communication (75 seconds) ran out. Subjects could not observe other pairs’ messages.

Once all subjects had chosen their prices, a feedback screen informed them of their own price and the opponent price. In supergames with a leniency programme in effect, subjects reaching a (High, High) outcome were prompted to choose whether to report or not to report their collusive activities to the competition authority (see Figure 5). (If the leniency programme was not in effect, or if any other outcome was reached, subjects were simply asked to click a button to continue.) The next screen told subjects the complete result for the round, including (if applicable) whether either subject in the pair reported collusive behaviour, whether collusive
behaviour was discovered, and both subjects’ profits for the round. The round then ended with subjects being notified whether the supergame would continue for another round or end with the current round.

At the end of the fifth or sixth supergame, the experimental session ended. For each subject, one round from each of the supergames was randomly chosen, and the subject was paid his/her earnings in those rounds. Each “lab pound” was exchanged for £0.50 of real money in the UK, and for 150 yen in Japan (rounded to the nearest £0.50 or 1 yen); there was no show–up fee. Total earnings averaged about £15 in Aberdeen and 2400 yen in Kyoto, for a session typically lasting 60–75 minutes.
2.6 Hypotheses

The theoretical analysis of our three treatments (no leniency, partial leniency and full leniency), with implications outlined in Section 2.3, allow us to formulate the following null hypotheses.

**Hypothesis 1**  The fraction of subject pairs reaching at least one (High, High) outcome is the same in all three treatments.

**Hypothesis 2**  The overall frequency of (High, High) outcomes is the same in all three treatments.

**Hypothesis 3**  The overall frequency of High price choices is the same in all three treatments.

**Hypothesis 4**  The frequency of reporting, conditional on a (High, High) outcome, is the same in the two treatments with reporting.

**Hypothesis 5**  Average profits are the same in all three treatments.

Of these five null hypotheses, two have natural directional alternatives. Hypothesis 1 relates to the incentives to collude. As noted in the introduction and discussed in Section 2, a potential indirect effect of leniency programmes is that by reducing the down–side risk to colluding, they lead to more cartels forming. If so, then we would expect the fraction of pairs with at least one (High, High) outcome to increase from the baseline to the partial–leniency treatment to the full–leniency treatment. Similarly, Hypothesis 4 refers to reporting frequency, and the analysis of the model suggests that reporting might be more likely under full leniency than under partial leniency. The remaining three hypotheses – Hypotheses 2, 3 and 5 (concerning the extent of collusion, price levels, and firms’ profits respectively) – have no clear directional alternative, as the direct effect of increased reporting and the indirect effect of increased incentives to collude pull in opposite directions.

3 Experimental results

In total, there were 23 experimental sessions with 384 subjects (see Table 2 below). Recall that there was no leniency programme in effect in Part 1 (the first three supergames) of any session, while in Part 2 (the remaining supergames), the leniency programme varied as shown in the table.

Throughout our analysis, unless stated otherwise, our unit of observation is a supergame–pair: an entire supergame played between a pair of subjects. This is more useful for our purposes than the typical analysis focussing on individual subject choices in each round, as it allows us to control for differences in supergame length within and across sessions (and treatments), and avoids excessively weighting results from the longer supergames. In Section 3.1, we present some descriptive statistics about collusive attempts, successful collusion, and the breakdown of collusion. In Section 3.2, we use parametric regressions to test the suggestive results we find in the descriptive statistics.\(^{23}\)

\(^{23}\)We do not analyse the messages sent by subjects during the pre–play communication stage here, but we do so in Appendix B.
Table 2: Session information

<table>
<thead>
<tr>
<th>Location</th>
<th>Treatment</th>
<th>Sessions</th>
<th>Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aberdeen</td>
<td>No leniency</td>
<td>3</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>(SEEL) Partial leniency</td>
<td>6</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>Full leniency</td>
<td>7</td>
<td>90</td>
</tr>
<tr>
<td>Kyoto</td>
<td>No leniency</td>
<td>2</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>(KEEL) Partial leniency</td>
<td>2</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>Full leniency</td>
<td>3</td>
<td>82</td>
</tr>
</tbody>
</table>

3.1 Descriptive statistics

Table 3 reports some aggregates from the experiment, along with \( p \)-values from tests of pairwise significant differences between treatments in Part 2. In computing these \( p \)-values, we use only session–level data and two–sided rejection regions, even when the alternative hypothesis is directional. This means that our non–parametric tests are very conservative (i.e., they under–state significance). To get a full picture of the significance of our treatment effects, these results should be viewed in conjunction with the regression results in later sections (which tend to over–state significance).

Table 3: Descriptive statistics, (supergame–pair)–level data

<table>
<thead>
<tr>
<th></th>
<th>Part 1 baseline</th>
<th>Part 2 partial leniency</th>
<th>Part 2 full leniency</th>
<th>pairwise ( p )-values base/</th>
<th>pairwise ( p )-values base/</th>
<th>pairwise ( p )-values part/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconditional</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At least one H choice (collusion attempt)</td>
<td>.851</td>
<td>.919</td>
<td>.862</td>
<td>.914</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>At least one HH outcome (cartel formed)</td>
<td>.424</td>
<td>.586</td>
<td>.538</td>
<td>.555</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>Fraction of L choices</td>
<td>.009</td>
<td>.010</td>
<td>.011</td>
<td>.004</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>Fraction of H choices (average price)</td>
<td>.444</td>
<td>.603</td>
<td>.440</td>
<td>.450</td>
<td>0.047</td>
<td>0.085</td>
</tr>
<tr>
<td>Fraction of HH outcomes (collusion extent)</td>
<td>.283</td>
<td>.463</td>
<td>.354</td>
<td>.322</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>Gross average payoff</td>
<td>6.430</td>
<td>7.576</td>
<td>6.487</td>
<td>6.550</td>
<td>0.047</td>
<td>0.085</td>
</tr>
<tr>
<td>Net average payoff</td>
<td>6.294</td>
<td>7.299</td>
<td>6.076</td>
<td>6.284</td>
<td>0.040</td>
<td>0.077</td>
</tr>
<tr>
<td>Conditional on at least one HH outcome</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reporting (by at least one firm)</td>
<td>—</td>
<td>—</td>
<td>.218</td>
<td>.450</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Under–cutting (by at least one firm)</td>
<td>.347</td>
<td>.241</td>
<td>.141</td>
<td>.101</td>
<td>n.s.</td>
<td>0.040</td>
</tr>
<tr>
<td>Endogenous breakdown (report, under–cut)</td>
<td>.347</td>
<td>.241</td>
<td>.359</td>
<td>.541</td>
<td>n.s.</td>
<td>0.008</td>
</tr>
</tbody>
</table>

