UNIVERSITY OF ALBERTA
FACULTY OF ARTS
Department of Economics

Working Paper No. 2018-13

Ordered Leniency: An Experimental Study of Law Enforcement with Self-Reporting

Claudia M. Landeo
University of Alberta

Kathryn E. Spier
Harvard Law School

September 2018

Ordered Leniency: An Experimental Study of Law Enforcement with Self-Reporting*

Claudia M. Landeo† and Kathryn E. Spier‡

September 13, 2018

Abstract

This paper reports the results of an experiment designed to assess the ability of an enforcement agency to detect and deter harmful short-term activities committed by groups of injurers. With ordered-leniency policies, early cooperators receive reduced sanctions. We replicate the strategic environment described by Landeo and Spier (2018). In theory, the optimal ordered-leniency policy depends on the refinement criterion applied in case of multiplicity of equilibria. Our findings are as follows. First, we provide empirical evidence of a "race-to-the-courthouse" effect of ordered leniency: Mild and Strong Leniency induce the injurers to self-report promptly. These findings suggest that the injurers' behaviors are aligned with the risk-dominance refinement. Second, Mild and Strong Leniency significantly increase the likelihood of detection of harmful activities. This fundamental finding is explained by the high self-reporting rates under ordered-leniency policies. Third, as a result of the increase in the detection rates, the averages fines are significantly higher under Mild and Strong Leniency. As expected when the risk-dominance refinement is applied, Mild Leniency exhibits the highest average fine.

KEYWORDS: Law Enforcement; Ordered Leniency; Self-Reporting; Experiments; Leniency; Coordination Game; Prisoners' Dilemma Game; Risk Dominance; Pareto Dominance; Equilibrium Selection; Non-Cooperative Games; Harmful Externalities; Corporate Misconduct; White-Collar Crime; Securities Fraud; Insider Trading; Market Manipulation; Whistleblowers; Plea Bargaining; Tax Evasion; Environmental Policy Enforcement

JEL Categories: C72, C90, D86, K10, L23

*We acknowledge financial support from the National Science Foundation (NSF Grant SES-1155761). We thank Tom Brennan, Dan Coquillette, Nick Feltovich, John Golberg, Christine Jolls, Louis Kaplow, Max Nikitin, Jack Ochs, and Abe Wickelgren for insightful discussions and comments. We are grateful for suggestions from participants at the 2018 NBER Summer Institute in Law and Economics and the Harvard Law School Faculty Workshop. We thank Susan Norton for administrative assistance, Tim Yuan for programming the software used in this study, and the Harvard Decision Science Laboratory for assistance in conducting the experimental sessions.

†University of Alberta, Department of Economics. Henry Marshall Tory Building 7-25, Edmonton, AB T6G 2H4. Canada. landeo@ualberta.ca, tel. 780-492-2553.

‡Harvard Law School and NBER. 1575 Massachusetts Ave., Cambridge, MA 02138. United States. kspier@law.harvard.edu, tel. 617-496-0019.

1 Introduction

Illegal activities, including corporate crimes and securities fraud, are often committed by groups of wrongdoers rather than by individuals working in isolation. Law enforcement agencies have established leniency policies to encourage wrongdoers to self-report, with the goals of improving detection of socially-harmful activities and strengthening deterrence.[1] A common feature of these policies is that wrongdoers who cooperate with the agency early in the process receive reduced sanctions, i.e., law enforcement agencies apply *ordered-leniency* mechanisms (Landeo and Spier, 2018). For instance, in 2009, the Securities and Exchange Commission (SEC) brought charges against Raj Rajaratnam and other co-conspirators for insider trading at several hedge funds including Galleon Management LP and New Castle Funds LLC.[2] Anil Kumar, an early cooperator, signed a cooperation agreement and was granted leniency. In contrast to other co-conspirators, he paid a reduced fine and received no prison time. At Kumar's sentencing, U.S. Attorney Prett Bharara stated: "Kumar's *immediate* cooperation warrants special mention and recognition" (Southern District of New York, Sentencing Memorandum, 2012, p. 12; emphasis added).[3]

Kaplow and Shavell (1994) provide seminal theoretical work on the control of harmful externalities with self-reporting.[4] In the context of a single injurer, they show that enforcement with self-reporting can induce individuals to report their harmful acts without compromising deterrence. This is accomplished by allowing those who self-report to pay a sanction equal to (or slightly less) than the expected sanction they would face if they did not report the act. Given that enforcement efforts do not need to be allocated to identify the injurers who voluntarily self-report, the enforcement agency can economize on investigatory costs and social welfare rises.[5]

Landeo and Spier (2018) extend this theoretical literature by studying the design of optimal enforcement policies with self-reporting for illegal short-term activities committed by *groups of injurers*.[6] They focus their analysis on ordered-leniency mechanisms where the degree of

[1]The Securities and Exchange Commission's Cooperation Program is an example of such a leniency policy.

[2]*SEC v. Rajaratnam,* 622 F.3d 159 (2nd Cir. 2010).

[3]*United States v. Kumar*, Case 1:10-cr-00013-DC, Document 47-1 (S.D.N.Y. July 16, 2012). Additional legal cases involving ordered leniency are discussed in Landeo and Spier (2018).

[4]See Becker (1968) and Polinsky and Shavell (1984) for early work on law enforcement policies.

[5]See also Malik (1993) and Innes (1999).

[6]Short-term illegal activities do not involve an ongoing relationship among group members. They are also referred as "occasional" activities (Buccirossi and Spagnolo, 2006). In game-theoretic terms, they correspond

leniency granted to an injurer is determined by his or her position in the self-reporting queue. Landeo and Spier (2018) show that granting a penalty reduction to the first injurer to report, and possibly (albeit lower) to the subsequent injurers, generates a so-called "race to the courthouse" where, in equilibrium, all injurers self-report immediately.[7] Importantly, by inducing self-reporting, ordered leniency increases the likelihood of detection of harmful acts without raising the enforcement costs. As a result of the higher likelihood of detection, the expected fine rises, deterrence is strengthened and social welfare is improved.[8] Our paper contributes to this literature by exploring these issues in a laboratory setting. To the best of our knowledge, there are no previous experimental analyses of law enforcement with ordered leniency for short-term group activities.

Landeo and Spier (2018) demonstrate that the optimal degree of leniency for self-reporting depends critically on the refinement criterion for equilibrium selection when multiple equilibria arise. When the enforcement agency grants relatively small discounts for self-reporting (Mild Leniency), the strategic environment faced by the injurers is a coordination game with two Nash equilibria: One where all injurers self-report, the risk-dominant equilibrium (Harsanyi and Selten, 1988); and, one where no injurer self-reports, the Pareto-dominant equilibrium. When the Pareto-dominance refinement is applied, Mild Leniency is a totally ineffective policy. In that case, the enforcement agency must grant larger discounts for self-reporting (Strong Leniency). With Strong Leniency, the strategic environment faced by the injurers is a prisoners' dilemma game where self-reporting by all injurers is the unique Nash equilibrium.

