



**Do Sanctions or Moral Costs Prevent the Formation
of Cartel-Type Agreements?**

Béatrice BOULU-RESHEF, Constance MONNIER-SCHLUMBERGER



Mise en ligne / Online : 11/04/2023

Remplace une version précédente du même numéro / Replaces a previous version of the same DR LEO number

Cette version remplace celle du / This replaces the previous version dated : XX/XX/XXXX

Titre précédent / A previous version was titled : "Remplacer par l'ancien titre"

A paraître dans / Forthcoming in : Nom du journal, vol, numéro, pages

Paru dans / Published in : Nom du journal, vol, numéro, pages

Do Sanctions or Moral Costs Prevent the Formation of Cartel-Type Agreements?

Béatrice Boulu-Reshef¹, Constance Monnier-Schlumberger²

Cartel decisions are made by managers. These decisions result in an increase in the company earnings at the expense of the earnings of a third party, typically consumers or business partners. Using laboratory experiments, this article studies the behavioral foundations of individuals' willingness to engage in such cartel-type agreements in a context in which engaging in such agreements increases the cartelists' earnings but reduces the earnings of a third participant. This reduction in earnings is a loss for the third participant but it can also be seen as a moral cost for the cartelists. The impact of sanctions – monetary, leniency, compliance, and exclusion – is investigated in a within-subject design and the impact of moral cost is tested in a between-subject design and at various levels of moral costs, which also allows us to test the robustness of the effects of sanction schemes in these different moral cost conditions. The results show that compliance and exclusion sanction schemes are vastly more effective deterrents than monetary and leniency sanction schemes. These results are robust to varying levels of moral costs, but also to varying levels of probabilities of detection and fines. Although statistically significant, the moral cost has an impact but its magnitude is limited, showing that participants are mainly sensitive to the monetary gains associated with cartel-type agreements and less sensitive to the reduction in earnings imposed on the third participant. Finally, being a woman and being risk-averse is associated with a lower propensity to engage in such agreements.

Key words Antitrust, cartel, fraud, moral behavior, sanctions.

JEL code C91 K21 K42 L41 M12

¹ Béatrice Boulu-Reshef, Laboratoire d'Economie d'Orléans, Université d'Orléans, Rue de Blois, 45067 Orléans, +33 (0)2 38 41 70 37, beatrice.boulu-reshef@univ-orleans.fr, <https://orcid.org/0000-0001-6265-2926>.

² Constance Monnier-Schlumberger, Prism-Sorbonne, Université Paris 1 Panthéon-Sorbonne, [17, rue de la Sorbonne 75005 Paris](https://www.univ-paris1.fr/), +33 (0)1 40 46 31 70, constance.monnier@univ-paris1.fr, <https://orcid.org/0000-0003-0831-4883>.

1. Introduction

Cartels are formed when individuals, typically managers, jointly decide to engage in anti-competitive price-fixing arrangements. In most cases, as for the recent truck cartels condemned by the European commission in 2021 (European Commission, 2021), these practices entail an agreement to charge higher prices than those that would have been set by a more competitive marketplace and these higher prices lead to the reduction of earnings of a third party, for instance consumers in Business-to-Consumers settings, or business partners in Business-to-Business settings. These decisions have been initially studied within the utilitarian framework (Becker 1968), which argues that a firm is more likely to engage in illicit practices when the gains derived from these practices exceed their costs, as, in its simplest form, the expected cost associated to cartel-type decisions is its sanction multiplied by the associated probability of detection. Nonetheless, cartel determinants are complex and they have been the objects of study.

First, further developments on the role of sanctions have highlighted that the difficulty to disincentive cartel-type practices is influenced by the fact that the interests of managers may not be aligned with those of shareholders (Jensen and Meckling 1976). Recent works on remuneration schemes show that they can generate perverse economic incentives, as well as psychological and social responses that motivate counterproductive behaviors such as fraud (Larkin and Pierce 2016) and anticompetitive practices (Fleckinger *et al.* 2013; Fonseca *et al.* 2022). Second, in addition to the role of incentives and sanctions in cartel-type decisions, previous works have investigated the role of psychological and behavioral factors in explaining such decisions. For instance, Parasuram *et al.* (2017) show that rewards can lead to more cheating when rewards are large enough. However, the relationship between reward and dishonest behavior is complex and individuals may display internal costs of illicit behavior or dishonesty that may be compensated by financial reward and as shown by Gino and Pierce (2009), financial self-interest cannot fully explain individuals' tendency to engage in dishonest behavior. In the context of cartel-type behavior, researches in law and economics finds that individual managers' characteristics play a role in cartel-type decisions, as surveyed by Combe and Monnier (2016), suggesting that psychological or other individual specific features may play a role in determining such decisions. Cartel decisions are thus determined by many factors, which complicates the analysis of the effectiveness of sanctions against them. In that perspective, laboratory experiments can highlight the most effective measures to deter these illegal practices. For instance, recent experimental literature shows that sanction schemes have effect in deterring collusive behavior (see for instance Bigoni *et al.* 2015, Chowdhury *et al.* 2018). But they only test basic sanctions. Thus, our article aims at completing this literature by first testing more sophisticated and recent sanction schemes, and second including the

effect of a moral cost (in the form of a loss inflicted on others) on the propensity of individuals to engage in cartel practices.

Using laboratory experiments, our paper combines two strands of literature by studying financial and behavioral determinants of cartel-type agreement decisions in a context in which individuals decide to engage in cartel-type decisions that, if jointly agreed, increases their earnings at the expense the earnings of a third party. Such cartel-type agreements increase the cartelists' earnings but reduce the earnings of a third party. This reduction in earnings for the third party can be seen as a moral cost for the cartelists and as a loss for the third party. In Business-to-Consumers settings, third parties are typically consumers. In Business-to-Business settings, they are typically firms. In addition, the sanctions mechanisms are diversified and enriched, including recent tools developed by competition policy. In the context of the experiment, participants are matched in groups of three and the reduction in the third player's earnings when cartel-type agreements are chosen by the two cartelists with whom the third player is matched. The objective of this investigation is to experimentally test the propensity of two individuals to form a cartel with the intent of increasing their earnings (if not detected) at the expense of another individual, while simultaneously, facing various sanctions aimed at deterring such behavior, and particularly recent ones such as exclusion and compliance. Thus, the experiment does not only rely on participants' decision to collude and subsequently face potential penalties, but on participants' decision to form a cartel-type agreement which imposes a real reduction in earnings of a third participant who is a person present in the laboratory. Thus, this paper fills a double gap in the literature as the question of the impact of the moral cost of cartelists remains, to the best of our knowledge, unstudied (if we exclude tacit collusion experiments which do not imply agreements); and as innovative sanctions schemes have not either been tested yet.

Hence, the paper addresses two main research questions. First, the paper studies the impact of financial motivations on the propensity to impact of various types of sanctions on the decision to enter into a cartel-type arrangement. Previous research has explored the role of sanctions on cartel formation through such means as monetary fines and whistleblowing. While these previously studied types of anti-cartel sanctions are prominent in pro-competitive policies, other types of sanctions, such as compliance programs and exclusion schemes, which rely mainly on disqualification or dismissal from one's employment, have emerged. To cover this extended range of sanctions, this paper examines in a controlled environment the impact of four types of anti-cartel sanctions on cartel-type decision agreements. In the experiment, monetary sanctions are implemented via a monetary fine in case of a detection; leniency is implemented via the option to report an agreement after agreeing on its formation; compliance is implemented via the removal of the agreement possibility in the period following a detection; and exclusion is implemented via the removal of both decisions options in the period following a detection. Although monetary sanctions and leniency have been studied in previous

experimental works – reviewed in subsection 2.1., the temporary removal of one decision as a form of forced compliance and the removal of both options as a form of exclusion or disqualification, are two new experimental manipulations. Second, in addition to sanction types, the paper investigates the impact of the size of the reduction in earnings imposed on a third party by the cartel-type agreement. In particular, the size of the reduction in earnings for the third party is varied which allows us to check the robustness of the results. Thus, the paper examines the combined influence of sanction schemes on cartel decision choices, as well as the impact of the size of the loss imposed on a third party. This paper uses a between-subject design to estimate the impact of sanctions schemes as well as the impact of the size of the loss. The robustness of these effects is investigated with the use of varying fine levels and probabilities of detection. Additionally, we control for two important individual level characteristics which are gender and risk aversion.

The paper starts with the experimental design. Section 3 reports the related literature and hypotheses. The results are listed in Section 4. Section 5 concludes.

2. Experimental methodology

Participants are matched randomly in groups of three individuals composed of two Type 1 participants and one Type 2 participant. Participants remained in the same group for the whole experiment and the roles are fixed for the whole experiment. Type 1 participants have to choose between two possible choices: Choice A and Choice B. Type 1 participants make their choices simultaneously and without communication. Choice A grants each Type 1 participant 10 experimental currency units (hereafter ECUs). Choice B grants each Type 1 participant 20 ECUs if both Type 1 participants choose Choice B. If only one Type 1 participant chooses Choice B and the other chooses Choice A, then Choice A is implemented for both participants. Choice A corresponds to a competitive strategy that leads to lower earnings compared to the cartel-type agreement and, as such, it is exposed to sanction schemes. Choice B corresponds to the choice of a cartel-type agreement that leads to higher earnings compared to a competitive strategy but it is exposed to sanction schemes. A cartel-type agreement is formed if both Type 1 participants choose Choice B. If Choice A is implemented, Type 2 participants each earn 10 ECUs. If Choice B is implemented, Type 2 participants earn $(10 - \text{the "loss" i.e. the "moral cost"}$, a loss that is imposed on the Type 2 participant). For simplicity, when referring to the reduction in earnings for the Type 2 participant, we will use the term “loss” and when referring to the psychological cost that is imposed on Type 1 participants in conditions in which a cartel-type agreement, we will use the term “moral cost”.

Using a between-subject design, four levels of losses, which are reduction in earnings imposed on the Type 2 participants due to the cartel-type agreements, are tested in separate experimental sessions.