\( p \)-values from two–sided robust rank–order tests on session–level data (n.s.: \( p > 0.20 \))

We define a “collusion attempt” as a High price choice by either subject in any round of a supergame–pair. So for example, the frequency of 0.919 in Part 2 of the baseline treatment means that in just under 92 percent of supergame–pairs, there was at least one High price choice, while in the other 8 percent, no–one chose High at any time. Similarly, “cartel formed” is a (High, High) outcome (which we also abbreviate as HH) occurring in one or more consecutive rounds at any time during the supergame for that pair of subjects; we also call this “successful collusion”. This is distinct from “collusion extent”, which is the fraction of rounds within a supergame–pair that the HH outcome occurs, and serves as one measure of harm to consumers.

Collusion can last until the supergame ends, or it can break down in one of three ways: reporting by one or both of the subjects in the pair, exogenous detection by the competition authority, or by “under–cutting”, which we define as any choice by either subject of a Medium or Low price following a previous–round HH outcome that had not been detected. Under–cutting, unlike reporting or exogenous detection, allows the possibility of future collusion, so it is possible for a supergame–pair to successfully collude more than once. Finally, since nearly all price choices were either Medium or High (Low prices represented fewer than 1 percent of price choices overall: 1.2 percent when punishment was in effect, and 0.9 percent when all three prices were available), the fraction of High choices will serve as our measure of “average price”.

In Part 1 of the experiment, where there is no leniency programme, most pairs attempt to collude, with about 85 percent having at least one High price choice by a firm. Just under half of pairs successfully collude at least once, and overall they collude in almost 30 percent of rounds. In Part 2 of the baseline sessions, both high prices and collusion become more frequent, with nearly 60 percent of pairs colluding at least once and HH outcomes occurring just under half of the time. Since there is no difference in any strategic aspect of the game between Part 1 and Part 2 of the baseline, these increases seem to be due to increased cooperation between subjects as they become more experienced in this setting.

The table shows no evidence of the conjectured “indirect effect” of leniency programmes, as neither leniency treatment shows any increase in attempts to collude or successful cartel formation. There is, however, a significant decrease in prices in both leniency treatments compared to the baseline, as well as a decrease (though insignificant) in the fraction of rounds in which collusion occurred. These decreases are attributable to the “direct effect” of the leniency programme: just over 20 percent of pairs in the partial–leniency treatment, and nearly half in the full–leniency treatment, end collusion by reporting, with significantly more reporting under full leniency than under partial leniency. There is a corresponding decrease in payoffs under either leniency programme compared to the baseline, though this decrease (like the decrease in the extent of collusion) is smaller than it should be, because the increase in reporting is partly offset by a decrease in under–cutting. This suggests, that some – though not all – of reporting is done as a substitute for under–cutting, though it must be emphasised that the total frequency of endogenous cartel dissolution (reporting or under–cutting) does increase from no leniency to partial leniency to full leniency (though only the former increase is significant). Finally, even the reduced payoffs in the leniency treatments are well above the competitive level of 4.

Figure 6 shows some additional descriptive statistics about collusion. Here we depart slightly from taking the supergame–pair as the unit of observation, and instead take each episode of successful collusion as an individual

25To be precise, there are 478 supergame–pairs with exactly one successful collusion, while 19 pairs (3.7 percent) collude exactly twice, 3 (0.6 percent) collude three times and 1 pair (0.2 percent) colludes four times.
The distinction matters since a small fraction of supergame–pairs successfully colludes more than once (see Note 25). The left panel shows that the vast majority of cartels (over 85 percent) are formed in the first round of a supergame. The fraction is even higher (nearly 90 percent) in Part 2 of the experiment, and does not vary substantially across treatments (90 percent in the baseline, 92 percent under partial leniency, 88 percent under full leniency). The right panel shows that collusion is often successful for only a single round; one–round cartels form a majority of all cartels overall, and nearly a majority (47 percent) in Part 2. There is some variation here across treatments, consistent with leniency programmes’ reducing cartel stability: one–round cartels comprise 33 percent of all Part–2 cartels in the baseline, 42 percent under partial leniency, and 59 percent under full leniency.

Cartels that last for only one round not only predominate, they are different in nature from longer–lasting cartels. Figure 7 categorises all episodes of collusion according to whether they ended endogenously (due to subject decisions – either reporting or price under–cutting) or otherwise (exogenously due being detected by the competition authority without having been reported, or only by the supergame ending, which for present purposes we classify as an exogenous ending). Over all supergames and treatments, over half of the cartels lasting for exactly one round ended endogenously. In Part 2, 45 percent of cartels ended endogenously in the baseline, compared to 65 percent under partial leniency and 77 percent under full leniency. Also varying was the proportion of endogenous breakdowns due to reporting versus under–cutting: 73 percent of endogenous breakdowns were due to reporting in the partial–leniency treatment and 83 percent under full leniency, as compared to zero in the baseline. By contrast, once collusion survived into a second round, it was unlikely to break down due to decisions by the firms themselves: only about one–fifth of cartels lasting 2–4 rounds ended endogenously, and none of the cartels lasting for 5 rounds or longer did.

These properties of collusion, along with the descriptive statistics reported in Table 3, suggest that introducing a leniency programme does lead to changes in firm behaviour. It does not appreciably affect the likelihood of cartels forming, but it reduces their stability. A substantial fraction of cartels break down via reporting, though
this fraction overstates the leniency programmes’ effect on collusion, since there is a partially offsetting decrease in under-cutting.