Since Landeo and Spier's (2018) framework involves multiple equilibria, and since the optimal enforcement policy with ordered leniency depends on the refinement criterion for equilibrium selection, it is appropriate to use experimental economics methods. Our experimental environment replicates Landeo and Spier's (2018) theoretical setting. Our experimental design includes three leniency conditions: Strong Leniency, where the first injurer to report receives a large reduction in the penalty; Mild Leniency, where the first injurer to report receives a small reduction in the penalty; and, No Leniency, where penalty reductions for self-reporting are not

to one-shot strategic environments. Leniency programs have been also applied to long-term illegal activities. See Spagnolo and Marvão (2016) for a survey on this literature.

[7] In civil litigation cases, the expression "race to the courthouse" refers to the superior rights granted to the first action filed. In Landeo and Spier (2018), early reporting increases the likelihood of getting the first position in the self-reporting queue and, hence, of getting a larger reduction in the penalty.

[8] Feess and Walzl (2004) theoretically study enforcement with self-reporting for criminal teams, focusing on the effects of injurers' cooperation on self-reporting.

granted. Across leniency conditions, the strategic environment involves two potential injurers. The timing of the game is as follows.

First, the private benefits from committing the act and the enforcement policy are communicated to the potential injurers. The enforcement policy involves a penalty, probabilities of detection that depend on the number of injurers who self-report (the higher the number of injurers who self-report, the higher the likelihood of detection), and reduced penalties for self-reporting that depend on the injurers' positions in the self-reporting queue (ordered leniency).

Next, given the enforcement policy, the potential injurers play a two-stage game of complete information. In Stage 1, the potential injurers decide whether to participate in the act. If both potential injurers elect to participate, then the act is committed and Stage 2 starts; otherwise, the game ends. In Stage 2, the injurers decide whether and when to report themselves to the authorities. The decision of an injurer to self-report hinges on the likelihood of detection if he remains silent, which itself depends on the self-reporting decision of the other injurer. There are negative externalities in the self-reporting stage: The likelihood that an injurer will be detected is higher when the other injurer reports the act. Finally, the injurers, if detected, are sanctioned.

Our main results are as follows. First, we provide empirical evidence that ordered-leniency policies create a race to the courthouse where the majority of injurers self-report promptly. The result that ordered leniency significantly raises the likelihood of self-reporting for Mild as well as Strong Leniency suggests that the injurers' behaviors are aligned with the risk-dominance refinement. Second, Mild and Strong Leniency significantly increase the likelihood of detection of harmful activities. This fundamental finding is explained by the high self-reporting rates under ordered-leniency policies. Third, as a result of the increase in the detection rates, the averagel fines are significantly higher under Mild and Strong Leniency. As expected under the risk-dominance refinement, Mild Leniency exhibits the highest average fine. Fourth, our experimental results indicate that some injurers systematically underestimate the likelihood and severity of sanctions when making their decisions about participating in the harmful act. These findings might suggest the presence of self-serving bias on the injurers' beliefs about securing the first position in the self-reporting queue. As a consequence, the deterrence power of ordered-leniency policies is weakened, and harmful acts are committed more frequently than predicted.[9]

[9]Our findings on deterrence are aligned with previous literature on crime and deterrence. See Nagin and Pogarsky (2003) for experimental evidence of the association between individual's perception of the risk associ-

Important policy implications are derived from our analysis and findings. Contrary to the common beliefs held by legal practitioners and policy makers, our results suggest that the proverbial prisoners' dilemma game is not the only relevant strategic environment for the design of optimal enforcement policies with self-reporting and the design of optimal cooperation agreements with wrongdoers. In fact, when the wrongdoers are sufficiently distrustful of each other, the enforcement agency can induce maximal cooperation by implementing a coordination game instead. Our findings underscore the importance of combining experimental and behavioral observation with theoretical modeling.

Our paper is motivated by insider trading and securities fraud. We believe, however, that the main insights derived from our study might apply to other contexts as well. The control of harmful externalities and the implementation of law-enforcement policies with self-reporting are relevant to environments such as plea bargaining with criminal defendants (Landes, 1971; Grossman and Katz, 1983; and, Kobayashi, 1992), corporate criminal liability and third-party enforcement (Kraakman, 1986; Arlen and Kraakman, 1997; and Arlen, 2012), federal government misbehavior and the *qui tam* whistleblower program (Engstrom, 2012), corporate misconduct and the whistleblower mechanisms under the Dodd-Frank Wall Street Reform and Consumer Protection Act (Greenberg, 2011), environmental policies and standards (Livernois and McKenna, 1999), and tax evasion (Andreoni, 1991; and Malik and Schwab, 1991), among other settings.[10] Our results regarding the effects ordered-leniency policies might be useful for the design of optimal law-enforcement mechanisms in these environments too.

Our work is also related to the literature on the enforcement of competition policies for illegal long-term activities. Motta and Polo (2003) and Spagnolo (2005) provide seminal theoretical work on leniency programs for cartels.[11] They demonstrate that although leniency policies strengthen deterrence, they might also be exploited by cartel members.[12] In particular, Spagnolo (2005) shows that leniency policies might be used by cartel members as a punishment tool, and hence contribute to the stability of collusive agreements. Bigoni et al. (2012) study the effects of leniency programs in the lab. Their findings suggest that leniency policies strengthen deterrence (i.e., reduce cartel formation) but they also contribute to the

ated with the participation in harmful activities, self-serving bias, and participation in harmful activities. See Chalfin and McCrary (2017) for a survey on the economics and criminology literature on crime and deterrence. See Section 6 for further discussion.

[10]See Landeo and Spier (2018) for further discussion.

[11]See also Aubert et al. (2006).

[12]See also Chen and Rey (2013).

stabilization of surviving cartels.[13] Landeo and Spier's (2018) theoretical framework can be a useful component in the analysis of optimal enforcement with self-reporting for harmful long-term group activities. Therefore, our experimental insights might be relevant for the design of enforcement mechanisms in antitrust environments, too.

Another strand of literature related to our paper is that on contract design in the presence of externalities among contract recipients. Rasmusen, et al. (1991) and Segal and Whinston (2000) study exclusive dealing contracts in a theoretical framework. Their analysis demonstrates that incumbent monopolists can use exclusive contracts to deter efficient entry when there are economies of scale in production. Landeo and Spier (2009, 2012) provide experimental evidence of the effects of exclusive dealing contracts on market foreclosure. Their results suggest that the incumbent monopolists can design profitable exclusive-dealing contracts by exploiting the negative externalities among buyers when there are economies of scale in production.[14]

The rest of the paper is organized as follows. Section 2 outlines the theoretical model and predictions. Section 3 presents the experimental design. Section 4 discusses the qualitative hypotheses to be tested. Section 5 examines the results from the experimental sessions. Section 6 provides further discussion of our findings, proposes avenues for future research, and concludes the paper.