The four levels of losses are as follows: A loss of 0 leading to earnings of $10 - 0 = 10$ ECUs for Type 2 participants (i.e. a baseline with “no loss”, i.e. “no moral cost” for Type 1 participants), a loss of 3 leading to earnings of $10 - 3 = 7$ ECUs for Type 2 participants (i.e. a “low loss” environment, i.e. “low moral cost” for Type 1 participants), an loss of 7 leading to earnings of $10 - 7 = 3$ ECUs for Type 2 participants (i.e. a “high loss” environment, i.e. “high moral cost” for Type 1 participants) and a loss of 10 leading to earnings of $10 - 10 = 0$ ECU for Type 2 participants (i.e. a “full loss” environment, i.e. “full moral cost” for Type 1 participants). Type 2 participants are informed of the decisions that Type 1 participants have made. In the experiment, Type 2 participants do not make any decisions but they answer a survey on their attitudes toward competition policies and firms’ practices, and they are informed of their earnings.

An agreement may only be detected during the period in which the agreement is formed. If a cartel-type agreement is formed and is not detected, then no sanction is implemented. An agreement may be sanctioned if detected. In the experiment, decisions are made in four sanction scheme environments which are implemented in a within-subject design. Thus, all participants face all sanction schemes. These sanction schemes correspond to the four main sanction schemes that are used in antitrust policies.

- The *Monetary sanction scheme* (labeled “Basic” in the instructions) corresponds to simple monetary fines. In the experiment, in case a cartel agreement is formed and in case it is detected, a monetary fine is implemented.
- The *Compliance sanction scheme* (labelled “Basic + Compliance” in the instructions) allows to test the impact of being under threat of greater surveillance and thus lowering the option to form cartel-type agreements on the propensity to engage in cartel-type practices. In the experiment, in case a cartel agreement is formed and in case it is detected, a monetary fine is implemented and, in addition, Choice B is no longer available in the period following the detection. In that period following the detection, the earnings of Type 1 participants are 10 ECUs and the earnings of Type 2 participants are 10 ECUs.
- The *Exclusion sanction scheme* (labeled “Basic + Exclusion” in the instructions) allows to measure the impact of being under threat of temporary incapacitation or dismissal on the propensity to engage in cartel-type practices. In the experiment, in case a cartel agreement is formed and in case it is detected, a monetary fine is implemented and, in addition, the ability to take a choice is no longer available in the period following the detection. In that period following the detection, the earnings of Type 1 participants are 0 ECU and the earnings of Type 2 participants are 10 ECUs.
- The *Leniency sanction scheme* (labelled “Basic + Leniency” in the instructions) aims at capturing the essence of leniency programs organized by competition authorities to promote

denunciation. In the experiment, in case a cartel agreement is formed, Type 1 participants are given the possibility to denounce the cartel-type agreement.

- If no Type 1 participant denounces the cartel-type agreement and in case it is detected, a monetary fine is implemented.
- If only one Type 1 participant denounces the agreement, this Type 1 participant does not pay his/her monetary fine, but the other Type 1 participant does.
- If both Type 1 participants denounce the agreement, the fastest participant to denounce the agreement gets his/her fine cancelled. The other Type 1 participant, on the other hand, must pay his/her fine, but not the other Type 1 participant's fine.

All decisions are made in two probability regimes: in a low/high probability environment (1/10 or 9/10) and in a medium low/medium high probability environment (1/3 or 2/3). The probability regimes are presented in reverse order to control order effects.

In each probability regimes, if detected, there are 3 possible types of fines. There is a low fine of 5 (leading to earnings for Type 1 participants of $20 - 5 = 15$), a medium fine of 10 (leading to earnings for Type 1 participants of $20 - 10 = 10$) and a high fine of 15 (leading to earnings for Type 1 participants of $20 - 15 = 5$, resulting in a loss compared to the earnings resulting from Choice A).

There are a total of 2 probabilities in each probability regimes * 2 probability regimes * 3 fines * 4 sanction schemes = 48 decision environments.

In addition, in order to test the robustness of the results by controlling for the impact of repetition as well as controlling for the impact of knowing versus not knowing the environment that will be dismissed in the Compliance and Exclusion sanction schemes, each environment is repeated once. In these two schemes, when making the decision the first time around in each environment (labeled 'Repetition 1/2' in the experimental interface), they know the environment that may be dismissed. When making the decision the second time around (labeled 'Repetition 2/2' in the experimental interface), they do not know the environment in which possible sanctions of the Compliance and Exclusion type will take place. This doubling of the decision allows to check the robustness of the result to the varying level of information that is implied by the fact that participants have more information the first time around than in that same decision environment the second time around. Therefore, in total, they make $48 * 2 = 96$ decisions.

The design allows to pull in one experiment using a within-subject design the four sanction schemes. By doing so, it combines the monetary scheme and the leniency sanction scheme that had been studied in previous works separately, as detailed in Section 4. Moreover, it adds the exclusion and

compliance schemes to the study of cartel-type agreements. In addition, the impact of the level of loss is studied using a between-subject design, resulting in four cohorts of participants who face the same sanction schemes but at different levels of losses.

The experiments were carried out at the LEEP (<http://leep.univ-paris1.fr/accueil.htm>), the experimental economics laboratory of the University Paris 1 Panthéon-Sorbonne and the Paris School of Economics. Participants were drawn from the database of volunteer participants. Their participation fee was of 6 euros and their additional earnings varied according to their decisions and the decisions of the participants in their group. In total, they earned on average 16.78 euros (st. dev. = 1.34). Sessions lasted about 1 hour and 15 minutes. In total 288 individuals participated in the experiments, 192 participants held Type 1 participant positions and 96 participants held Type 2 participant positions. This results in 192 participants, 96 decisions each, leading to 18,432 observations in total, and as there are four levels of losses, each level of loss is documented by 4,608 observations.

3. Related literatures and hypotheses

This article analyzes the efficacy of sanction schemes in experimental settings while accounting for the effect of moral cost on cartel-type agreement decisions and agreement formation and while controlling for individual traits. This section reviews the literatures related to the experimental design and lays out the hypotheses that are tested.

2.1 Game structure

Experimental studies have been carried out to better understand collusive mechanisms and the effect of competition policies. This literature, which was started by Plott (1982) and Holt (1995), is surveyed in Davis and Holt (1998) and Normann and Müller (2015). Previous researches study the effect of communication, price announcements and leniency on the formation and stability of cartels and price levels. For instance, Fonseca and Normann (2012) find that communication facilitates cartel formation in experimental markets (see also Andres *et al.* 2023).

Most experiments on cartels have been conducted with Bertrand games and have been focusing on collusion via the level of price generated after non-bidding agreements. Apesteguia *et al.* (2007) use a one-shot homogeneous goods Bertrand game between three firms, with the possibility to communicate before the price choice and to be detected and punished by an antitrust authority. Hinloopen and Soetevent (2008) implement a repeated version of Apesteguia *et al.* (2007) game. Bigoni *et al.* (2012, 2015) use a repeated differentiated-goods Bertrand duopoly game. In their article, participants can decide, before choosing prices, whether to form a cartel by communicating on prices. Chowdhury *et al.*

(2018) use an experimental market with three firms competing in a repeated Bertrand game with inelastic demand and constant marginal cost in which participants making the firms' decisions can form via communication a non-binding price cartel agreement. Chowdhury *et al.* (2020) rely on a setting similar, but after an initial phase of communication, the ability to agree on price-fixing is removed and participants are only able to tacitly collude. Fonseca *et al.* (2022) rely on an infinitely repeated Bertrand oligopoly game and employ a laboratory experiment in which cartels are operationalized through the ability of managers to verbally communicate. Andres *et al.* (2021) use a discrete three-player Bertrand price-setting game and introduce an experimental design that includes the option to communicate without automatically running the risk of being convicted for forming a cartel. In their experiment, firms interact repeatedly and a chat window opens at the beginning of each round. Each market includes a participant in the role of the competition authority as an active player. They find that leniency does not affect cartelization.

Hamaguchi *et al.* (2009) do not rely on a Bertrand game. In their study, participants are not given a choice on whether or not they want to participate in a cartel, but they are offered possibilities to denounce it or not. Thus, this article studies the functioning of leniency programs after participants are placed in a position in which they have formed cartels. As in Hamaguchi *et al.* (2009), our experiment relies also on a limited market environment. In our experimental setting, participants must decide whether or not to form a cartel-type agreement. Hence, our paper does not document the level of collusion (measured via the price increase) but it studies the impact of the conditions participants are in on their propensity to form a cartel-type agreement. As the effect of sanctions schemes and moral costs on cartel formation is complex, the experimental setting used here allows studying variations of sanctions schemes and moral costs that cannot be cleanly introduced in a market game as the returns from a cartel-type agreement have to be held constant in the experimental design, which is the case in our design.

2.2 Matching procedure

Matching procedures in the above-cited articles rely on fixed matching. This design best captures the features of actual markets in which firms interact with each other repeatedly in the same market. Hamaguchi *et al.* (2009) and Apesteguia *et al.* (2007) rely on matching that is fixed throughout the experiment. But in some articles, the fixed is altered in the course of the experiment. Chowdhury *et al.* (2020) allocate participants who interacted with the same two other participants throughout the experiment, apart from a treatment in which participants at some pre-announced point in time are rematched into new groups. In Bigoni *et al.* (2012, 2015), there is a rematching procedure in that at the end of each period, participants were rematched with the same competitor with a probability of 85%. This kind of re-matching treatment allows accounting for sanctions such as debarment or imprisonment

which shortens the expected length of interaction between managers. Nevertheless, these individual sanctions are usually for a short period, and most of the time, managers still work in the same markets or even firms (Stephan, 2011). To match and contribute to the preexisting literature, our experiment relies on fixed matching.

2.2 The influence of sanctions schemes on collusion and cartels

The effects of sanctions and detection probabilities on collusion have been studied by Bigoni *et al.* (2012), Bigoni *et al.* (2015), Chowdhury *et al.* (2018), Chowdhury *et al.* (2020) and Fonseca *et al.* (2022). Bigoni *et al.* (2012) demonstrate that fines have a significant deterrent effect, but also have a pro-collusive effect because the presence of sanctions schemes is associated with an increase in cartel prices. Treatments differ in the presence of a cartel prohibition with positive expected fines for infringers, and in the possibility of obtaining leniency or a reward by self-reporting before an investigation is opened. They show that antitrust without leniency reduces cartel formation but increases cartel prices. With rewards, prices fall at the competitive level. Overall, their results suggest a strong cartel deterrence potential for well-run leniency and reward schemes. Bigoni *et al.* (2015) ran several treatments differing in the probability of detection the level of the fine, and the possibility to report. Participants can self-report both before and after price choices become public information, as in reality. They test two levels of fine and probability of detection and show that the size of the fine in itself plays a role in deterring cartels, regardless of the probability of detection. They show that leniency policies offering immunity to the first reporting party and a high fine is the main determinant of deterrence, having a strong effect even when the probability of detection is zero. Absent leniency, the probability of detection and the expected fine matter more, and low fines are exploited to punish defections. Chowdhury *et al.* (2018) conduct an experiment involving two combinations of fines and detection probabilities. They find that in the absence of leniency, the probability of detection and fines are substitutable. Chowdhury *et al.* (2020) implement a fine treatment in which participants face an exogenous detection probability of 16%. Detection is possible either in the period in which the agreement is formed or in subsequent periods provided that it has not been detected before. Fonseca *et al.* (2022) explore the impacts of different antitrust regimes on managers' labor contracts and they compare legal regimes in which firms are fined with legal regimes in which managers are prosecuted. Their results suggest that there is less collusion in legal regimes in which managers are prosecuted.