3.2 Parametric statistical analysis

We next present results from several regressions. We estimate eight models, each with a different left-hand-side variable, using either the Part–2 data or, when appropriate, the subset of supergame–pairs that reached an HH outcome (i.e., a successful collusion) at least once. The Part–1 data are used only to construct a variable to control for subjects’ intrinsic proclivity toward collusion, as described below.

Our first model is meant to shed light on the incentives to collude induced by the leniency programmes; the dependent variable is “HH indicator”, equal to one if the pair achieves an HH outcome in any round. Our second model looks at the overall prevalence of collusion; the dependent variable is “HH fraction”, the fraction of rounds in which collusion took place. Our third model assesses the effects on prices overall; as mentioned earlier, since Low price choices almost never occur, the fraction of High (versus Medium) price choices can be taken as a measure of the average price. Our next three models look at the effects of the leniency programmes on the two endogenous ways of breaking cartels individually (using indicators for reporting and under-cutting), and for either kind of endogenous breakdown (an indicator with value 1 if either subject reported or under-cut). Our last two models look at the leniency programmes’ effects on firm profitability; dependent variables are gross and net profits.

We used similar right-hand-side variables in each regression. Our main explanatory variables are indicators for the partial– and full–leniency treatments (except for the “Report” regression, which does not use the data from the baseline treatment, and hence we dropped the partial–leniency treatment indicator). We include the supergame number to control for changing behaviour over time, an indicator for the session taking place in Kyoto to control for any subject–pool effects, and a constant term. Finally, we included an index of the pair’s
collusive or cooperative tendencies, called “collusiveness”; this was simply the fraction of High choices within the pair in the first round of each supergame in Part 1 of the session (thus taking on values that are whole-number multiples of one-sixth). We use Tobit models for the two profit variables, H fraction and HH fraction, and probits for the others. All of the models were estimated using Stata (version 12) with robust standard errors.

The results, displayed in Table 4, largely confirm what was observed in the descriptive statistics.26 The

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>HH indicator</th>
<th>HH fraction</th>
<th>H fraction</th>
<th>Report</th>
<th>Undercut</th>
<th>Endog. break.</th>
<th>Gross profit</th>
<th>Net profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subsample:</td>
<td>Full sample</td>
<td>At least one HH outcome</td>
<td>Full sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>partial leniency</td>
<td>-0.073</td>
<td>-0.352**</td>
<td>-0.252***</td>
<td>-0.076</td>
<td>0.129</td>
<td>-1.018***</td>
<td>-1.238***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.068)</td>
<td>(0.180)</td>
<td>(0.072)</td>
<td>(0.049)</td>
<td>(0.093)</td>
<td>(0.276)</td>
<td>(0.277)</td>
<td></td>
</tr>
<tr>
<td>full leniency</td>
<td>-0.099</td>
<td>-0.493***</td>
<td>-0.275***</td>
<td>0.318***</td>
<td>-0.117**</td>
<td>0.311***</td>
<td>-1.126***</td>
<td>-1.193***</td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td>(0.168)</td>
<td>(0.065)</td>
<td>(0.057)</td>
<td>(0.054)</td>
<td>(0.082)</td>
<td>(0.241)</td>
<td>(0.244)</td>
</tr>
<tr>
<td>Kyoto session</td>
<td>-0.013</td>
<td>-0.000</td>
<td>0.049</td>
<td>-0.071</td>
<td>0.026</td>
<td>-0.026</td>
<td>0.077</td>
<td>0.125</td>
</tr>
<tr>
<td></td>
<td>(0.050)</td>
<td>(0.126)</td>
<td>(0.051)</td>
<td>(0.057)</td>
<td>(0.047)</td>
<td>(0.066)</td>
<td>(0.203)</td>
<td>(0.202)</td>
</tr>
<tr>
<td>supergame #</td>
<td>0.016</td>
<td>0.042</td>
<td>0.009</td>
<td>-0.090</td>
<td>0.009</td>
<td>-0.089</td>
<td>0.056</td>
<td>0.078</td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>(0.090)</td>
<td>(0.037)</td>
<td>(0.040)</td>
<td>(0.031)</td>
<td>(0.045)</td>
<td>(0.147)</td>
<td>(0.146)</td>
</tr>
<tr>
<td>collusiveness</td>
<td>0.751***</td>
<td>2.135***</td>
<td>0.854***</td>
<td>-0.130</td>
<td>-0.130</td>
<td>-0.292*</td>
<td>3.276***</td>
<td>3.015***</td>
</tr>
<tr>
<td></td>
<td>(0.105)</td>
<td>(0.300)</td>
<td>(0.105)</td>
<td>(0.128)</td>
<td>(0.104)</td>
<td>(0.154)</td>
<td>(0.363)</td>
<td>(0.367)</td>
</tr>
<tr>
<td>N</td>
<td>442</td>
<td>442</td>
<td>442</td>
<td>245</td>
<td>245</td>
<td>245</td>
<td>442</td>
<td>442</td>
</tr>
<tr>
<td>[ln(L)]</td>
<td>276.12</td>
<td>433.92</td>
<td>356.10</td>
<td>122.59</td>
<td>98.38</td>
<td>154.55</td>
<td>950.17</td>
<td>943.10</td>
</tr>
</tbody>
</table>

* (**, ***): Coefficient significantly different from zero at the 10% (5%, 1%) level. Endog. break = endogenous breakdown (report or undercut)

insignificant and near–zero marginal effects of the partial– and full–leniency variables on the HH indicator confirm that despite potentially providing incentives for firms to collude, these leniency programmes do not make cartels more likely to form. Moreover, they lead to lower average prices, as shown by their negative effect on the H fraction, and they make cartels less stable, as shown by their negative effect on the HH fraction. As with the aggregate data, we can see that these effects on prices and the extent of collusion are driven primarily by reporting, which is significantly more likely under the full–leniency programme than under partial leniency. Under–cutting is significantly less likely under full leniency than in the baseline, while there is no significant difference between partial leniency and the baseline (though there is also no significant difference between partial and full leniency). Endogenous breakdowns are more likely under full leniency than either the baseline or under partial leniency (p ≈ 0.012 for the latter comparison). Both leniency programmes have negative effects on both gross and net profits, but these effects are smaller than they might have been due to the decrease in under–cutting mentioned in the discussion of aggregate results.