2 Theoretical Framework

2.1 Model Setup

Landeo and Spier (2018) consider a general model of complete information. Their benchmark model includes three risk-neutral players: Two identical potential injurers and an enforcement agency. The potential injurers maximize their private net benefits from committing a harmful act and the enforcement agency maximizes social welfare.[15] They assume that the enforcement

[13] See Apestegui et. (2007), Hinloopen and Soetevent (2008), Bigoni et al. (2015) and Feltovich and Hamaguchi (2018) for additional experimental work on leniency and cartels.

[14] See Smith (2011) for additional experimental work on exclusive contracts. See Landeo and Spier (2015) for experimental work on the design of incentive contracts for teams in the presence of externalities among team members.

[15] Social welfare includes the aggregation of the benefits to the injurers. It also includes the social costs: The harm inflicted on others (externalities associated with the harmful activities) and the cost of enforcement.

agency cannot costlessly identify the parties responsible for committing the harmful act. The timing of the game is as follows.

First, the enforcement agency publicly commits to an enforcement policy with ordered leniency for self-reporting. The enforcement policy includes a fine or monetary sanction f (measured per injurer), leniency multipliers $r_1, r_2 \in [0, 1]$ that correspond to the first and second positions in the self-reporting queue, respectively.[16] The discount for position i in the reporting queue is then $1 - r_i$, $i = 1, 2$.

Second, the potential injurers play a two-stage game. In Stage 1, they decide simultaneously and independently whether to participate in a socially-harmful group activity. The benefit of committing the act is b for each injurer. The value of b is observed by both potential injurers before they decide whether to participate in the activity. If both injurers decide to participate, then the act is committed. If the act is committed, then Stage 2 begins; otherwise, the game ends. In Stage 2, the injurers decide simultaneously and independently whether and when to report the harmful act to the enforcement agency. Each injurer can choose to report at any time in an interval, $t \in [0, \bar{t}]$.[17]

Third, the injurers, if detected, are accurately identified by the enforcement agency and are forced to pay sanctions. The probabilities of detection and the corresponding sanctions are as follows. If neither injurer self-reports, then the harmful act is detected with probability p_0 and each injurer pays fine f.[18] If exactly one injurer reports the act, then the injurer who reports pays fine $r_1 f$ and the injurer who remains silent is detected with probability p_1 and fully sanctioned (i.e., pays fine f).[19] If both injurers report the act, then the first injurer pays fine $r_1 f$ and the second injurer pays fine $r_2 f$.[20] In the event of a tie, then a coin flip (equally weighted) determines the injurers' positions in the self-reporting queue. The strategic-form representation of the self-reporting subgame is presented in Figure 1.[21]

To minimize subjects' computational costs, and given that the purpose of this study is to

[16]Multipliers $(r_1, r_2) = (1, 1)$ imply that the enforcement policy does not grant leniency for self-reporting.

[17]In theory, $t = 0$ represents immediate reporting and $t > 0$ represents delayed reporting.

[18]Each injurer gets an expected payoff equal to $b - p_0 f$.

[19]The injurer who self-reports the act gets a payoff equal to $b - r_1 f$, and the silent co-conspirator gets an expected payoff equal to $b - p_1 f$.

[20]If both injurers self-report the act, then they are equally likely to get the first and second position in the self-reporting queue. Hence, the expected payoff for each injurer is equal to $b - (\frac{r_1 + r_2}{2}) f$.

[21]This figure, which replicates Landeo and Spier's (2018) Figure 1, describes the components that determine the injurers' expected payoffs.

Figure 1: Strategic-Form Representation of the Self-Reporting Subgame (Expected Payoffs)

	No Report (NR)	Report (R)
No Report (NR)	$b - p_0 f$, $b - p_0 f$	$b - p_1 f$, $b - r_1 f$
Report (R)	$b - r_1 f$, $b - p_1 f$	$b - \left(\frac{r_1+r_2}{2}\right) f$, $b - \left(\frac{r_1+r_2}{2}\right) f$

assess the determinants of self-reporting, time to report, detection and deterrence, our experimental design focuses on the potential injurers' decisions to participate in the act and the injurers' decisions to self-report. We assign particular numerical values to the model parameters.[22] Across leniency environments, the parameter values are as follows: $b \in [200, 1600]$; $f = 900$; $p_0 = 0.4$ and $p_1 = 0.9$; and, $t \in [0, 90]$ (measured in seconds). When ordered leniency for self-reporting is granted, the leniency multipliers are: $r_1^S = 0.333$ and $r_2^S = 1$, for Strong Leniency; and, $r_1^M = 0.466$ and $r_2^M = 1$, for Mild Leniency. The leniency multipliers under Strong and Mild Leniency generate maximal detection and deterrence when the Pareto-dominance and risk-dominance refinements are applied, respectively.[23] When leniency for self-reporting is not granted, the leniency multipliers are $r_1^N = r_2^N = 1$.

2.2 Equilibrium Analysis

The equilibrium concept is subgame-perfect Nash equilibrium. We focus on pure-strategy equilibria. We apply backward induction and start by analyzing the injurers' self-reporting decisions in Stage 2. Then, we analyze the potential injurers' decisions regarding participating in the act in Stage 1.

2.2.1 Stage 2: Decision to Report the Act and Time to Report

The strategic-form representation of the self-reporting subgame for the Strong Leniency (S), Mild Leniency (M), and No Leniency (N) environments under the chosen numerical values for the model parameters is presented in Figure 2.

[22]Our numerical example satisfies satisfies all the model's assumptions and, therefore, the predictions derived from these assumptions hold.

[23]The construction of the leniency multipliers follows Landeo and Spier (2018): By Landeo and Spier's (2018) Proposition 3 (Case 2, $p_1 > \frac{1+p_0}{2}$), the leniency multipliers that generate maximal detection and deterrence for these parameter values, i.e., optimal ordered leniency, are $(r_1^S, r_2^S) = (0.400, 1)$ and $(r_1^M, r_2^M) = (0.533, 1)$. To break indifference, we deduct $\varepsilon = 0.067$ from r_1^S and r_1^M.