Experimental results on leniency further show that such whistleblowing settings help deter collusive practices. Chowdhury *et al.* (2020) also implement a fine reduction procedure for leniency applications. They show that rematching in the experiment is found to be an effective mechanism to prevent post-cartel tacit collusion as well as to reduce cartel stability. Hence, provided that debarment programs have similar disruptive indirect enforcement effects on collusion in the field as in the lab, these policy tools

may help to minimize the harm caused by post-cartel tacit collusion. Apesteguia *et al.* (2007) test a fine equal to 10% of the revenue with positive probability. They also study an implementation of leniency in which self-reporting firms receive a fine reduction; and bonus, in which they are rewarded with a share of the fines paid by other firms. In the presence of a leniency sanction scheme, they find that prices are significantly lower but that the rate of cartel formation remains high. Hinloopen and Soetevent (2008) open up the possibility for participants to “race to report” by giving early applicants larger fine reductions. In each period they face a probability of 15% of being detected by an antitrust authority and a fine equals 10% of their revenue. They obtain that leniency reduces the rate of cartel formation. In Hamaguchi *et al.* (2009) players who reported the information can receive a partial exemption, full exemption or reward. They use a treatment in which leniency is given to all deviators and a treatment in which leniency is given to only the first deviator. The reward of whistleblowers has been found to help deter collusive practices. Their results also show that the larger the number of cartel members in a group, the weaker their ability to maintain the collusion is.

Concerning sanction schemes, our experiment contributes to the literature in the following way. First, it focuses on individual sanctions as participants’ earnings are directly impacted by their collusive behavior in all our sanctions schemes. Moreover, it includes a greater variability of fines and detection probabilities which allows to cleanly assess the role of fines and probability of detection. Second, besides leniency, the experiment adds two sanctions schemes which have never been evaluated in laboratory for cartel-type agreements: compliance – in its simplest form, i.e. disqualification – and exclusion. In particular, our experimental setting aims at testing if stronger sanctions, such as compliance and even more so exclusion, lower the propensity to choose the cartel-type agreement option. To the best of our knowledge, there is no experimental evidence on the impact of compliance in cartel-type settings in the laboratory. However, evidence on the effect of exclusion with reintegration – which resembles our implementation – has been collected in other games. For instance, Soldà and Villeval (2020) find in a taking game that a long exclusion has a high disciplining effect on cooperation after reintegration, but only when the length of exclusion is not chosen by group members. Concerning the treatment with exclusion, previous experiment using exclusion over one period and find that the presence of exclusion favors the choice of cooperative behavior (Masclot 2003, Kerr *et al.* 2009, Koike *et al.* 2018). In line with the pre-existing experimental literature, we hypothesize that the effects of the two additional sanction schemes – compliance and exclusion – are associated with a lower propensity to choose the cartel-type agreement option. In line with the literature, we hypothesize:

H1 *The stronger the sanction schemes, the lower participants’ propensity to choose the cartel-type agreement option. The propensity to choose the cartel-type agreement is highest in the monetary sanction scheme, followed by leniency, compliance and finally, exclusion.*

2.3 The role of moral costs in cartel-type decisions and formation

Previous research has shown that market framing may increase the likelihood of unethical behavior; for instance, the role of competition in markets in creating strong incentives for unethical practices has been established (Shleifer 2004, Bennett *et al.* 2013). Using experimental methods, Falk and Szech (2013) find that markets can erode moral values. Sutter *et al.* (2020) further explain that market framing increases the acceptability of unethical behavior and that it is associated with a dilution of responsibility. In the case of cartel decisions, Stucke (2011) stresses that the fact that damages inflicted on third parties are not always visible to the managers involved in cartel-type decisions may undermine their sense of responsibility. To the best of our knowledge, the role of a moral cost in the form of a reduction in earnings imposed on a third party in the context of cartel-type decisions has not been investigated explicitly. However, experimental and behavioral economists have long studied the conditions under which participants engage in reducing other participants' earnings if it comes at the expense of their earnings. The seminal "money burning" experiment of Zizzo and Oswald (2002) shows that participants are willing to pay to reduce other participants' earnings. Abbink *et al.* (2002) report a bribery experiment in which a treatment with a "negative externality" is tested; the negative externality is a monetary damage imposed on all other participants in the session except the two participants interacting in a pair and they find no effect of the presence of the externality compared to the baseline with no externality. Murphy *et al.* (2011) propose a "social value orientation" scale documents that some participants are willing to reduce others' earnings in order to increase their own earnings. In an experimental framework, Kirchler *et al.* (2016) find that the threat of monetary punishment and the removal of anonymity, making participants identifiable, promote moral behavior in settings in which participants can generate a negative externality imposed on third parties. These results are explained by how individuals perceive responsibility for their behavior. Sutter *et al.* (2020) experimentally examine negative externalities by looking at voided donations for a potentially life-saving vaccine provided by UNICEF. The authors find lower trading volume in markets with a negative externality; however, there is some trading, suggesting that at least to some participants, the presence of a negative externality does not prevent them from trading.

Finally, concerning moral costs, cartel-type settings have the particularity to place cartelists in a situation in which they have to arbitrage between loyalty and fairness; loyalty towards their co-cartelists who benefit from the cartel and fairness towards third parties who suffer from the cartel. This loyalty-fairness dilemma has been described in the context of whistleblowing (Elliston 1982, Jubb 1999). Concerning the intention to blow the whistle, Waytz *et al.* (2013) validate the hypothesis that loyalty towards co-cartelist may be associated with a lower propensity to reject cartel-type agreements as they find that individuals who value fairness are more likely to report wrongdoings than individuals who value loyalty.

Previous studies show that the reduction in earnings imposed on a participant may contribute to reducing the propensity to choose the cartel-type agreement option, but the loyalty-fairness dilemma could mitigate this effect. This leads to the following hypothesis:

H2 *The presence of a moral cost in the form of a reduction of earnings imposed on another participant lowers slightly the propensity to choose the cartel-type agreement option.*

2.2 The role of individual traits in cartel decisions and formation

The literature on business ethics examines the role of individuals' traits as a factor of corporate fraud. Zahra *et al.* (2005) developed a framework in which individual variables (age, experience, education, gender and self-control) influence the likelihood of corporate fraud. Concerning the corporate fraud literature, it appears that risk aversion or risk appetite is likely to play a role in the choice of individuals. Tan and Yim (2014) show that being more risk-averse decreases the propensity for tax evasion. Bernhardt and Rastad (2016) find that cartelists with high risk will collude less. Empirical papers also focus on managers' personality traits. Cohen *et al.* (2011) analyze evidence from press articles covering corporate fraud cases and show that the personality traits of managers appear to be a major fraud-risk factor. The propensity to engage in corruption has been attributed to individual predispositions, including lack of integrity (Frost and Rafilson, 1989), moral identity (Reed and Aquino, 2003), self-control (Marcus and Schuler, 2004) and empathy (Eisenberg, 2000). Studies have depicted the CEO ethical profile as a determinant of managerial fraud (Carson 2003, Chen 2010). Zona *et al.* (2013) show that a complex set of personal/ psychological attitudes of a CEO affect ethical and legal behaviors. Rijsenbilt and Commandeur (2012) find a positive relationship between proxies for CEO narcissism and fraud. Other studies focus on gender and find that women have a greater aversion to unethical behavior than men (Robinson *et al.* 2000) and that their reasoning is more contextual than men, which includes ethical matter (Van Staveren 2014). Ramdani and Van Witteloostuijn (2012) show that bribery is more likely to occur when the principal owner is male rather than female. Rivas (2013) also finds in an experimental study that women are less corrupt than men. The experimental study by Hamaguchi *et al.* (2009) conclude that women are more likely than men to end a cartel by denouncing them through a leniency program.

Our experiments contribute to these literatures by documenting the role of two individual traits, i.e. risk attitude and gender. In line with the above-mentioned literature, we hypothesize:

H3a *Risk-aversion lowers the propensity to choose the cartel-type agreement option.*

H3b *Women are less likely to choose the cartel-type agreement option than men.*

The following section reports the results.

4. Results

First, we report the effect of sanctions schemes, fines and the probability of detection. Second, we document the impact of the size of the loss imposed on Type 2 participants. Third, the effects of gender and risk aversion are described. The results focus mainly on two output measures, the cartel-type agreement choice, i.e. Choice B, and cartel-type agreement formation, which is implemented if both Type 1 participants simultaneously choose Choice B.

4.1. The effect of sanctions schemes, fines and detection probability

Sanction schemes have strong and persistent effects on Choice B and cartel-type agreements. Table 1 and Figure 1 report the effectiveness of the four sanction schemes implemented in the experiment. The monetary sanction scheme and the leniency sanction scheme are not effective in deterring the selection of Choice B, as 61.96% and 65.45% of choices are Choice B, in the monetary and the leniency sanction scheme, respectively. Compliance is more effective as this proportion drops to 51.54%. Exclusion is the most effective sanction scheme as this proportion drops to 42.30%. The proportion of cartel-type formation in each sanction scheme is coherent with these behavioral results: it is lowest in the exclusion sanction scheme with 27.82%, followed by the compliance sanction scheme with 39.67% and by the monetary sanction scheme and the leniency sanction scheme with 47.31% and 50%, respectively. These differences are statistically significant (see Appendix Table A1, Panel a.).