26 We are primarily interested in the effects of the partial– and full–leniency treatments, but some of the other variables’ effects are worth a brief mention. The strong positive effects of our collusiveness variable on prices, collusion and payoffs (and weak negative effects on the breakdown variables) suggests that there may be intrinsic subject heterogeneity in propensity to collude. The Kyoto–session variable has no significant effects, suggesting that subject–pool effects were limited. Finally, the supergame also has no significant effects, suggesting minimal time variation.
4 Summary and discussion

Whistle-blowing incentives – where individuals with inside knowledge of illegal or unethical activity receive some reward for reporting this activity – are an important and growing facet of government policy in areas such as competition and corruption. Intuitively, it is reasonable to expect that introducing such incentives, or increasing them further, should make whistle-blowing more likely (the incentives’ direct effect. However, a perverse indirect effect is also possible: if these incentives reduce the cost of future exit from some illegal or unethical arrangement, individuals may perceive the potential risks of forming such an arrangement in the first place as lower, and thus may be more willing to enter into one.

In this paper, we have focussed on one kind of whistle-blowing incentive: leniency programmes aimed at price-fixing firms. We developed a simple theoretical model that captures the essence of firms’ decision making under competition policy, with or without a leniency programme. While multiple equilibria make sharp point predictions impossible, there is support for both effects: raising the incentive to blow the whistle makes it easier to support equilibria in which firms report, but also makes it easier to support equilibria in which they collude in the first place. The net impact on the extent of collusion, and on associated market variables such as prices and firms’ profits, are therefore ambiguous: they may increase, decrease or remain the same.

Because of this intuitive and theoretical indeterminacy, we conduct a laboratory experiment in order to gain empirical evidence from a controlled setting about the relationship between leniency programmes and anti-competitive behaviour. There are three treatments: one with no leniency programme, one where leniency involves a reduced fine (partial leniency) and one where it involves complete immunity from the fine (full leniency). The main results are as follows.

**Result 1** There is no increase (significant or otherwise) in the fraction of subject pairs successfully colluding (i.e., reaching at least one (High, High) outcome) from the baseline treatment to either leniency treatment.

**Result 2** The overall extent of collusion (i.e., frequency of (High, High) outcomes) is significantly lower under full leniency than in the baseline. The frequency is also lower under partial leniency than in the baseline, but the difference is not significant.

**Result 3** Average prices (i.e., the frequency of High price choices) are significantly lower under either full or partial leniency than in the baseline.

**Result 4** The frequency of reporting, conditional on successful collusion, is significantly higher under full leniency than under partial leniency.

**Result 5** Average profits are significantly lower under either partial or full leniency than in the baseline.

These results correspond to Hypotheses 1–5 in Section 2.6 respectively. Additional results are (a) the increase in reporting observed under either leniency programme compared to the baseline is partly off-set by a decrease in under-cutting; (b) when cartels form, they typically form immediately; and (c) both reporting and under-cutting tend to involve relatively new cartels (both fall off quickly after the first round of collusion, and neither occurs after five rounds of collusion).
As always, caution should be taken in drawing conclusions about the outside world based on the outcome of an experiment, given its admittedly artificial laboratory environment, and the many simplifying assumptions made about the available decisions, timing, parameter values, and so on. But with that caveat acknowledged, our results might be interpreted as follows. The effects of introducing a leniency programme, or increasing its generosity, are positive on balance: the direct effect of increased break-down of collusion via greater whistle-blowing is fairly strong, and the potential indirect effect of reduced deterrence is not observed (suggesting that this may be at worst a second-order effect). The net impact includes less time spent under collusion, lower average prices, and firm profits closer to competitive levels. However, several qualifications apply. First, some of the increased reporting comes from firms that would have deviated from collusion anyway (that is, they report instead of under-cutting the rival’s price); this crowding-out implies that the leniency programme’s visible effects will be smaller than the reporting frequency on its own would indicate. Second, most of the gains from introducing a leniency programme can be achieved with partial leniency; the additional benefits from full leniency are smaller and typically not significantly different from zero. Finally, neither leniency programme has much ability to dissolve collusion that has become well established.

On balance, our results should inspire optimism about the potential for leniency programmes to reduce anti-competitive behaviour. Some of our findings – such as the decreases in prices and collusion under partial leniency compared to no leniency, the increased reporting to the competition authority under partial leniency, and the further increase in reporting under full leniency – are broadly in line with results seen elsewhere in the experimental literature (see Table 1). Since – as we have noted repeatedly – the definition of collusion we use is substantially different from that used in several of these previous experiments, it is reassuring that the broad effects of leniency programmes are robust to how collusion is measured.

By contrast, our findings of further decreases in prices and collusion under full leniency (compared to partial leniency) are at odds with several other studies findings of increases when moving from partial to full leniency or to bonuses. We do not view our results as contradicting theirs, especially given that the differences found in these studies have typically not been significant (and ours was also insignificant). At most, these results illustrate a fairly general point about how details of the leniency programme, and other aspects of the strategic environment, can influence the implications of a lab experiment based on that leniency programme – just as they would influence outcomes in the outside world. In both their full–leniency and bonus treatments, Hamaguchi et al. (2009) did not give subjects a choice of price; rather, high prices (and thus collusion) were imposed until collusion was detected, so the increased collusion they found in their full–leniency treatment simply reflected a decrease in reporting (and conversely, the decreased collusion in their bonus treatment reflected an increase in reporting). Apesteguia et al. (2007) used a one–shot setting, so the effects of cartel dissolution on future prices was not factored into their calculation of average prices, which would likely have been lower if they had. Bigoni et al. (2012) conjecture that the increased collusion in their bonus treatment was due precisely to the bonus (that is, it would not have been seen under full leniency alone). In their words, subjects “may have formed cartels with the intent of fooling their competitor by undercutting the agreed–upon price and simultaneously reporting the cartel so as to cash in the reward” (p. 385). Clearly, further research is still needed to improve our understanding of how leniency programmes should be designed in order to maximise their benefits.
References


Mezzanotte, F.E. (2009), “Can the Commission use Article 82EC to combat tacit collusion?” CCP working paper 09–5, University of East Anglia.