Figure 2: Strategic-Form Representation of the Self-Reporting Subgame
(Expected Payoffs per Leniency Environment)

Strong Leniency (S)

	NR	R
NR	$b-360$, $b-360$	$b-810$, $b-300$
R	$b-300$, $b-810$	$b-600$, $b-600$

Mild Leniency (M)

	NR	R
NR	$b-360$, $b-360$	$b-810$, $b-420$
R	$b-420$, $b-810$	$b-660$, $b-660$

No Leniency (N)

	NR	R
NR	$b-360$, $b-360$	$b-810$, $b-900$
R	$b-900$, $b-810$	$b-900$, $b-900$

When Strong Leniency is implemented, the injurers confront a "prisoners' dilemma" game, where the unique Nash equilibrium is the less-preferred outcome where both injurers self-report, i.e., the (R, R) outcome. When Mild Leniency is implemented, the injurers confront a "coordination game" where the two Nash equilibria are (NR, NR) and (R, R). The Pareto-dominant Nash equilibrium is (NR, NR) and the risk-dominant Nash equilibrium is (R, R). Then, self-reporting by both injurers occurs only when both injurers fail to coordinate on their preferred outcome. Given that only the first injurer to report receives leniency under Strong Leniency and Mild Leniency, both injurers have an incentive to minimize the time to report in order to secure the first position in the self-reporting queue.[24] In other words, ordered-leniency policies exhibit a race-to-the-courthouse effect. As a result, each injurer will report the act immediately in equilibrium, $t = 0$. When leniency for self-reporting is not granted (No Leniency), the unique Nash equilibrium involves no reporting by both injurers, i.e., the (NR, NR) outcome.

Proposition 1 characterizes the pure-strategy Nash equilibria of the self-reporting subgame for the three leniency environments.[25]

[24] We assume that in case of indifference, an injurer decides to self-report immediately.

[25] For general versions of the content of this proposition and formal proofs, see Landeo and Spier's (2018) Lemmas 1 and 2.

Proposition 1. *The pure-strategy Nash equilibria of the self-reporting subgame are as follows.*

1. *Suppose Strong Leniency for self-reporting is granted. There is a unique pure-strategy Nash equilibrium where both injurers self-report immediately, (R, R).*
2. *Suppose Mild Leniency for self-reporting is granted. There are two pure-strategy Nash equilibria, one where both injurers self-report immediately, and one where neither injurer self-reports. (R, R) Pareto dominates (NR, NR) and (R, R) risk dominates (NR, NR).*
3. *Suppose No Leniency for self-reporting is granted. There is a unique pure-strategy Nash equilibrium where neither injurer self-reports, (NR, NR).*

Detection Rate

In equilibrium, the enforcement agency achieves higher detection rates with ordered leniency. Specifically, when Strong Leniency is granted, or Mild Leniency is granted and the risk-dominance refinement is applied, both injurers self-report in equilibrium. Hence, the enforcement agency will detect an injurer with certainty. When leniency for self-reporting is not granted, or when Mild Leniency for self-reporting is granted and the Pareto-dominance refinement is applied, neither injurer self-reports in equilibrium. Hence, the enforcement agency will detect an injurer with a likelihood equal to $p_0 = .40$. Corollary 1 summarizes the equilibrium detection rate for the three leniency environments.

Corollary 1. *Across b-values, an injurer is detected with* 40% *likelihood under No Leniency or Mild Leniency with the Pareto-dominance refinement. Across b-values, an injurer is always detected under Strong Leniency or Mild Leniency with the risk-dominance refinement.*

Expected Fine

As discussed above, ordered-leniency policies increase the likelihood that harmful group activities are detected. Although ordered leniency involves fine discounts, a potential injurer will confront a higher expected fine when Strong Leniency or Mild Leniency (with the risk-dominance refinement) are implemented. In particular, when Strong Leniency or Mild Leniency is granted and the risk-dominance refinement is applied, a potential injurer confronts expected fines equal to $\left(\frac{r_1^S+r_2^S}{2}\right)f = 600$ and $\left(\frac{r_1^M+r_2^M}{2}\right)f = 660$ in equilibrium, for Strong and Mild Leniency, respectively. When leniency for self-reporting is not granted or when Mild Leniency for self-reporting is granted and the Pareto-dominance refinement is applied, a potential injurer confronts a expected fine equal to $p_0 f = 360$ in equilibrium. Hence, the highest expected fine corresponds to Mild Leniency when the risk-dominance refinement is applied. Corollary 2 summarizes the equilibrium expected fine for the three leniency environments.

Corollary 2. *Across b-values, the expected fine under Mild Leniency with the risk-dominance refinement is equal to* 660. *The expected fine under Strong Leniency is equal to* 600. *The expected fine under No Leniency or Mild Leniency with the Pareto-dominance refinement is equal to* 360.

2.2.2 Stage 1: Decision to Participate in the Act

Given the equilibrium expected fines, the highest individual benefits to induce an injurer not to participate in the act, i.e, the deterrence thresholds $\hat{b}^i$ $(i = N, S, M)$, are as follows:[26] $\hat{b}^N = \hat{b}^M = p_0 f = 360$, for the No Leniency environment and the Mild Leniency environment when the Pareto-dominance refinement is applied; $\hat{b}^S = \left(\frac{r_1^S + r_2^S}{2}\right) f = 600$, for the Strong Leniency environment; and, $\hat{b}^M = \left(\frac{r_1^M + r_2^M}{2}\right) f = 660$ for the Mild Leniency environment when risk-dominance is applied. Proposition 2 characterizes the equilibrium decisions in Stage 1 for the three leniency environments.[27]

Proposition 2. *Each potential injurer will decide to participate in the act under the following conditions.*

1. *Suppose Strong Leniency for self-reporting is granted. The potential injurer decides to participate if and only if* $b > \hat{b}^S = 600$.

2. *Suppose Mild Leniency for self-reporting is granted. If the Pareto-dominance refinement is applied, then the potential injurer decides to participate if and only if* $b > \hat{b}^M = 360$. *If the risk-dominance refinement is applied, then the potential injurer decides to participate if and only if* $b > \hat{b}^M = 660$.

3. *Suppose No Leniency for self-reporting is granted. The potential injurer decides to participate if and only if* $b > \hat{b}^N = 360$.

[26] We assume that in case of indifference, a potential injurer decides not to commit the act.

[27] For a general version of this proposition and a formal proof, see Landeo and Spier (2018), Lemma 3.

Table 1 – Theoretical Point Predictions

Leniency Environment	Report		Detection	Expected Fine[d]	Deterrence Rate	
	Rate[a]	Time[b]	Rate[c]	($\hat{b}$)	$b \le \hat{b}$	$b > \hat{b}$
Strong Leniency (S)	1	0	1	600	1	0
Mild Leniency (M)						
• Risk Dominance	1	0	1	660	1	0
• Pareto Dominance	0	–	.40	360	1	0
No Leniency (N)	0	–	.40	360	1	0

Notes: [a]Report rate conditional on committing the act; [b]report time (in seconds) conditional on committing the act and reporting; [c]detection rate conditional on committing the act; [d]expected fine conditional on committing the act.