Table 1 – Choice B and cartel-type agreements by sanctions scheme

Treatment	Baseline	Compliance	Exclusion	Leniency	Total
<i>Choices</i>					
Number of Choice A	1,753	2,233	2,659	1,592	8,237
Number of Choice B	2,855	2,375	1,949	3,016	10,195
Number of observations	4,608	4,608	4,608	4,608	18,432
Proportion of Choice A	.38	.48	.58	.35	.45
Proportion of Choice B	.62	.52	.42	.65	.55
<i>Resulting choices implemented</i>					
Competitive choice implemented	2,428	2,780	3,326	2,304	10,838

Cartel-type agreement implemented	2,180	1,828	1,282	2,304	7,594
Observations	4,608	4,608	4,608	4,608	18,432
Proportion competitive choice	.53	.60	.72	.50	.59
Proportion of cartel-type agreements	.47	.40	.28	.50	.41

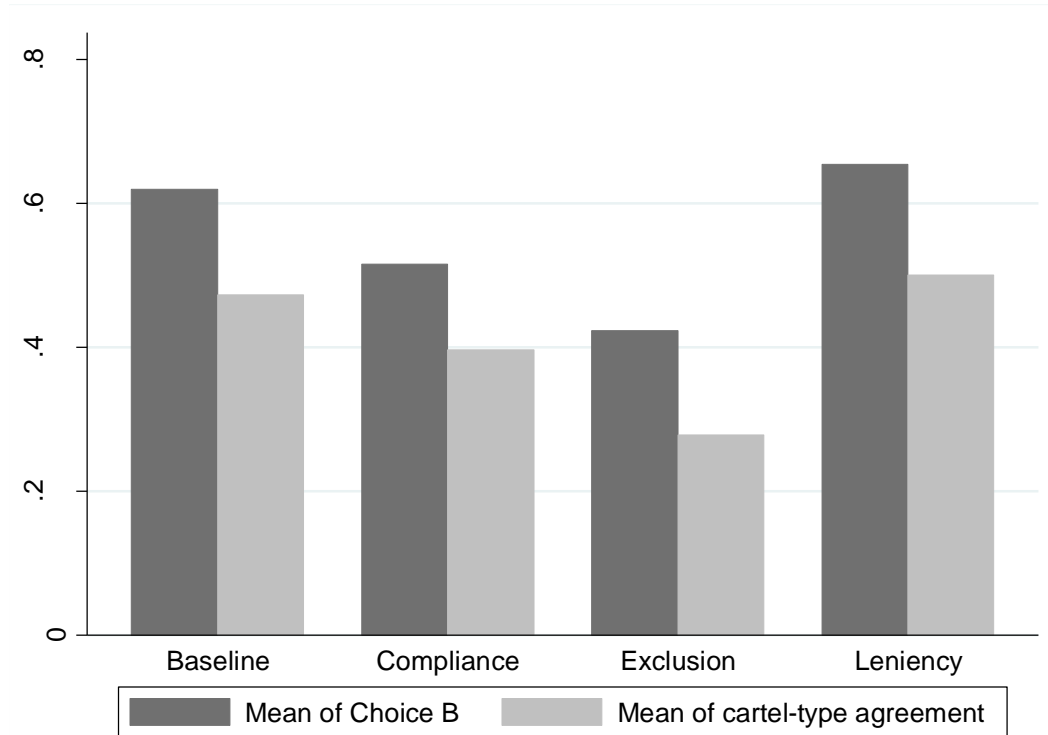


Figure 1 – Proportions of Choice B and cartel-type agreements by sanctions scheme

Result 1 (sanctions schemes). *Compliance and exclusion sanctions schemes are very effective at deterring cartel-type agreements. Monetary and leniency sanction schemes are less effective at deterring cartel-type agreements. The leniency sanction scheme leads to an increase in the rate of detection.*

Note that in the compliance and the exclusion sanction schemes, the decrease in Choice B and in cartel-type agreements can in part be explained by a mechanical effect: since Choice B is removed in the period following a detection, participants have fewer instances in which they may select Choice B. In the exclusion sanction scheme, Choice B was selected 45.12% of the time when Choice B was available. The resulting cartel-type agreement rate is 29.68%. The decrease in Choice B and in cartel-type agreement remains statistically different from the monetary sanction scheme (with a p -value = 0.0000 in both cases). In the compliance sanction scheme, Choice B was selected 59.12% of the time when Choice B was available. The cartel-type agreement rate is then of 45.57%. The decrease in Choice B remains statistically different from the monetary scheme (with p -value = 0.0075). However, the decrease in cartel-type agreements is not statistically different from the monetary sanction scheme. The

effectiveness of the compliance sanction scheme is thus more a matter of *ex post* monitoring than *ex ante* deterrence, unlike exclusion which is efficient *ex ante*. The leniency sanction scheme is ineffective in deterring cartels. However, it is found to be effective in terms of detection; under leniency, the rate of detection is 89.58 %.

Fines impact behaviors in a consistent fashion as higher fines are associated with lower propensities in both output measures. Table 2 reports the results. The deterrent effect of fines is greater when it is increased from 10 to 15 (15 being the maximum fine level) than when it increases from 5 to 10. These differences in Choice B and cartel-type agreements across fines are statistically significant (see Table A1, panel a.). In the monetary sanction scheme, with a low fine (of 5), 74% of the participants choose B and the proportion of cartel-type agreements amounts to 61%. With a high fine (of 15), these proportions are 44% and 29% respectively. These rates are lowest with exclusion and with a high fine, respectively at 31% and 17%. Overall, by multiplying the amount of the fine by three from 5 to 15, the proportion of cartel-type agreements is reduced by half. The deterrence effect is highest when the fine exceeds the overprofit of the cartel (which is equal to 10). These results plead in favor of severe fines. The effect of a fine of 15 is noticeable in the case of the leniency sanction scheme as the reduction in cartel-type agreements is comparable to the one in the monetary sanction scheme.

Result 2 (fines). *Higher fines are associated with lower propensities to choose B and to engage in cartel-type agreements in all sanction schemes, especially when it exceeds cartel overprofit.*

Table 2 – Proportions of Choice B and cartel-type agreements by sanction type and fine

Fine	5		10		15	
	Choice B	Cartel-type agreement	Choice B	Cartel-type agreement	Choice B	Cartel-type agreement
All data	0.67	0.54	0.58	0.44	0.39	0.24
Baseline	0.74	0.61	0.68	0.52	0.44	0.29
Compliance	0.63	0.52	0.57	0.45	0.35	0.23
Exclusion	0.53	0.38	0.43	0.28	0.31	0.17
Leniency	0.79	0.67	0.69	0.54	0.48	0.29

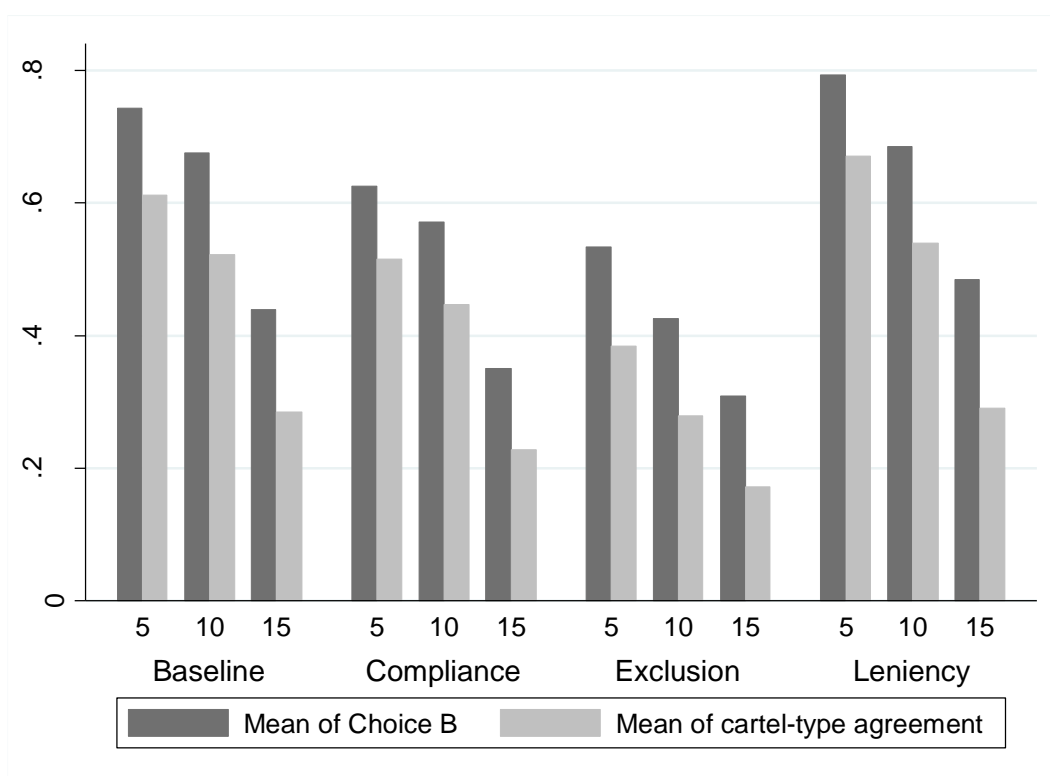


Figure 2 – Proportions of Choice B and cartel-type formation by sanction type and fine

Probabilities of detection have the expected effect in that higher probabilities of detection are associated with a lower propensity to choose Choice B and to engage in cartel-type agreements. Table 3, Figure 3 and Panel b. in Table A1 report the results.