A Extension to the theoretical analysis

In this section, we extend the theoretical analysis of Section 2. There, we had motivated our hypotheses by examining changes in the discount factor $\delta$ needed to support some aspect of behaviour, with the maintained assumption that as some behaviour is supported by a larger set of discount factors, it is more likely to occur. Here, we take a different approach, by associating the likelihood of a particular behaviour with the “level of trust” players need to have in each other’s ability and willingness to sustain cooperation. This approach is used by Bigoni et al. (2015) (see also Fehr 2009 and Sapienza et al. 2013).

Suppose firms imperfectly trust each other to cooperate. Specifically, suppose that in any situation where a choice of the High price by the rival firm is anticipated, the firm believes the rival actually will choose High with probability $\beta \in (0, 1)$, and defects by choosing Medium with probability $(1 - \beta)$. (So, as $\beta \to 1$, our analysis reduces to that in the main text.)

As a simple example, apply our extension to the repeated underlying game in Figure 1 – without competition policy or leniency. It is straightforward to verify that (Medium, Medium) is still the unique stage–game Nash equilibrium (since $\beta > 0$ by assumption); this holds as well for the baseline and leniency games used in our experiment. So there continue to be pro–competitive subgame perfect equilibria when $\beta$ is less than 1. But as before, our interest is in when anti–competitive equilibria also exist.

Suppose both firms choose the trigger strategy “Choose High until either firm chooses Medium or Low; then choose Medium”. If $\beta = 1$, this would result in the collusive outcome (High, High) being played in every round. For $\beta < 1$, a firm’s subjective expected discounted payoff is given by

$$\pi_c = \beta (10 + \delta \pi_c) + (1 - \beta)(2 + 4\delta + 4\delta^2 + ...).$$  \hspace{1cm} (6)

With probability $\beta$, collusion is maintained until the next round, so that the continuation payoff is the current payoff, discounted once. With probability $(1 - \beta)$, the rival firm defects, and both choose Medium from the next round on. Solving (6) for $\pi_c$ yields

$$\pi_c = \frac{2 + 8\beta + \delta (2 - 12\beta)}{(1 - \beta\delta)(1 - \delta)}. \hspace{1cm} (7)$$

Optimally under–cutting this strategy by choosing Medium in every round yields either 11 or 4 (with probability $\beta$ and $(1 - \beta)$ respectively) in the current round and 4 in each future round, for a subjective expected discounted payoff of

$$\pi_u = 11\beta + 4(1 - \beta) + 4\delta + 4\delta^2 + ... = 7\beta + \frac{4}{1 - \delta}. \hspace{1cm} (8)$$

So, the condition for a subgame perfect equilibrium is that $\pi_c \geq \pi_u$. Combining (7) and (8), and solving for $\beta$, yields

$$\beta \geq \frac{-1 + (1 + 56\delta)^{1/2}}{14\delta} \equiv \beta^*. \hspace{1cm} (9)$$

Note that $\beta^*$ is decreasing in $\delta$: the lower the discount factor, the higher the level of trust is needed in the rival to support collusion. This result – which will be general to all of the equilibrium conditions we consider here – motivates the “discount factor” reasoning we had used in the main text.
Now consider the games used in our experiment, where there is now competition policy and perhaps reporting and leniency. The collusive strategy has to be modified slightly: “Choose High until either firm chooses Medium or Low or punishment goes into effect; then choose Medium”. Against itself, the collusive strategy earns
\[
\pi_c = \beta (1-p)(10 + \delta \pi_c) + \beta p (10 - F + 4\delta + 4\delta^2 + ...) + (1 - \beta)(2 + 4\delta + 4\delta^2 + ...),
\] (10)
while the under-cutting strategy still earns
\[
\pi_u = 7\beta + \frac{4}{1 - \delta},
\] (11)
Then the SPE condition is obtained from \(\pi_c \geq \pi_u\), using (10) and (11). In our experiment, we have \(p = 0.08\), \(F = 6\), and \(\delta = 0.8\); substituting these yields \(\beta_{SPE \ collusion}^* = 0.575\).

Next, consider the optimal response involving reporting: “Choose High in first round, report if the rival also chose High, then choose Medium forever”. Against the collusive strategy, this earns
\[
\pi_r = \beta \left(10 - F + R + \frac{4}{1 - \delta}\right) + (1 - \beta) \left(2 + \frac{4}{1 - \delta}\right) = 2 + \frac{4}{1 - \delta} + \beta(8 - F + R),
\] (12)
which for \(F \geq R\) is strictly less than \(\pi_u\), and also \(\pi_u\) for any \(\beta > 0\). So whenever collusion can be broken by reporting, it can also be broken by under-cutting – even under leniency – so that collusion is supported as long as it does not pay to under-cut (just as in the original analysis).

Now, suppose we are interested in all Nash equilibria rather than only those that are subgame perfect. Then we can modify the collusive trigger strategy to include harsher punishments: “Choose High until either firm chooses Medium or Low or punishment goes into effect; then choose Medium if collusion was detected by the competition authority, or Low otherwise”. Against itself, this new collusive strategy earns
\[
\pi_c = \beta (1-p)(10 + \delta \pi_c) + \beta p (10 - F + 4\delta + 4\delta^2 + ...) + (1 - \beta)(2 + 2\delta + 2\delta^2 + ...)
\] (13)
while the under-cutting strategy earns
\[
\pi_u = 11\beta + 4(1 - \beta) + 2\delta + 2\delta^2 + ... = 4 + 7\beta + \frac{2\delta}{1 - \delta}
\] (14)
and the reporting strategy earns
\[
\pi_u = \beta(10 - F + R) + (1 - \beta)2 + 2\delta + 2\delta^2 + ... = 2 + (8 - F + R)\beta + \frac{2\delta}{1 - \delta}.
\] (15)
As before, \(\pi_u > \pi_r\) as long as \(F \geq R\), so we only need to check that \(\pi_c \geq \pi_u\). With the parameters from our experiment, this happens when \(\beta\) is larger than \(\beta_{NE \ collusion}^* = 0.418\). Note that this threshold level of trust is lower than the one supporting collusion in subgame perfect equilibrium, illustrating another general result: collusion is easier to maintain if Nash (incredible) punishments are allowed than if only subgame perfect (credible) punishments are allowed.