Deterrence Rate

Remember that a harmful (group) act is committed only if *both* potential injurers agree to participate in the act. Hence, across leniency environments, a harmful act will be deterred with certainty when the individual benefit from committing the act is not greater than the deterrence threshold, i.e., when $b \le \hat{b}^i$ $(i = N, S, M)$. Corollary 3 summarizes the equilibrium deterrence rate for the three leniency environments.

Corollary 3. *A harmful act is always deterred when* $b \le \hat{b}^i$ $(i = N, S, M)$ *and never deterred when* $b > \hat{b}^i$ $(i = N, S, M)$*.*

Table 1 outlines the theoretical point predictions (individual level).[28]

3 Experimental Design

In assessing the validity of the qualitative predictions derived from the theory, our study analyzes the effects of ordered-leniency policies on self-reporting, detection, individual fines, and deterrence, and investigates whether ordered-leniency policies exhibit a race-to-the-courthouse

[28]In theory, the rate of deterrence of individuals from participating in harmful group activities and the rate of deterrence of harmful group activities are the same. In the lab, however, they might be different. Our statistical analysis will be focused on the deterrence of harmful group activities, i.e., on the percentage of groups where one or both potential injurers decided not to participate in the activity. See Section 5.2.5.

effect. The experimental design consists of three leniency environments: Strong Leniency (S), where the fine reduction for self-reporting is large; Mild Leniency (M), where the fine reduction for self-reporting is small; and, No Leniency (N), where a fine reduction for self-reporting is not granted.

Next, we present a description of the laboratory implementation of the theoretical environments.

3.1 The Games

Procedural regularity is accomplished by developing a software program that allows the subjects to play the game by using networked personal computers. The software, constructed using the Java programming language, consists of 3 versions of the game, reflecting the three experimental conditions: Strong Leniency, Mild Leniency, and No Leniency.[29]

To ensure control and replicability, a free-context environment is implemented. Specifically, neutral labels are used to denote the subjects' roles: Players B1 and B2 (potential injurers 1 and 2, respectively). The "Act" is described as an economic decision involving potential benefits (associated with Stage 1) and potential losses (associated with Stage 2).[30] The players' choices are also labeled in a neutral way: Decision whether "To Agree to Jointly Commit the Act" or "Not to Agree to Jointly Commit the Act;" and, decision whether "To Report the Act" or "Not to Report the Act." The game includes 5 practice matches and one actual match. The practice matches allow the subjects to experiment with the different options and hence, learn about the consequences of their choices. Only the actual match is considered in the subject's payment.

The benchmark game corresponds to the Strong Leniency condition (S). Subjects play the role of Player B1 or Player B2. The roles of Players B1 and B2 are similar. Each match involves two stages. In Stage 1, each player independently decides whether to participate in the act. The players have 90 seconds to make their decisions in Stage 1. After the decisions are made, both players are informed about the other player's decision. If both players agree to commit the act, then Stage 2 starts. Otherwise, the game ends.

[29]The use of a JAVA software especially designed for this study allows us to have flexibility in the design of randomization processes and the design of user-friendly screens. See Supplementary Material for a Sample of Software Screens. A complete set of software screens is available from the authors upon request.

[30]See the Appendix for a sample of the instructions (Mild Leniency condition). See Supplementary Material for a complete set of instructions for the three experimental conditions.

In Stage 2, each player independently decides whether to report or not report the act. Each player has 90 seconds to decide whether to report or not report the act and submit his or her chosen action.[31] When both players decide to report at the same time, the computer randomly assigns the first position in the self-reporting queue to each player with equal probability. After the decisions are made, both players are informed about the decision of the other player and the payoffs for both players, and the game ends. The payoffs reflect the Strong Leniency policy: The first player to self-report receives a fine reduction. It is worth noting that our experimental design also allows us to collect data on the time to submit the chosen action. These data are used to assess whether ordered-leniency policies exhibit a race-to-the-courthouse effect.[32]

Variations of the benchmark game satisfy the other experimental conditions. In the Mild Leniency condition (M) and No Leniency condition (N), the subjects play a similar game. The only difference across conditions refers to the fine reduction granted to the first player to self-report. Specifically, in the Mild Leniency condition (M), the first player to self-report receives a fine reduction which is lower than the one granted in case of the Strong Leniency condition (S). In the No Leniency condition (N), the first player to self-report does not receive a fine reduction.

Each experimental condition includes four 24-subject sessions. To achieve *independent observations in the actual match,* we use the following role and pairing procedure per session: (1) The total number of subjects are anonymously and randomly assigned to one of the following two groups, Group 1 and Group 2; (2) half of the subjects in each group is assigned the role of Player B1 and the other half is assigned the role of Player B2; (3) for each practice match, Players B1 from Group 1 are anonymously and randomly paired with Players B2 from Group 2, and Players B2 from Group 1 are anonymously and randomly paired with Players B1 from Group 2; (4) for the actual match, Players B1 and B2 from Group 1 are anonymously and randomly paired, and Players B1 and B2 from Group 2 are anonymously and randomly paired. The same protocol for pair formation is applied across sessions and conditions. As a result, for each session of 24 subjects, 12 independent observations (pairs) are obtained, across sessions and conditions. Hence, 48 independent observations (pairs) per condition and 144 independent

[31] To control for any possible differences in connectivity time between the lab server and the terminals, we decided to place all subjects that submitted their chosen actions within the same second in the same "time interval." For instance, if two subjects submitted their actions between the first and second seconds, i.e., $t \in [1, 2)$ seconds, then the recorded "time interval" will be the same for both subjects.

[32] See Section 4.2 for details.

Table 2 – Experimental Conditions

b-Value Segments	Strong Leniency (S)	Mild Leniency (M)	No Leniency (N)
$b \in [200, 360]$	8	8	8
$b \in (360, 600]$	22	22	22
$b \in (600, 660]$	22	22	22
$b \in (660, 1600]$	44	44	44
Total Number of Individuals	96	96	96

observations (pairs) in total are obtained. Given our experimental design, the observations within a pair are also independent.[33] Hence, 288 independent individual observations in total are obtained.

Table 2 summarizes the information regarding the experimental conditions and observations per b-value segment for the actual match. The theoretical deterrence thresholds guide the design of the distribution of b-values. To ensure *comparability across conditions*, we randomly predetermine the b-values used in the actual match of each of the four sessions of a condition, and apply these values to each condition.[34] For each condition, the total number of b-values for the actual match is equal to 48 (12 values per session; 4 sessions per condition).[35] To ensure *consistency across sessions and conditions*, we randomly predetermine the b-values for each of the five practice matches, and apply these values across sessions and conditions.[36]

[33] Two features of the experimental design are relevant. (1) Pairs are anonymously formed, across sessions and conditions. (2) When an act is committed by a pair, the only information that each pair member receives about the other pair member at the end of Stage 1 is that the other pair member also decided to participate in the act. The information is provided in a general and structured format: "You agreed to commit the act and Player B2 agreed to commit the act." (See the Sample Software Screens in the Supplementary Materials.) The exact same information is provided to the subjects across pairs, sessions and conditions. As a result, the specific pair where a subject is assigned does not affect the decisions of that subject, and hence, the observations within a pair are independent.