Result 3 (probability of detection). *High probabilities of detection are associated with lower cartel choice and formation propensities to choose Decision B.*

Table 3 – Proportions of Choice B and cartel-type agreements by sanction type and probability of detection

Cartel Probability	1/10		1/3		2/3		9/10	
	Choice	Cartel-type agreement	Choice	Cartel-type agreement	Choice	Cartel-type agreement	Choice	Cartel-type agreement
Combined	0.80	0.68	0.67	0.52	0.42	0.26	0.33	0.19
Baseline	0.86	0.75	0.75	0.60	0.48	0.31	0.40	0.24
Compliance	0.78	0.69	0.64	0.51	0.37	0.24	0.28	0.15
Exclusion	0.73	0.60	0.52	0.36	0.26	0.10	0.18	0.5
Leniency	0.82	0.69	0.76	0.61	0.56	0.38	0.48	0.32

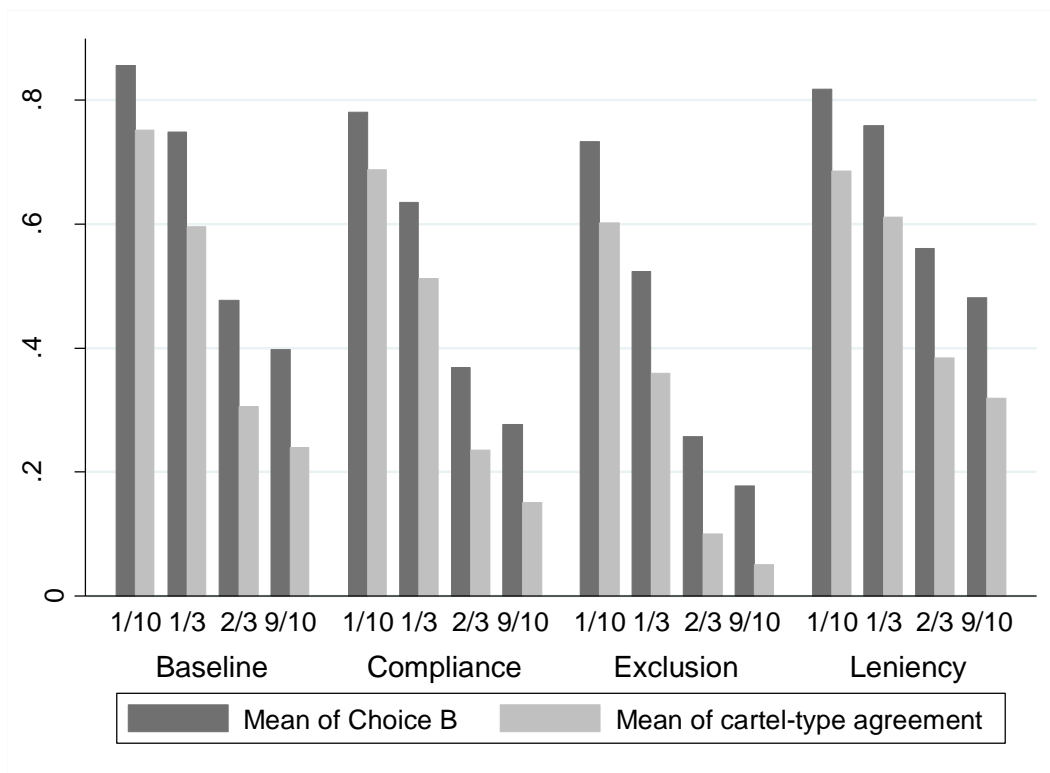


Figure 3 – Proportions of Choice B and cartel-type agreements by sanction type and probability of detection

Table 4 reports a logistic regression of the likelihood of Choice B by moral costs which shows the robustness of the effects of sanction schemes, fines and the probabilities of detection across moral cost conditions. These results attest that participants are impacted by these sanctions in a systematic fashion.

Table 4 – Logistic regression of the likelihood of Choice B by moral costs (marginal effects)

	No loss	Low loss	High loss	Full loss
Risk aversion	-.075	-.045	-.047	-.016
	(.049)	(.056)	(.086)	(.073)
Gender (men = 1)	.011	.075	.048	.010
	(.057)	(.070)	(.061)	(.046)
<i>Benchmark – Fine of 5</i>				
Fine of 10	-.110***	-.102***	-.108***	-.118***
	(.021)	(.021)	(.025)	(.024)
Fine of 15	-.359***	-.341***	-.305***	-.353***
	(.036)	(.045)	(.044)	.044

<i>Benchmark – Probability of detection</i> (1/9 or 2/3)	-	-	-	-
Probability of detection (0.66)	-.409*** (.044)	-.349*** (.041)	-.328*** (.039)	-.361*** (.049)
Probability of detection (0.9)	-.457*** (.040)	-.464*** (.044)	-.399*** (.049)	-.448*** (.054)
<i>Benchmark – Monetary</i>	-	-	-	-
Compliance	-.160*** (.027)	-.134*** (.024)	-.150*** (.023)	-.084*** (.034)
Exclusion	-.282*** (.030)	-.284*** (.040)	-.245*** (.031)	-.169*** (.032)
Leniency	.011 (.036)	.026 (.031)	.045 (.041)	.097*** (.032)
Repetition of the decision in the decision environment	-.003 (.015)	.023 (.015)	.009 (.014)	-.023 (.014)
Probability of Choice B	.587	.553	.565	.560
N	4608	4608	4608	4608
Log likelihood	-2527.544	-2574.996	-2677.853	-2609.497
P-value	.000	.000	.000	.000

Another way to investigate the impact of sanctions is to study the combination of fines and probabilities of detection as this combination allows us to measure the sensitivity of participants to the cartel-type agreement expected overprofit. Table 5 reports the cartel-type agreement expected overprofit, computed using the following method:

- In the *Monetary sanction scheme*, the cartel expected profit is equal to $(20 - \text{fine} * \text{probability of detection}) - 10$.
- In the *Compliance sanction scheme*, the cartel overprofit is equal to the expected overprofit in the monetary sanction scheme – an opportunity cost corresponding to the fact that in case of detection, Player 1 will not be given the option to choose the cartel option in the following round. This opportunity cost is equal to the probability of detection times the cartel expected overprofit in the following round.
- In the *Exclusion sanction scheme*, the cartel overprofit is equal to the expected overprofit in the monetary sanction scheme - an opportunity cost corresponding to the fact that in case of detection, player 1 will not be given the option to play. This opportunity cost is equal to the probability of

detection*(normal competitive profit (i.e. 10) plus cartel expected overprofit in the following round).

- In the *Leniency sanction scheme*, in absence of denunciation, the cartel expected overprofit is equal to the cartel expected overprofit in the baseline. In the case of denunciation, both players have ex-ante the same chance to obtain immunity as it is based on the speed of denunciation. To avoid biasing the computation, the probability that a Type 1 participant will denounce the cartel-type agreement is set at 50%.

Table 5 - Cartel-type agreement expected overprofit amount

Fine	Probability of fine	Baseline	Compliance (repetition 1)	Compliance (repetition 2)	Exclusion (repetition 1)	Exclusion (repetition 2)	Leniency
5	0.9	5.5	0.55	1	-8.45	-8	6.5
10	0.9	1	0.1	-3.5	-8.9	-12.5	3
15	0.9	-3.5	-0.35	-8	-9.35	-17	0.5
5	0.66	6.67	2.22	3.33	-4.44	-3.33	7.08
10	0.66	3.34	1.11	0	-5.55	-6.67	4.17
15	0.66	0	0	-3.33	-6.67	-10	1.25
5	0.33	8.34	5.56	6.67	2.22	3.33	7.91
10	0.33	6.67	4.44	5	1.11	1.67	5.83
15	0.33	5	3.33	3.33	0	-0.5	3.75
5	0.1	9.5	9	9	7.55	8	8.5
10	0.1	9	8.5	8.5	7.1	7.5	7
15	0.1	8.5	8	8	6.65	7	5.5

The rationale behind the use of cartel-type agreement expected overprofit is that if Type 1 participants are sensitive to the cartel-type agreement expected overprofit, they should take Decision B when the cartel-type agreement expected overprofit is greater than zero. Figure 4 reports the propensity to take Choice B by sanction type and by Type 2 participant loss. It shows that this propensity to take Choice B increases as cartel-type agreements' overprofit increases. The results show that although Player 1 participants are more likely to take Choice B when the expected overprofit is high, their propensity to take Choice B is greater than zero when the expected overprofit is negative.

Result 4 (expected overprofit). *Type 1 participants' propensity to choose decision B increases as cartel-type agreement expected overprofit increases. Their propensity to take Choice B is greater than zero when the cartel-type agreement expected overprofit is negative.*

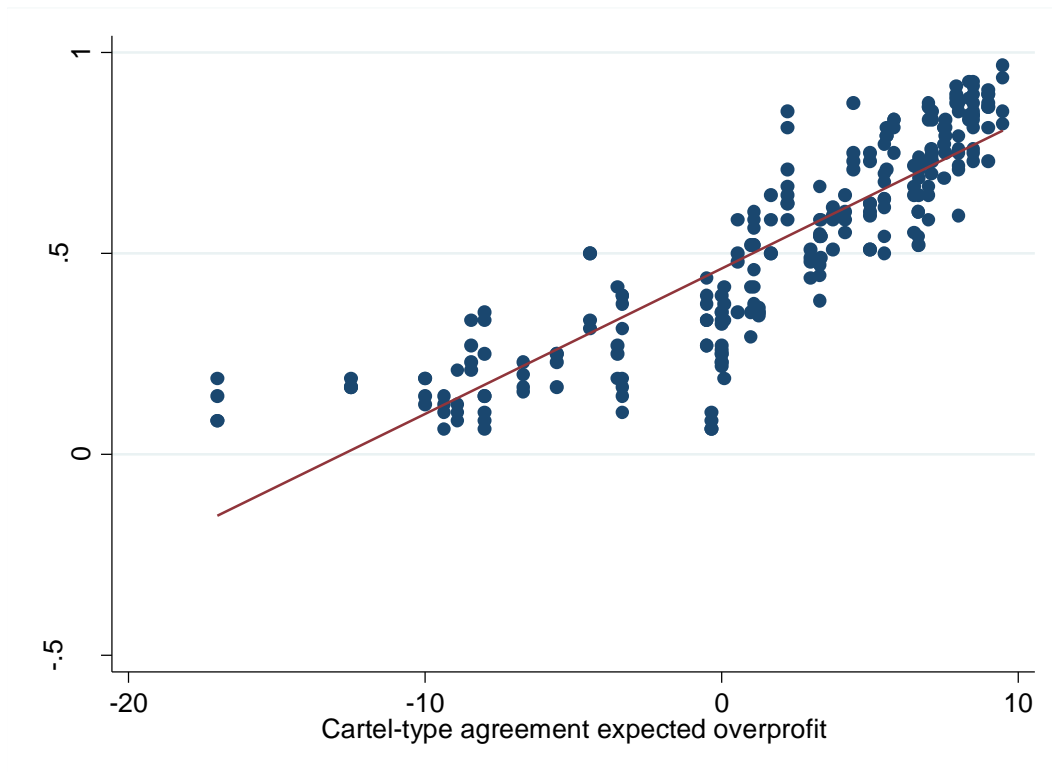


Figure 4 – Choice B by cartel-type agreement by expected overprofit

4.2. Moral costs

In addition to the studying of sanction schemes, the studies the impact of moral cost, which was implemented in a between-subject design in the baseline and the three treatments. The between-subject design allows us to test the robustness of the results on sanctions schemes as well as documenting the impact of moral costs on the propensity to engage in a cartel-type agreement in a clean fashion. Table 6 and Figure 5 document the proportion of Choice B by sanction scheme and moral cost. Overall, the results on the effectiveness of sanction schemes are robust to the introduction of varying levels of moral cost: exclusion is the most effective sanction scheme, followed by the compliance sanction scheme, regardless of the level of loss imposed on Type B participants.