Finally, consider the reporting trigger strategy “Choose High and report if possible, then choose Medium if both firms chose High in the first round, otherwise choose Low”. This strategy cannot be part of a subgame
perfect equilibrium, but might be part of a Nash equilibrium. Against itself (but switching to best responses when necessary), it earns

$$\pi_r = \beta(10 - F + R + 4\delta + 4\delta^2 + ...) + (1 - \beta)(2 + 2\delta + 2\delta^2 + ...),$$

(16)

while the optimal under-cutting strategy earns

$$\pi_u = 11\beta + 4(1 - \beta) + 2\delta + 2\delta^2 + ... = 4 + 7\beta + \frac{2\delta}{1 - \delta}$$

(17)

and colluding without reporting earns

$$\pi_c = \beta(10 - F + 4\delta + 4\delta^2 + ...) + (1 - \beta)(2 + 2\delta + 2\delta^2 + ...).$$

(18)

Note that $\pi_c < \pi_r$ if $R > 0$, so for any positive fine reduction, reporting is strictly better than non-reporting. So, we only need to determine when reporting is better than under-cutting. Substituting (16) and (17) into the inequality $\pi_r \geq \pi_u$ and simplifying yields

$$\beta \geq \beta_{NE \ reporting}^* \equiv \frac{2}{1 - \delta - F - 1 + R}$$

(19)

(note that (19) reduces to (5) in Section 2.2 as $\beta \to 1$). It is straightforward to show that $\frac{\partial \beta_{NE \ reporting}^*}{\partial R} < 0$. Thus, less trust is needed to support cartel formation and reporting as the degree of leniency increases. In particular, for the parameters used in our experiment, this equilibrium is supported when $\beta > 0.222$ under full leniency ($R = 6$) and when $\beta > 0.333$ under partial leniency ($R = 3$), while in a hypothetical analogue to our baseline treatment with reporting but no leniency ($R = 0$), the equilibrium would still be supported when $\beta > 0.667$.

B Analysis of pre-play communication

In the main text of the paper, we focus on the effects of our treatments on various measures of the implementation of anti-competitive behaviour, relating to pecuniary gains by firms at the expense of consumers. However, recall that each supergame in the experiment is preceded by a round of cheap talk for each matched pair of subjects. So, our experimental data also allow us to examine the impact of these treatments on verbal agreements to collude, as well as the relationship between any such agreements and subjects’ subsequent behaviour.27 To do this, we hired research assistants (RAs) to convert the qualitative data (sequences of free-form messages) to quantitative data, by classifying each individual “conversation” – and therefore each supergame-pair – according to what events took place during it. The events we focussed on were:

- Any non-blank message;

27We abuse terminology slightly by referring to “verbal” agreements even though all communication took place as text messages sent and received via computer.
- Both firms in a pair sending a non–blank message (within our framework, this corresponds most closely to the “mutual agreement to communicate” definition of collusion used in previous experiments by Apesteguia et al. (2007), Hinloopen and Soetevent (2008) and Bigoni et al. (2012));
- Any message about prices (irrespective of whether it was a suggestion, a factual statement, a question, etc.);
- Any suggestion to collude (a proposal of a joint pricing scheme that included a positive frequency of HH pairs, irrespective of whether and how it was responded to);
- An agreement on any pricing scheme, collusive or otherwise (a proposal made by one member of the pair, assented to by the other, without any subsequent questioning or disagreement about the scheme); and
- An agreement to collude (agreement on a pricing scheme with a non–zero frequency of HH pairs).

These variables were coded as binary (e.g., either a conversation contained a mention of prices or it did not; we were not interested in how many mentions there were after the first). Also recorded were the total number of distinct messages, and the total number of characters over all messages, for each conversation.

For the Japanese data, only one RA was available, so in the analysis below, we used the classifications directly from that RA. For the UK data, we used three RAs, so we treated an event (e.g., suggestion to collude) as having occurred if at least two of the three RAs classified it as having occurred.

Table 5 shows some descriptive statistics concerning the nature of subjects’ communication and its connection with behaviour in the subsequent supergame. Overall, just under half of the supergames were preceded by a verbal agreement to collude by the two subjects, and an additional quarter of the supergames had both of them sending messages but no verbal agreement. While reaching a verbal agreement is neither necessary nor sufficient for actual collusion, there is a strong positive association between the two. Collusion attempts and cartel formation are more likely, and collusion and high prices occur a greater fraction of the time, following a verbal agreement (case (i) in the bottom half of the table) than without one (cases (ii) and (iii)), according to a Wilcoxon signed–ranks test (session–level data, \( p \approx 0.003 \) for collusion attempts, \( p < 0.001 \) for the other three statistics); for example, collusion occurs in 76 percent of supergames after a verbal agreement to collude but only 27 percent of supergames with no such agreement. There is even evidence that cartels are more stable following a verbal agreement, with endogenous breakdowns significantly less likely (\( p \approx 0.034 \)), though there is only an insignificant decrease in under–cutting (\( p > 0.20 \)) and weakly significantly less reporting (\( p \approx 0.082 \)) after a verbal agreement. Finally, both gross and net profits are significantly higher following a verbal agreement (\( p < 0.001 \)). By contrast, comparison of cases (ii) and (iii) highlights that when there is no verbal agreement to collude, mere agreement to communicate has little systematic effect on any of the measures of collusive behaviour or breakdown.\(^{28}\)

Table 6 shows the extent to which various ways of defining collusion are correlated with each other. The first definition we use is (a) the “harm to consumers” one from the current paper: at least one HH outcome

\(^{28}\)The cases where the difference between cases (ii) and (iii) appears largest arise from small samples (e.g., the difference in reporting under partial leniency between (ii) and (iii) is based on one observation and five observations in the two samples).
Table 5: Communication outcomes – descriptive statistics