[34] The chosen distribution of b-values has the following features: Four b-value segments are considered, $[200, 360]$, $(360, 600]$, $(600, 660]$, and $(660, 1600]$; the segments include 8, 23, 23 and 46% of the total b-values, respectively; for each segment, the specific b-values are randomly chosen (equally likely values).

[35] The adopted distribution of b-values allows us: (1) To collect a sufficiently high number of observations to perform statistical analysis of deterrence across conditions, and across relevant b-value segments within each condition; and, (2) to collect a sufficiently high number of observations in which Stage 2 occurs with certainty in equilibrium across conditions ($b > 660$) to perform statistical analysis of detection across conditions.

[36] For each practice match, at least one b-value pertains to each of the four b-value segments; and, the majority of b-values pertain to the last b-value segment, which in theory, elicits a self-reporting stage with

3.2 The Experimental Sessions

We ran twelve 80-minute sessions of 24 subjects each (four sessions per condition; 96 subjects per condition; 288 subjects in total) at Harvard University.[37] Each session was conducted by two research assistants at the Harvard Decision Science Laboratory.[38] Subjects were recruited using the lab's Sona computer program and the lab's subject pool. Subjects were allowed to participate in one experimental session only, and received information only about the game version that they were assigned to play. The participant pool included undergraduate and graduate students from Harvard University, Boston University and Northeastern University, from a wide variety of fields of study. A laboratory currency called the "token" (29 tokens = 1 U.S. dollar) was used in our experiment. To avoid negative payoffs, each subject received an initial endowment equal to 700 tokens.[39] The show-up fee was equal to \$10. The average game earnings was equal to \$32. Hence, the average total payment was equal to \$42 (average game earnings plus participation fee) for an 80-minute session.

At the beginning of each session, written instructions were provided to the subjects (see the Appendix). The instructions about the game and the software were verbally presented by the experimenter to create common knowledge. Specifically, subjects were informed: (1) about the game structure, possible choices, and payoffs; (2) about the random process of allocating roles; (3) about the randomness and anonymity of the process of forming pairs;[40] (4) about the token/dollar equivalence, and that they would receive the dollar equivalent of the tokens they held at the end of the session. Finally, subjects were asked to complete a set of exercises to ensure their ability to read the information tables. The answers to the exercises were read aloud by the research assistants. Questions about the written instruction and questions about the exercises were answered by the research assistants privately and before the beginning of the practice matches. The rest of the session was entirely played using computer terminals and the software designed for this experiment. After the actual match, subjects were required to fill

certainty. Hence, the distribution of b-values ensures that the subjects will get enough experience regarding the self-reporting stage.

[37]See Supplementary Material for detailed description of the procedures followed to implement the experimental sessions and detailed information about the subjects used in this study.

[38]See Supplementary Material for a description of the procedure followed to train the research assistants.

[39]Note that the minimum possible b-value was equal to 200 tokens and the maximum possible fine was equal to 900 tokens. Then, the minimum possible match payoff was equal to 0 tokens.

[40]In particular, subjects were informed that they would not play with the same partner in any practice match; and, that they would not play with any of their previous partners in the actual match.

out a short questionnaire with general demographic questions. At the end of each experimental session, subjects privately received their monetary payoffs in cash.

4 Qualitative Hypotheses

The qualitative hypotheses are presented below. Cooper et al. (1990) suggest that risk-dominance is generally the equilibrium selection criterion chosen by subjects in the lab when there are multiple equilibria.[41] Then, we might expect that the majority of subjects will apply the risk-dominance refinement in our experiment. Hence, the hypotheses related to Mild Leniency are constructed under the risk-dominance refinement.

4.1 Report Rate

Hypothesis 1. *Strong and Mild Leniency increase the report rate, with respect to No Leniency.*

In theory, Strong Leniency and Mild Leniency (when the risk-dominance refinement is applied) induce both injurers to self-report, i.e., (R, R) is the unique N.E. of the reporting subgame. In contrast, No Leniency induces both injurers to no-report, i.e., (NR, NR) is the unique N.E. of the reporting subgame. Therefore, the implementation of ordered leniency raises the report rate. When the Pareto-dominance refinement is applied instead, Mild Leniency induces both injurers to no-report, i.e., (NR, NR) is chosen by both injurers. Then, Mild Leniency and No Leniency exhibit the same zero report rate.

Hypothesis 2. *Strong and Mild Leniency exhibit the same report rate.*

As mentioned, Strong Leniency and Mild Leniency (when the risk-dominance refinement is applied) induce both injurers to self-report. Then, both leniency policies exhibit the same 100% report rate, and the same zero no-report rate. When the Pareto-dominance refinement is applied instead, Mild Leniency induces both injurers to no-report. Then, Mild Leniency increases the no-report rate (100% v. zero, for Mild and Strong Leniency, respectively).

[41] Landeo and Spier (2009) offer important evidence in favor of the risk-dominance refinement in contractual settings with multiple equilibria. Burton and Sefton (2004) provide powerful evidence of the role of riskiness in the choice of a strategy. See Ochs (1995) for a survey of seminal work on coordination games.

4.2 Race-to-the-Courthouse Effect

Hypothesis 3. *Strong and Mild Leniency (weakly) reduce the time to submit the chosen action, with respect to No Leniency. There is a race-to-the-courthouse effect.*

In theory, Strong Leniency and Mild Leniency (when the risk-dominance refinement is applied) incentivize the injurers to minimize their reporting times to increase their chances to get the first position in the self-reporting queue. In other words, the *incentives* provided by ordered-leniency policies generate a race-to-the-courthouse effect. As a result, all injurers self-report immediately, $t = 0$. In the lab, subjects have between zero and 90 seconds to decide an action (R or NR) and to submit the chosen action. Then, a prompt submission of the chosen action will be represented by a time $t = 0 + \varepsilon$ ($\varepsilon > 0$, small number).[42]

We assess the race-to-the-courthouse effect of ordered-leniency policies by comparing the times to submit the chosen action under Strong and Mild Leniency with the time to submit the chosen action under No Leniency. In theory, Strong Leniency and Mild Leniency (when the risk-dominance refinement is applied) induce both injurers to self-report, i.e., (R, R) is the unique N.E. of the reporting subgame. In other words, the chosen action by both injurers is R in equilibrium. In contrast, No Leniency induces both injurers to no-report, i.e., (NR, NR) is the unique N.E. of the reporting subgame, i.e., the chosen action by both injurers is NR in equilibrium. When the Pareto-dominance refinement is applied instead, Mild Leniency induces both injurers to no-report, i.e., (NR, NR) is chosen by both injurers, i.e., the chosen action is NR in equilibrium.