Table 6 – Proportion of Choice B by sanction scheme and moral cost

Treatment	Baseline	Leniency	Compliance	Exclusion	Total
-----------	----------	----------	------------	-----------	-------

Choice B	No loss (0)	0.65	0.66	0.53	0.43	0.57
	Low loss (3)	0.62	0.64	0.52	0.39	0.54
	High loss (7)	0.63	0.66	0.50	0.42	0.55
	Full loss (10)	0.58	0.66	0.51	0.45	0.55
Cartel-type agreements	No loss (0)	0.50	0.51	0.40	0.30	0.43
	Low loss (3)	0.49	0.51	0.41	0.26	0.42
	High loss (7)	0.46	0.47	0.38	0.26	0.39
	Full loss (10)	0.44	0.51	0.40	0.30	0.41

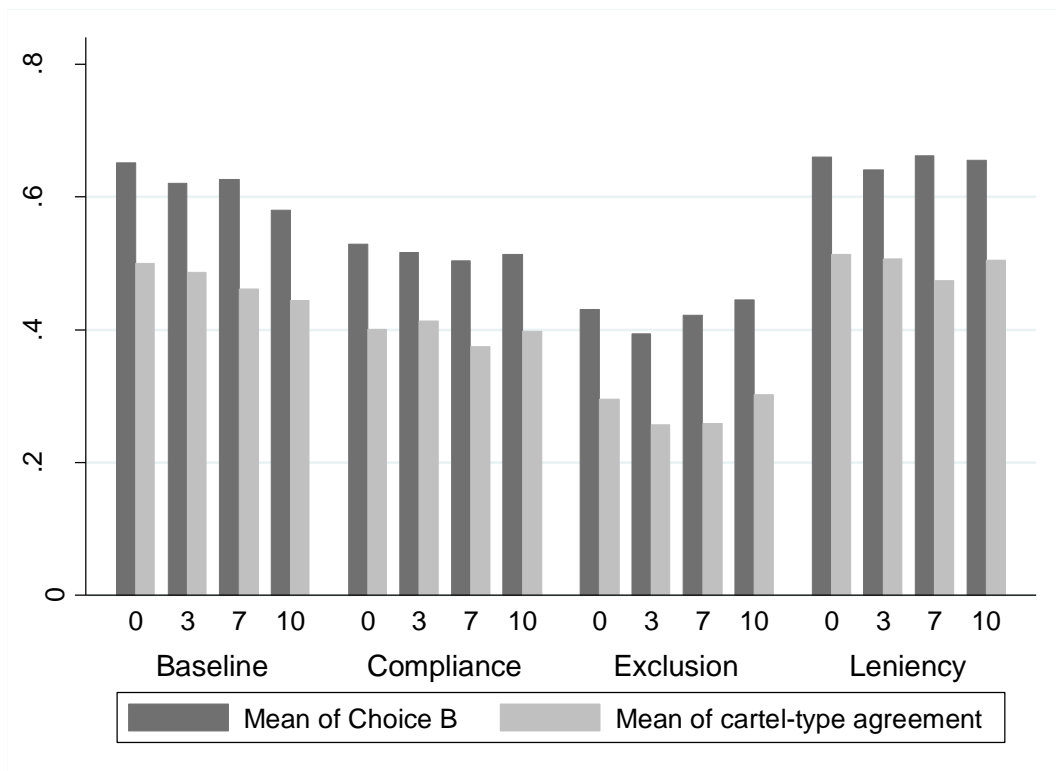


Figure 5 – Choice B by sanction scheme and moral costs

The effect of moral costs is statistically significant but small in magnitude, as, any loss imposed on Type B participants as opposed to no loss causes a reduction in the propensity to choose Decision B of about 2%. Indeed, the propensity to choose Choice B in any loss condition (loss of 3, 7 or 10) is statistically lower than in the No loss (loss of 0) condition (p -value = 0.0234), indicating that participants take, at least to some extent, moral cost into account. When considering pairwise comparisons, the propensity to choose Choice B in the Low loss (loss of 3) is statistically lower than in the No loss (loss of 0) condition (p -value = 0.0179). This propensity is also lower in the Full loss (loss of 10) condition compared to the No loss (loss of 0) condition but only at the 10% level (p -value =

0.065). This leads to Result 5:

Result 5 (moral cost). *Moral cost reduces significantly Type 1 participants' propensity to choose Decision B but in a magnitude that is small as it causes a reduction in the propensity to choose Decision B of about 2%.*

Table 7 reports the effect of moral cost controlling for fines, probabilities, individual characteristics and the repetition of the decision by sanction scheme. It documents that, *ceteris paribus*, the impact of loss remains only significant in the Baseline condition and when there is a full loss imposed on Type 2 participants. This confirms that the effect of moral cost is, relative to the effect of sanctions, small. In addition to the non-parametric tests, it allows us to document that the effect of moral cost is significant only when in the baseline and when the loss is total for Type 2 participants. This suggests that moral cost impacts decisions when the sanction scheme is simple, and when the loss imposed on Type 2 participants is high. Outside these conditions, sanction schemes are the only tools that reduce the propensity to choose Decision B and that prevent the formation of cartel-type agreements.

Table 7 – Logistic regression of the likelihood of Choice B by sanction schemes (marginal effects)

	Baseline	Compliance	Exclusion	Leniency
<i>Benchmark – No loss (0)</i>	-	-	-	-
Low loss (3)	-.044 (.022)	-.018 (.023)	-.047* (.022)	-.029 (.021)
High loss (7)	-.033 (.022)	-.036 (.023)	-.005 (.022)	-.004 (.021)
Full loss (10)	-.095*** (.023)	-.026 (.023)	.026 (.023)	-.017 (.021)
<i>Benchmark – Fine of 5</i>	-	-	-	-
Fine of 10	-.089*** (.020)	-.067*** (.020)	-.128*** (.018)	-.139*** (.018)
Fine of 15	-.367*** (.019)	-.326*** (.018)	-.260*** (.017)	-.354*** (.018)
<i>Benchmark – Probability of detection (1/9 or 2/3)</i>	-	-	-	-
Probability of detection (0.66)	-.390*** (.018)	-.361*** (.016)	-.345*** (.013)	-.272*** (.019)

Probability of detection (0.9)	-.465***	-.445***	-.418***	-.354***
	(.017)	(.015)	(.013)	(.018)
Risk aversion	-.039*	.006	-.084***	-.019
	(.018)	(.019)	(.019)	(.017)
Gender (men = 1)	.041**	.037*	-.014	.079***
	(.015)	(.016)	(.016)	(.014)
Repetition of the decision in the decision environment	.024	-.050**	.001	.025
	(.015)	(.016)	(.016)	(.014)
Probability of Choice B	.651	.517	.400	.682
N	4608	4608	4608	4608
Log likelihood	-2505.65	-2664.78	-2604.13	-2575.01
P-value	0.0000	0.0000	0.0000	0.000

Result 6 (moral cost and sanction schemes). *Ceteris paribus and by sanction scheme, moral cost impacts in a statistically significant fashion the propensity to choose Decision B in the baseline and when the loss imposed on Type B participants is total.*

The relationship between the propensity to choose Decision B and cartel-type agreement expected overprofit – as reported in Figure 4 – is robust to the varying levels of moral cost, as shown in Figure 6. Figure 6 further shows that dispersion around the cartel-type agreement expected overprofit is mainly driven by the compliance sanction scheme and the exclusion sanction scheme, which means that there is higher heterogeneity in behavior when sanctions are strict. This result is coherent with Result 6, which suggests that simple sanctions are stronger deterrent, and that stricter sanctions accentuate differences in the loyalty/fairness arbitration. This leads to Result 7:

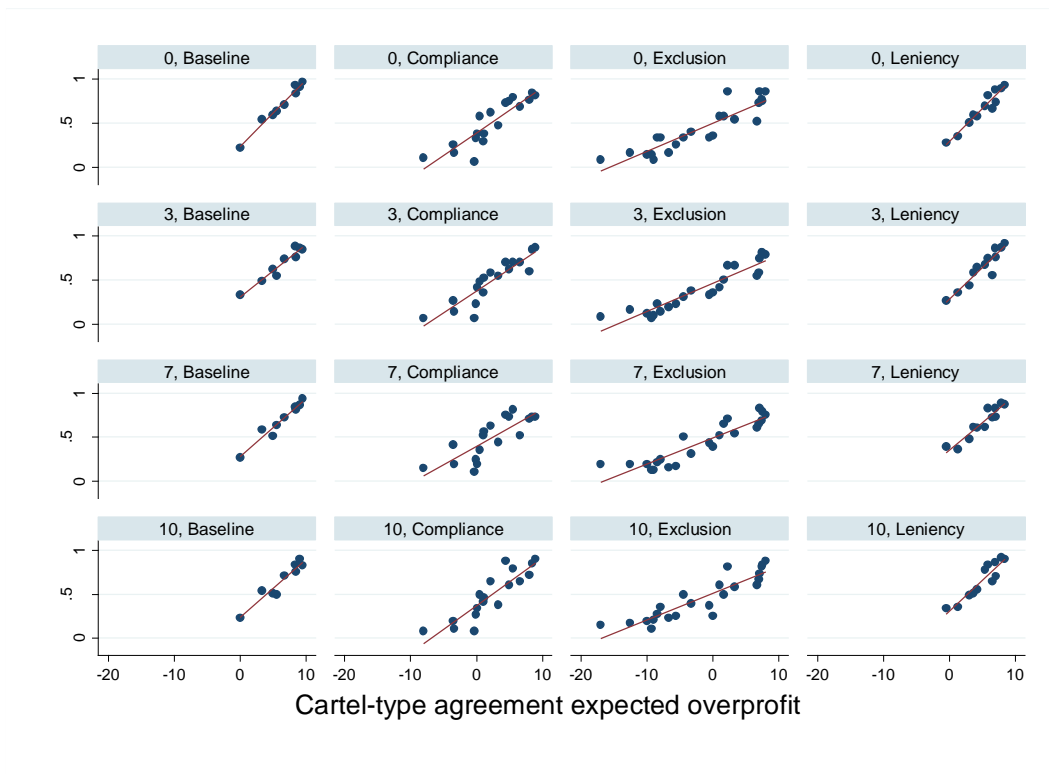


Figure 6 – Choice B by cartel-type agreement expected overprofit, by sanction scheme and by moral cost