<table>
<thead>
<tr>
<th>Frequency of...</th>
<th>Part 1</th>
<th>Part 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>baseline</td>
<td>partial leniency</td>
</tr>
<tr>
<td>any non–blank message</td>
<td>.851</td>
<td>.758</td>
</tr>
<tr>
<td>both send a non–blank message</td>
<td>.753</td>
<td>.717</td>
</tr>
<tr>
<td>any message about prices</td>
<td>.740</td>
<td>.687</td>
</tr>
<tr>
<td>suggestion to collude</td>
<td>.498</td>
<td>.535</td>
</tr>
<tr>
<td>agreement on any pricing strategy</td>
<td>.495</td>
<td>.535</td>
</tr>
<tr>
<td>agreement to collude</td>
<td>.367</td>
<td>.444</td>
</tr>
</tbody>
</table>

Frequency, given (i) agreement to collude; (ii) both send a non–blank message but no agreement; (iii) neither

<table>
<thead>
<tr>
<th>(i)/(ii)/(iii)</th>
<th>(i)/(ii)/(iii)</th>
<th>(i)/(ii)/(iii)</th>
<th>(i)/(ii)/(iii)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reporting (given cartel)</td>
<td>.000/.000/.000</td>
<td>.000/.000/.000</td>
<td>.171/1.000/.400</td>
</tr>
</tbody>
</table>

during the supergame. The second definition (b) is the “agreement to communicate” one used in previous competition policy experiments: both firms choosing to send a non–blank message. To these, we add three additional definitions: (c) a verbal agreement to collude; (d) both sending a non–blank message followed by an HH outcome (i.e., (a) plus (b)); and (e) a verbal agreement followed by an HH outcome (i.e., (a) plus (c)). All of these can be justified by some combination of competition law, policy and de facto enforcement.

Table 6: Correlations between alternative definitions of collusion (Part 2 supergames)

<table>
<thead>
<tr>
<th></th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) HH outcome</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) Both send non–blank message</td>
<td>.190</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) Verbal agreement to collude</td>
<td>.442</td>
<td>.577</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d) Both send non–blank message and HH outcome</td>
<td>.849</td>
<td>.499</td>
<td>.601</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>(e) Verbal agreement to collude and HH outcome</td>
<td>.743</td>
<td>.437</td>
<td>.757</td>
<td>.875</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Although all of the correlations in the table are positive (and significant with \( p < 0.001 \)), the weakest correlation is between our definition and the “agreement to communicate” definition used in previous experiments, with both more highly correlated with definitions (c), (d) and (e) than with each other. Interestingly, the “agreement to communicate” definition is somewhat more strongly correlated with verbal agreements alone (though perhaps not surprisingly, since the former is a necessary condition for the latter), but it is much less strongly correlated with definitions (d) and (e) than our “harm to consumers” definition is, even though both are necessary
Finally, we examine the effects of our leniency programmes on verbal agreements to collude and on agreements to communicate, by estimating six additional probit models, each with the appropriate indicator as the dependent variable. Models 9–12 concern verbal agreements to collude. Model 11 uses the same set of right-hand-side variables and same sample (all Part–2 data) as in the probit for collusion (Table 4), for easy comparison. Model 12 is nearly the same, but includes some additional variables that might be associated with a verbal agreement: the total number and length of messages in the conversation, and the product of these with the Kyoto–session indicator (sentences in English and Japanese have different lengths, which might affect the number of messages needed to agree on collusion). Two more models (9 and 10) use the entire data–set but exclude the leniency–treatment variables and the collusiveness variable (which was based on outcomes in Part 1). Models 13 and 14 involve agreements to communicate (both send a non–blank message), and correspond to Models 9 and 11 for collusive agreements. (We leave out the message–number and –length variables, due to high correlation with the dependent variable.) As before, these models were estimated using Stata (version 12) with robust standard errors. The results are displayed in Table 7.

The most obvious difference between these results and those in Table 4 is that here, the leniency programmes have a positive effect on both verbal agreements to collude and on agreements to communicate – and these effects are significant in the case of partial leniency – while we saw previously that neither programme had a

### Table 7: Regression results (marginal effects at means, std. errors in parentheses), supergame–level data

<table>
<thead>
<tr>
<th></th>
<th>[9]</th>
<th>[10]</th>
<th>[11]</th>
<th>[12]</th>
<th>[13]</th>
<th>[14]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dep. variable: verbal agreement to collude</td>
<td>Both send non–blank message</td>
<td>All supergames</td>
<td>Part 2 only</td>
<td>All supergames</td>
<td>Part 2 only</td>
<td></td>
</tr>
<tr>
<td>partial leniency</td>
<td>0.029***</td>
<td>0.152***</td>
<td>0.208***</td>
<td>(0.061)</td>
<td>(0.055)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>full leniency</td>
<td>0.106***</td>
<td>0.090</td>
<td>0.073</td>
<td>(0.062)</td>
<td>(0.055)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>Kyoto session</td>
<td>0.043</td>
<td>–0.015</td>
<td>0.031</td>
<td>0.259***</td>
<td>0.043</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.061)</td>
<td>(0.049)</td>
<td>(0.094)</td>
<td>(0.032)</td>
<td>(0.036)</td>
</tr>
<tr>
<td>supergame #</td>
<td>0.073***</td>
<td>0.059***</td>
<td>–0.077**</td>
<td>–0.048</td>
<td>0.073***</td>
<td>–0.112***</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.011)</td>
<td>(0.036)</td>
<td>(0.033)</td>
<td>(0.010)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>collusiveness</td>
<td>0.452***</td>
<td>0.381***</td>
<td>0.206***</td>
<td>(0.104)</td>
<td>(0.096)</td>
<td>(0.076)</td>
</tr>
<tr>
<td></td>
<td>(0.104)</td>
<td>(0.096)</td>
<td>(0.076)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>number of messages</td>
<td>0.029***</td>
<td></td>
<td>0.012</td>
<td>(0.006)</td>
<td>(0.008)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.008)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total length of messages</td>
<td>0.000</td>
<td></td>
<td>0.003***</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>988</td>
<td>988</td>
<td>430</td>
<td>430</td>
<td>988</td>
<td>430</td>
</tr>
<tr>
<td>ln(L)</td>
<td>655.07</td>
<td>614.43</td>
<td>278.66</td>
<td>256.92</td>
<td>655.07</td>
<td>197.25</td>
</tr>
</tbody>
</table>