In theory, given that an early submission of the NR action does not benefit the injurers (i.e., an early submission of the NR action does not involve any penalty reductions), the time to submit the chosen action under No Leniency and Mild Leniency (when the Pareto-dominance refinement is applied) will be $t \in [0 + \epsilon, 90]$ ($\varepsilon > 0$) in equilibrium. In contrast, the time to submit the R action under Strong and Mild Leniency (when the risk-dominance refinement is

[42]The value of ε is, of course, unknown. Then, a theoretical point prediction of "prompt" submission of the chosen action under Strong and Mild Leniency cannot be constructed. Hence, statistical analyses to assess the validity of theoretical point predictions are not possible. Alternatively, we could directly elicit the time to report by adding a question about the preferred time to report to the "report" choice. To avoid influencing subjects' choices, a similar type of question would need to be added to the "no report" choice. To ensure the truthful revelation of information, an additional payment mechanism would be required. Hence, this alternative would add unnecessary complexity to the experimental environment, and hence, noise to the collected data.

applied) will be $t = 0 + \varepsilon$ ($\varepsilon > 0$) in equilibrium.[43] Hence, Strong Leniency or Mild Leniency when the risk-dominance refinement is applied (weakly) decrease the time to submit the chosen action, with respect to No Leniency.[44] In other words, ordered-leniency policies exhibit a race-to-the-courthouse effect.

4.3 Detection Rate

Hypothesis 4. *Strong and Mild Leniency increase the detection rate, with respect to No Leniency.*

In theory, Strong Leniency and Mild Leniency (when the risk-dominance refinement applies) induce both injurers to self-report after committing the act. As a result, an injurer is detected with certainty in equilibrium. In contrast, with No Leniency, neither injurer self-reports. Then, an injurer is detected with only 40% chance (p_0). Hence, the implementation of ordered-leniency policies raises the detection rate. When the Pareto-dominance refinement is applied instead, neither injurer self-reports under Mild Leniency. So both Mild and No Leniency exhibit the same low 40% detection rate.

Hypothesis 5. *Strong and Mild Leniency exhibit the same detection rate.*

As mentioned above, Strong Leniency and Mild Leniency (when the risk-dominance refinement is applied) induce both injurers to self-report after committing the act. Therefore, an injurer is detected with certainty under both leniency policies. When the Pareto-dominance refinement is applied instead, neither injurers self-reports under Mild Leniency. As a result, the likelihood of detection is 40% only. Hence, Strong Leniency will increase the detection rate.

[43] The logic of the proof of Landeo and Spier's (2018) Lemma 1 for $r_1 = r_2 = 1$ and $r_1 < r_2 = 1$, respectively, applies here.

[44] In addition to assuming that the risk-dominance refinement applies and that the equilibrium actions are chosen, a behavioral assumption is implicitly used in the construction of Hypothesis 3. We assume that the decision-making processes associated with the choices of actions R and NR involve the same amount of time. Remember that the choice of action R and the choice of action NR involve the review of the *same* information. Then, this assumption is appropriate.

4.4 Expected Fine

Hypothesis 6. *Mild Leniency increases the expected fine, with respect to Strong Leniency and No Leniency. Strong Leniency increases the expected fine, with respect to No Leniency.*

In theory, given the detection rates, the implementation of ordered-leniency policies raises the expected fine. The expected fines are equal to 660 for Mild Leniency (when the risk-dominance refinement is applied), 600 for Strong Leniency, and 360 for No Leniency and Mild Leniency (when the Pareto-dominance refinement is applied).

4.5 Deterrence Rate

Hypothesis 7. *Within each leniency environment, the deterrence rate is lower when the benefit from the harmful act is greater than the deterrence threshold.*

According to the theoretical predictions, the benefit associated with the harmful act incentivizes each potential injurer to participate in the act. In equilibrium, the act is committed only when the benefit is greater than the deterrence threshold (expected fine). In other words, when $b > \hat{b}^i$ $(i = N, S, M)$, the deterrence rate is lower.

Hypothesis 8. *Mild Leniency increases the deterrence rate, with respect to Strong Leniency and No Leniency. Strong Leniency increases the deterrence rate, with respect to No Leniency.*

In theory, a harmful act is committed only when $b > \hat{b}^i$ $(i = N, S, M)$. In equilibrium, the deterrence threshold for Mild Leniency (when the risk-dominance refinement is applied) is higher than the expected fine for Strong Leniency and No Leniency (660 v. 360, 660 v. 600, respectively); and, the expected fine for Strong Leniency is also higher than the expected fine for No Leniency (600 v. 360). As a result, Mild Leniency exhibits the highest deterrence rate and No Leniency exhibits the lowest deterrence rate. When the Pareto-dominance refinement is applied instead, the expected fine for Mild and No Leniency will be the same (360). Then, Mild and No Leniency exhibit the same deterrence rates, and the highest deterrence rate is achieved with Strong Leniency. Across b-values, in theory, the deterrence rates are 31, 54, and 8%, for Strong Leniency, Mild Leniency, and No Leniency, respectively.[45]

[45] Given the number of observations per b-value segment (see Table 2), the deterrence rates are computed as follows: 4/48, 15/48 and 26/48, for Strong Leniency, Mild Leniency, and No Leniency, respectively.

It is worth noting that, in theory, Strong and Mild Leniency (when the risk-dominance refinement is applied) have the property of incentivizing both injurers to be the first to report. As a result, both injurers will report immediately, and hence, they will be equally likely to get the first position in the self-reporting queue (i.e., the chance of each injurer to get the first position will be equal to 50%).[46] In the lab, however, some subjects might exhibit cognitive biases, such as self-serving bias (Babcock et al., 1995),[47] and hence believe that their chances to be the first to report are greater than the chances of the other injurers (i.e., they might believe that their chances to get the first position in the self-reporting queue are greater than 50%). In the limiting case, some subjects might believe that they will always be the first to report. Then, under Strong Leniency, they might consider a fine equal to 300 instead of an expected fine equal to 600 when making their decision about committing the act.[48] Similarly, under Mild Leniency, some subjects might consider a fine equal to 420 instead of an expected fine equal to 660 when making their decision about committing the act.[49] As a result, we might observe deterrence rates lower than those predicted by the theory.[50]

[46] Remember that, in the Strong Leniency environment, the expected fine under (R, R) and equal likelihoods of getting the first or second position in the self-reporting queue is equal to .50(300) + .50(900) = 600; and, in the Mild Leniency environment, the expected fine under (R, R) and equal likelihoods of getting the first or second position in the self-reporting queue is equal to .50(420) + .50(900) = 660.