Result 7 (influence of sanctions scheme on the cartel overprofit/loss trade-off). *Strict sanction schemes, i.e. compliance and exclusion lead to more heterogeneous behavior than looser sanctions, i.e. the baseline and leniency. The impact cartel-type agreement expected overprofit is robust to the introduction of different moral costs.*

4.3. Individual characteristics

Risk attitudes can affect the decision to form a cartel. Risk aversion is associated with a lower probability of breaking the law. In terms of risk preference, Figure 7 reports that higher risk aversion is associated with a lower propensity to choose the cartel decision. A logistic regression of the impact of the number of safe choices in the Holt and Laury (2002) task leads to a small (-0.009, with a constant of 0.55 and standard error of 0.001) but statistically significant (p -value = 0.000) decrease in the likelihood to choose Decision B. This suggests that risk aversion plays a role in the choice of engaging in such cartel-type agreement, but that the magnitude of this effect is limited.

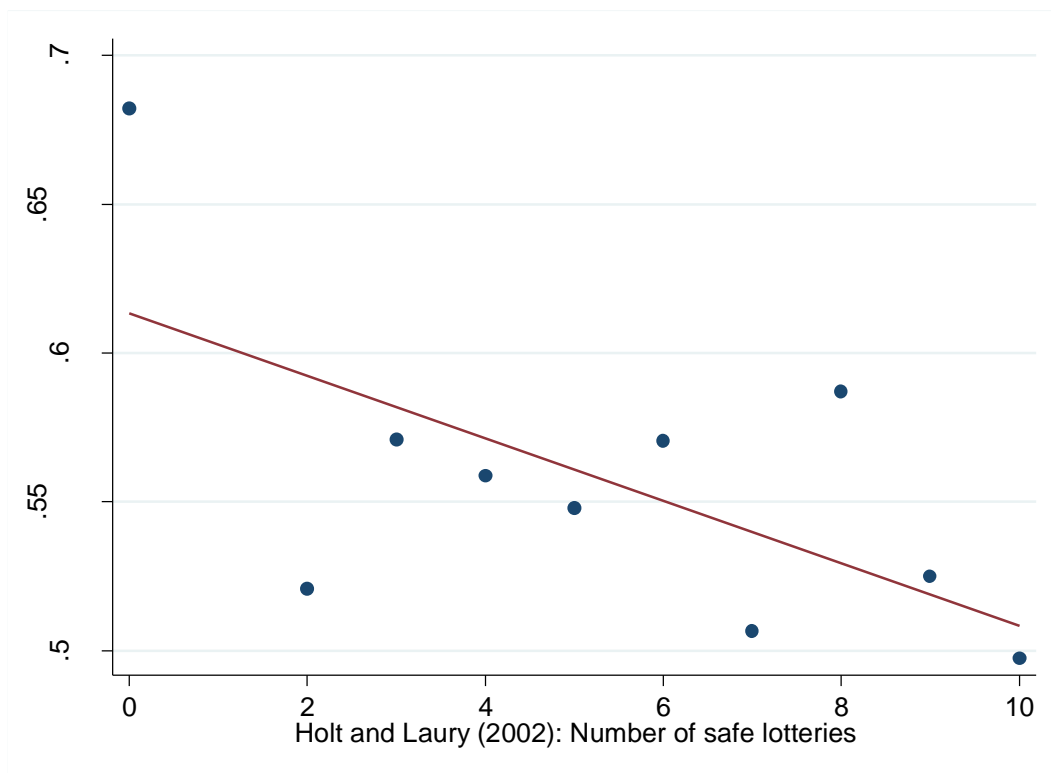


Figure 7 – Proportion of cartel decisions over the number of safe choices in a Holt and Laury (2002)

Result 8 (influence of risk aversion). *Risk aversion reduces the propensity to choose Choice B.*

The studying of the impact of gender shows that, when using all data, the propensity to choose Decision B is higher among men than women (p -value = 0.000). Men choose 56.72% of the time the cartel decision while women choose it 53.76% of the time. When comparing the individual propensities of men and women participants, the statistical significance disappears in all conditions except the baseline. The results are reported in Table 8.

Table 8 – Choice B by sanction scheme and moral cost

	Women (N=91)			Men (N=101)			Comparison (p -value)	
	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.	All data	Individual participant average
Combined	8736	0.53	0.49	9696	0.56	0.49	0.000	0.640
Baseline	2184	0.60	0.49	2424	0.63	0.48	0.044	0.033
Leniency	2184	0.62	0.49	2424	0.69	0.46	0.000	0.692
Compliance	2184	0.50	0.50	2424	0.53	0.50	0.062	0.942
Exclusion	2184	0.43	0.49	2424	0.42	0.49	0.665	0.953

Result 9 (influence of gender). *Men are found to have a higher propensity to choose Decision B.*

5. Conclusion

This paper proposes an experimental analysis of the impact of sanctions schemes and moral costs on the propensity to engage in cartel-type agreements. Our results show that compliance and exclusion sanctions schemes are more effective at deterring cartel-type agreements than monetary and leniency sanction schemes. However, the leniency sanction scheme leads to an increase in the rate of detection. Higher fines are found to be associated with lower propensities to engage in cartel-type agreements in all sanction schemes, especially when such fines exceed cartel overprofit. High probabilities of detection are also associated with lower cartel choice and formation propensities to engage in cartel-type agreements. Moral cost is found to reduce in a statistically significant fashion the propensity to engage in cartel-type agreements, notably when the sanction is simple as in the monetary sanction scheme. Concerning individual characteristics, risk aversion and being a woman is found to reduce the propensity to engage in cartel-type agreements.

Such results allow us to help document the conditions under which efficient competition policy may be envisioned. Thus, our behavioral approach contributes to better evaluate antitrust policy (see Bailey 2015 for a discussion on the relationship between behavioral economics and antitrust policy).

Concerning sanctions schemes, our results argue for severe fines as well as regular detections of cartels. Severe fines and high probabilities of detection may help improve the deterrent effect of the leniency sanction schemes, which is similar to the results reported by Chowdhury *et al.* 2018. Concerning detection, our results nuance the Beckerian substitutability between fines and probability of detection, as the level of the probability is found to matter. An important result of our experimental analysis is that exclusion is the most effective sanction scheme in deterring cartels. This result pleads for the development of sanctions to be imposed on the individuals who engage in cartel-type agreements, which is in line with the public policy discourse on the value of incapacitation sanctions, whose benefits have been highlighted in the literature (Hammond 2010). In addition, they are found to be better accepted by the general public than by prison sentences (Stephan 2015). Moreover, the results on the small but statistically significant effect of moral costs argue for strengthening corporate compliance programs which should emphasize the harm inflicted by these cartel-type agreements on customers or businesses that get harmed by these agreements.

Our experimental setting could be extended in future research. First, the fact that participants remain anonymous may impact behavior. A relevant development could be to remove anonymity as

Kirchler *et al.* (2016) find that the removal of anonymity, making participants identifiable, can have effects on market decisions as these change perceptions of responsibility. Second, experimental studies, notably of compliance and exclusion sanction schemes as they turn out to be more effective at deterring cartel, could further focus on training on legal risks and damages inflicted on third parties, i.e. consumers or businesses. It would therefore be relevant to replicate our experiment on individuals who have previously received training on antitrust regulations.

References

- Abbink, K., Irlenbusch, B., & Renner, E. (2002). An Experimental Bribery Game. *Journal of Law, Economics, & Organization*, 18(2), 428–454.
- Andres, M., Bruttel, L. & Friedrichsen, J. (2023). How Communication Makes the Difference between a Cartel and Tacit Collusion: A Machine Learning Approach. *European Economic Review*, 152(104331).
- Andres, M., Bruttel, L. & Friedrichsen J. (2021). The Leniency Rule Revisited: Experiments on Cartel Formation with Open Communication. *International Journal of Industrial Organization*, 76(102728).
- Apesteguia, J., Dufwenberg, M., & Selten, R. (2007). Blowing the Whistle. *Economic Theory*, 31(1), 143–166.
- Bailey, E. M. (2015). Behavioral Economics and U.S. Antitrust Policy. *Review of Industrial Organization*, Special Issue: Behavioral Industrial Organization, 4(3), 355–366.
- Becker, G. (1968). Crime and Punishment: An Economic Approach, *Journal of Political Economy*, 76(2), 169–217.
- Bennett, V., Pierce, L., Snyder, J., & Toffel, M. (2013). Customer-Driven Misconduct: How Competition Corrupts Business Practices, *Management Science*, 59(8), 1725–1742.
- Bernhardt, D. & Rastad, M. (2016). Collusion Under Risk Aversion and Fixed Costs, *The Journal of Industrial Economics*, 64(4), 808–834.
- Bigoni M., Fridolfsson, S.-O., Le Coq, C. & Spagnolo, G. (2012). Fines, leniency, and rewards in antitrust. *The RAND Journal of Economics*, 43(2), 368–390.
- Bigoni, M., Fridolfsson, S.-O., Le Coq, C. & Spagnolo, G. (2015). Trust, Leniency, and Deterrence, *The Journal of Law, Economics, & Organization*, 31(4), 663–689.
- Carson, T. L. (2003). Self-interest and business ethics: Some lessons of the recent corporate scandals. *Journal of Business Ethics*, 43, 389–394.