* (**,***): Coefficient significantly different from zero at the 10% (5%, 1%) level. Message variables also interacted with Kyoto–session dummy.
significant effect on actual collusion (indeed, the sign was actually negative). Thus, if instead of using the HH indicator as our measure of a successful collusion attempt, we had used either a verbal agreement to collude or the “agreement to communicate” of previous competition policy experiments, we would have reached a different and potentially misleading conclusion: we would have thought that our leniency programmes do in fact increase collusion, so that combined with the reduced cartel stability we also observed, we might mistakenly have concluded that the net effect of the leniency programmes was ambiguous.
C: Sample instructions (from partial-leniency treatment, UK sessions)

Instructions: first part of experiment

You are about to participate in a study of decision making. Please read these instructions carefully, as the amount of money you earn may depend on how well you understand them. If you have any questions, please feel free to ask the experimenter. We ask that you not talk with the other participants during the experiment.

These instructions are for the first part of the experiment. You will receive instructions for the second part after this part is finished. This first part is made up of several market games. You will be playing the role of a firm. At the beginning of each market game, you will be randomly matched to another participant, who also plays the role of a firm. You will be matched to the same firm for an entire market game, which will last for a number of rounds. In each round, you will choose the price of your product: High, Medium or Low. At the same time, the other firm will choose the price of its product. Your price and the other firm’s price, together, determine your profits from the market, as shown in the table.

<table>
<thead>
<tr>
<th>Your price</th>
<th>Other firm price</th>
<th>High</th>
<th>Medium</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Your profit: £10</td>
<td>Other firm profit: £10</td>
<td>Other firm profit: £11</td>
<td>Other firm profit: £0</td>
</tr>
<tr>
<td>Medium</td>
<td>Your profit: £11</td>
<td>Other firm profit: £2</td>
<td>Other firm profit: £4</td>
<td>Other firm profit: £0</td>
</tr>
<tr>
<td>Low</td>
<td>Your profit: £0</td>
<td>Other firm profit: £2</td>
<td>Other firm profit: £0</td>
<td>Other firm profit: £0</td>
</tr>
</tbody>
</table>

Chat: In the first round of each market game, whilst deciding on your price, you have the opportunity to send and receive messages with the other firm. The chat portion of the computer screen is shown below.

To write a message, make sure your cursor is active in the narrow rectangle (where “Compose messages here” appears), and type normally using the keyboard. To send, press the Enter key on your keyboard. You may send as many or as few messages as you wish. However, we ask that you NOT send messages containing:
(a) personal or identifying information about yourself;
(b) physical threats.
Sent messages will appear in the box above the narrow rectangle, on your screen and on the screen of the other firm. If the other firm sends you a message, this will also appear in the box on your screen. Other participants in the experiment will not be able to see your messages, and you will not be able to see theirs. Once either you or the other firm has chosen a price, it will be impossible to send additional messages, but you can still view messages until you’ve chosen your price or the time available for messages runs out.

**Competition law:** The government has ruled that choice by both firms of the High price is anti-competitive pricing. It has also established a competition authority to discover such behaviour. In any round in which both you and the other firm choose a High price, there is an 8% chance that the competition authority will discover this. If your anti-competitive behaviour is discovered, the punishment is:

(i) Both you and the other firm are fined £6 in the current round; this fine is subtracted from the profit you would have earned.
(ii) Both you and the other firm will have your prices restricted in all remaining rounds of this market game (only Medium and Low prices will be available).

If your anti-competitive behaviour is not discovered, or if either you or the other firm (or both) does not choose the High price, there is no punishment.

**Continuing or ending the market game:** At the end of each round, there is a 20% chance that the market game will end, and an 80% chance that it continues for at least another round. If the game ends, a new market game will begin, and you will again be randomly matched to another participant. Also, punishments for anti-competitive behaviour disappear when a new market game begins. If the market game continues, you play another round, matched to the same other firm.

**Sequence of play in a round:** The sequence of play in a round is as follows.

1. Sellers choose their prices. You will be able to send and receive messages at this time, if it is the first round of a market game.
2. You are informed of the other firm’s price.
3. You are informed of your profit for the round, including any fine incurred. Also, if you and the other firm chose High prices, you are told whether your anti-competitive behaviour was discovered.
4. The computer randomly determines whether the game will continue or end.

**Payments:** At the end of the experimental session, you will be paid based on your results. The computer will randomly select one round from each market game – in this part and the next part. You will be paid the total of your scores in those rounds, translated into real money at an exchange rate of 2 “lab-pounds” = £1. Payments are made privately and in cash.
Instructions: second part of experiment

The procedure in this part of the experiment is very similar to that in the first part. You will play several market games, each for a variable number of rounds. The participant matched with you will still be chosen randomly at the beginning of a market game, and remain the same until the market game has ended.

Leniency programme: The difference from the first part of the experiment is that the competition authority has instituted a leniency programme. In any round where both you and the other firm have chosen the High price, you are each given the opportunity to report your anti-competitive behaviour to the competition authority. If either you or the other firm does report, then the competition authority will (100% of the time) discover your anti-competitive behaviour, and both firms’ prices will be restricted (to Medium or Low) in all remaining rounds of this market game. However, a firm that reports incurs a reduced fine of £3:
- If you report, then you will incur the reduced fine for anti-competitive behaviour.
- If the other firm reports but you do not, then you will incur the same fine as in the first part of the experiment, while the other firm incurs the reduced fine.
- If both firms report, then both incur the reduced fine.

Sequence of play in a round: The new sequence of play in a round is as follows.
(1) Sellers choose their prices. You will be able to send and receive messages at this time, if it is the first round of a market game.
(2) You are informed of the other firm’s price. If you and the other firm chose High prices, you choose whether to report your anti-competitive behaviour to the competition authority.
(3) You are informed of your profit for the round, including any fine incurred. Also, if you and the other firm chose High prices, you are told whether your anti-competitive behaviour was discovered.
(4) The computer randomly determines whether the game will continue or end.