- Chen, S. (2010). The role of ethical leadership versus institutional constraints: A simulation study of financial misreporting by CEOs. *Journal of Business Ethics*, 93, 33–52.
- Chowdhury, S., Subhasish M., Wandshneider, F. (2018). Antitrust and the ‘Beckerian Proposition’: the Effects of Investigation of Fines on Cartels, In: Victor J. Tremblay & Elizabeth Schroeder & Carol Horton Tremblay (ed.), *Handbook of Behavioral Industrial Organization*, chapter 14, 368–403, Edward Elgar Publishing.
- Chowdhury, S. M., & Crede, C. (2020). Post-Cartel Tacit Collusion: Determinants, Consequences, and Prevention. *International Journal of Industrial Organization*, 70, 1–16.
- Combe, E. & Monnier-Schlumberger, C. (2016). Lutte contre les cartels and comportement des managers, *Revue Concurrences*, 4(81860), 51–68.
- Cohen, J., Ding, Y., Lesage, C. & Stolowy, H. (2011). Corporate Fraud and Managers’ Behavior: Evidence from the Press, *Journal of Business Ethics*, 95(2), 271–315.
- Davis, D. & Holt, C. (1998). Conspiracies and Secret Price Discounts, *Economic Journal*, 108(448), 736–756.
- Eisenberg, N. (2000). Emotion, regulation, and moral development, *Annual Review of Psychology*, 51, 665–697.
- Elliston, F.A. (1982). Anonymity and whistleblowing, *Journal of Business Ethics*, 1, 167–177.
- European Commission (2021). https://ec.europa.eu/commission/presscorner/detail/ro/IP_16_2582
- Falk, A. & Szech, N. (2013). Morals and markets. *Science*, 340(6133), 707–711.
- Fleckinger, P., Lafay, T. & Monnier, C. (2013). Rémunération des dirigeants et risque de fraude d'entreprise. *Revue économique*, 64, 457–467.
- Fonseca, M.A., Gonçalves, R., Pinho, J. & Tabacco, G.A., (2022). How do antitrust regimes impact on cartel formation and managers’ labor market? An experiment, *Journal of Economic Behavior and Organization*, 204(C), 643–662.

Fonseca, M.A. & Normann, H.T. (2012). Explicit vs. tacit collusion—The impact of communication in oligopoly experiments. *European Economic Review*, 56(8), 1759–1772.

Frost, A. G., & Rafilson, F. M. (1989). Overt integrity tests versus personality-based measures of delinquency: An empirical comparison. *Journal of Business and Psychology*, 3, 269–277.

Gino F. & Pierce, L. (2009). Dishonesty in the Name of Equity. *Psychological Science*, 20(9), 1153–1160.

Hamaguchi Y., Kawagoe, T, & Shibata, A. (2009). Group size effects on cartel formation and the enforcement power of leniency programs. *International Journal of Industrial Organization*, 27(2), 145–165.

Hammond, S. (2010). The evolution of criminal antitrust enforcement over the last two decades, Speech before the National Institute on White Collar Crime, <http://www.justice.gov/atr/public/speeches/255515.htm>.

Hinloopen, J. & Soetevent, A.R. (2008). Laboratory evidence on the effectiveness of corporate leniency programs. *RAND Journal of Economics*, 39(2), 607–616.

Holt, C.A. (1995) Industrial Organization: A Survey of Laboratory Research. *The Handbook of Experimental Economics*, In John H. Kagel and Alvin E. Roth (ed.), Princeton: Princeton University Press, pp. 349–444.

Holt, C.A. & Laury, S.K. 2002. Risk Aversion and Incentive Effects. *American Economic Review*, 92(5), 1644–1655.

Jensen, M. & Meckling, W. (1976). Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. *Journal of Financial Economics*, 3(4), 305–360.

Jubb, P.B. (1999). Whistleblowing: A Restrictive Definition and Interpretation. *Journal of Business Ethics*, 21, 77–94.

Kerr, N.L., Rumble, A. C., Park, E.S., Ouwerkerk, J.W., Parks, C.D., Gallucci, M., & Van Lange, P.A. (2009). How many bad apples does it take to spoil the whole barrel? : Social exclusion and toleration for bad apples. *Journal of Experimental Social Psychology* 45(4), 603–613.

Kirchler, M., Huber, J., Stefan, M. & Sutter, M. (2016). Market design and moral behavior, *Management Science*, 62(9), 2615–2625.

Koike, S., Nakamaru, M., Otaka, T., Shima, H., Shimomura, K.I., Yamato, T. (2018). Reciprocity and exclusion in informal financial institutions: An experimental study of rotating savings and credit associations. *PLOS ONE*, 13(8), e0202878.

Larkin, I., & Pierce, L. (2016). Compensation and employee misconduct: The inseparability of productive and counterproductive behavior in firms. In D. Palmer, K. Smith-Crowe, & R. Greenwood (Eds.), *Organizational Wrongdoing: Key Perspectives and New Directions* (Cambridge Companions to Management, pp. 270-304). Cambridge: Cambridge University Press.

Marcus, B., & Schuler, H. (2004). Antecedents of counterproductive behavior at work: A general perspective. *Journal of Applied Psychology*, 89(4), 647–660.

Masclet, D. (2003). Ostracism in work teams: a public good experiment. *International Journal of Manpower*, 24(7), 867–887.

Murphy, R.O., Ackermann, K.A., & Handgraaf, M.J.J. (2011). Measuring Social Value Orientation (SVO). *Judgment and Decision Making*, 6(8), 771–781.

Normann, H. T., & Müller, W. (2015). Experimental economics in antitrust. In R. D. Blair, & D. D. Sokol (Eds.), *The Oxford Handbook of International Antitrust Economics, Vol 1*, 229–253). (Oxford Handbooks in Economics). Oxford University Press.

Parasuram B., Bennett V. & Pierce L. (2017) The Wages of Dishonesty: The Supply of Cheating Under High-Powered Incentives. *Journal of Economic Behavior & Organization*, 137, 428–444.

Plott, C. (1982). Industrial Organization Theory and Experimental Economics, *Journal of Economic Literature*, 20(4), 1485-1527.

Ramdani, D. & Van Witteloostuijn, A. (2012). The Shareholder-Manager Relationship and Its Impact on the Likelihood of Firm Bribery. *Journal of Business Ethics*, 108(4), 495–507.

- Reed, A. I. I., & Aquino, K. F. (2003). Moral identity and the expanding circle of moral regard toward out-groups. *Journal of Personality and Social Psychology*, 84(6), 1270–1286.
- Rijsenbilt, A. & Commandeur, H. (2012). Narcissus Enters the Courtroom: CEO Narcissism and Fraud, *Journal of Business Ethics*, 117(2), 413–429.
- Rivas, M.F. (2013). An experiment on corruption and gender. *Bulletin of Economic Research*, 65(1), 10–42.
- Robinson, R., Lewick, J. & Donahue, E. (2000). Extending and testing a five factor model of ethical and unethical bargaining tactics: introducing the SINS scale, *Journal of Organizational Behavior*, 21(6), 649–664.
- Shleifer, A. (2004). Does competition destroy ethical behavior? *American Economic Review, Papers and Proceedings*, 94(2), 414–418.
- Soldà, A. & Villeval, M.C. (2020). Exclusion and Reintegration in a Social Dilemma. *Economic Inquiry*, 58(1), 120–149.
- Stephan, A. (2011). Disqualification Orders for Directors Involved in Cartels, *Journal of European Competition Law & Practice*, 2(6), 529–536.
- Stephan, A. (2015). Survey of Public Attitudes to Price Fixing in the UK, Germany, Italy and the USA, *Centre for Competition Policy UEA Law School University of East Anglia, CCP Working Paper 15–8*.
- Stucke M. (2011). Am I a Price-Fixer? A Behavioral Economics Analysis of Cartels, in Beaton-Wells C. & Ezrachi I. (dir.), *Criminalizing Cartels: A Critical Interdisciplinary Study of an International Regulatory Movement*, Oxford, Hart Publishing.
- Sutter, M., Huber, J., Kirchler, M., Stefan, M. & Walzl, M. (2020). Where to look for the morals in markets, *Experimental Economics*, 23, 30–52.
- Tan, F. & Yim, A. (2014). Can strategic uncertainty help deter tax evasion? An experiment on auditing rules, *Journal of Economic Psychology*, 40(C), 161–174.

Van Staveren, I. (2014). The Lehman Sisters hypothesis, *Cambridge Journal of Economics*, 38(5), 995–1014.

Waytz, A., Dungan, J. & Young, L. (2013). The whistleblower's dilemma and the fairness–loyalty tradeoff, *Journal of Experimental Social Psychology*, 49(6), 1027–1033.

Zahra, S.A., Priem, R.L., & Rasheed, A. (2005). The antecedents and consequences of top management fraud. *Journal of Management*, 31(6), 803–828.

Zizzo, D.J. & Oswald, A.J. (2001). Are People Willing to Pay to Reduce Others' Incomes? *Annales d'Économie et de Statistique*, 63/64, 39–65.

Zona, F., Minoja, M. & Coda, V. (2013). Antecedents of Corporate Scandals: CEOs' Personal Traits, Stakeholders' Cohesion, Managerial Fraud, and Imbalanced Corporate Strategy, *Journal of Business Ethics*, 113(2), 265–283.

Appendix

Table A1 – Comparison of choice B by fine (Panel a.) and by probability of detection (Panel b.) by moral cost (p-value - χ^2)

		Baseline		Leniency		Compliance		Exclusion	
		Choi ce	Cartel formation	Choi ce	Cartel formation	Choi ce	Cartel formation	Choi ce	Cartel formation
<i>a. Fine</i>									
All data	5 vs 10	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000
	10 vs 15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	5 vs 15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
No cost (0)	5 vs 10	0.045	0.003	0.000	0.001	0.269	0.061	0.001	0.001
	10 vs 15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	5 vs 15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Low cost (3)	5 vs 10	0.082	0.108	0.002	0.000	0.121	0.043	0.007	0.012
	10 vs 15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	5 vs 15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
High cost (7)	5 vs 10	0.002	0.000	0.000	0.000	0.214	0.246	0.014	0.018
	10 vs 15	0.000	0.000	0.000	0.000	0.000	0.000	0.021	0.005
	5 vs 15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Full cost (10)	5 vs 10	0.125	0.148	0.001	0.000	0.027	0.009	0.000	0.000
	10 vs 15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	5 vs 15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<i>b. Probability of detection</i>									
All data	1/3 vs 2/3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1/10 vs 9/10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
No cost (0)	1/3 vs 2/3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1/10 vs 9/10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Low cost (3)	1/3 vs 2/3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1/10 vs 9/10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1/3 vs 2/3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

High cost (7)	1/10 vs 9/10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Full cost (10)	1/3 vs 2/3 1/10 vs 9/10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000