[47] Self-serving bias is attributed to motivated reasoning, i.e., a propensity to reason in a way that supports the individual's subjectively favored beliefs by attending only to some available information (Kunda, 1990, 1987). See Babcock et al. (1995) and Landeo (2009) for experimental evidence of self-serving bias, and Landeo et al. (2013) for theoretical work on self-serving bias in incomplete-information environments. See also Landeo (2018) for further discussion of theoretical and experimental studies on self-serving bias.

[48] The expected fine for the biased injurer is 300 irrespective of the reporting decision of the other injurer. Importantly, under these biased beliefs, report is the dominant strategy for the biased injurer. This is because his expected fines under no-report are 360 or 810, depending on the choice of no-report or report by the other injurer, respectively. Then, as in the case of the environment with unbiased injurers, the biased injurer will choose to report.

[49] The expected fine for the biased injurer is 420 irrespective of the reporting decision of the other injurer. Importantly, under these biased beliefs, report will be the best response when the other injurer chooses to report ($420 < 810$) and no-report will be the best response when the other injurer chooses no-report ($360 < 420$). Then, as in the case of the environment with unbiased injurers, the biased injurer might choose to report or no-report, according to his beliefs about the strategy that the other injurer will choose and his assessment of the riskiness of each strategy.

[50] Nagin and Pogarsky (2003) provide experimental evidence of the association between individual's perception of the risk of participating in harmful activities, self-serving bias, and participation in harmful activities. See Chalfin and McCrary (2017) for a survey on the literature on crime and deterrence. See Section 6 for

Table 3 – Descriptive Statistics[a]

Condition	Report	Time to Submit Chosen Action[b]			Detection	Individual	Deterrence Rate	
	Rate	R	NR	R or NR	Rate	Fine[c]	$b \leq \hat{b}$	$b > \hat{b}$
Strong Leniency (S)	.90	2.14	15.25	3.49	.99	588.46	.47	.06
[96]		(1.30)	(19.62)	(7.25)		(307.47)		
Mild Leniency (M)	.76	2.71	12.29	4.99	.96	638.25	.31	.00
[96]		(2.70)	(17.28)	(9.51)		(268.71)		
No Leniency (N)	.06	16.75	16.21	16.24	.53	475.71	.75	.23
[96]		(22.23)	(17.67)	(17.77)		(452.51)		

Notes: [a]Total number of individuals in brackets; standard deviation in parentheses; report and detection rates conditional on committing the act; [b]mean time to submit chosen action (in seconds) conditional on committing the act; [c]mean individual fine conditional on committing the act.

5 Results

Given our experimental design, the collected observations are independent. Then, it is appropriate to use non-parametric statistical tests. Specifically, our analysis involves the use of the Fisher's exact test and the Wilcoxon rank-sum (Mann Whitney) test.

5.1 Data Summary

Table 3 provides descriptive statistics (individual-level data) for the report rate, report time, detection rate, fine, and deterrence rate (for b-values below and above the deterrence threshold $\hat{b}$. The time to submit the chosen action is defined as the average individual time to submit the chosen action (R or NR) in seconds, conditional on committing the act.[51] The report rate is defined as the percentage of individuals who decided to report the act, conditional on committing the act. The detection rate is defined as the percentage of individuals who were detected, conditional on committing the act. The individual fine corresponds to the average individual fine, conditional on committing the act. The deterrence rate when $b \leq \hat{b}$ (or $b > \hat{b}$)

further discussion.

[51]As mentioned in Section 3.1 (footnote 30), to control for any possible differences in connectivity time between the lab server and the terminals, we decided to place all subjects that submitted their chosen actions within the same second in the same "time interval." For instance, if two subjects submitted their chosen actions between the first and second seconds, i.e., $t \in [1, 2)$ seconds, then the recorded "time interval" to submit the chosen action will be the same for these two subjects. We use the midpoint of each time interval in the statistical analysis of results.

is defined as the percentage of groups with $b \leq \hat{b}$ (or $b > \hat{b}$) where one or both potential injurers decided not to participate in the act (i.e., the percentage of harmful acts that are deterred).

Overall, our results are aligned with the theoretical predictions under the risk-dominance refinement. Specifically, the data indicate that the implementation of ordered-leniency policies increased self-reporting (90 v. 6%, Strong Leniency and No Leniency, respectively; 76 v. 6%, Mild Leniency and No Leniency, respectively). The average time to submit the chosen action (R or NR) for Strong, Mild, and No Leniency were equal to 3.49, 4.99, and 16.24 seconds, respectively. In particular, the average times to submit the R action under Strong and Mild Leniency were equal to 2.14 and 2.71 seconds, respectively; and the average time to submit the NR action under No Leniency was equal to 16.21 seconds. These findings suggest that ordered-leniency policies incentivize the injurers to self-report promptly to increase their chances to get the first position in the self-reporting queue. In other words, the data indicate the presence of a race-to-the-courthouse effect under ordered-leniency policies.

Importantly, the data suggest that the implementation of ordered-leniency policies increased the rate of detection of harmful acts (99 v. 53%, Strong Leniency and No Leniency, respectively; 96 v. 53%, Mild Leniency and No Leniency, respectively). The high detection rates under Strong and Mild Leniency are explained by the relatively high self-reporting rates in these environments. Remember that the likelihood of detection of a silent co-conspirator increases when the other injurer self-reports (from 40 to 90%), and that detection occurs with certainty when both injurers self-report the act.

Our findings also indicate that Mild Leniency increased the average individual fine, with respect to Strong Leniency and No Leniency; and, Strong Leniency increased the average individual fine, with respect to No Leniency. As predicted when the risk-dominance refinement is applied, the highest and lowest average individual fines were obtained under Mild Leniency and No Leniency, respectively. These results might be explained by the higher detection rates under ordered-leniency policies.

Finally, the data suggest that the deterrence rates under Strong Leniency, Mild Leniency, and No Leniency for $b \leq \hat{b}^i$ ($i = S, M$) were lower than 100% (the theoretical predictions). These results might indicate that some injurers considered a different deterrence threshold when making their decisions about participating in the act, due for instance to self-serving bias.

Table 4 – Effect of Ordered Leniency on Self-Reporting[a]

Conditions	Number of Individuals[b]	Report Rate[c]	p-value[d]
Strong Leniency v. No Leniency	78, 70	.90 v. .06	$p < .001$
Mild Leniency v. No Leniency	80, 70	.76 v. .06	$p < .001$
Mild Leniency v. Strong Leniency	80, 78	.76 v. .90	$p = .020$

Notes: [a]"v." denotes "versus;" [b]number of individuals who committed the act; [c]report rate conditional on committing the act; [d]p-value corresponds to one-sided Fisher's exact test.

5.2 Analysis

The main findings will be presented in a series of results.

5.2.1 Report Rate

We start our analysis by studying the effects of the implementation of ordered-leniency policies on the injurers' decisions to self-report. Table 4 summarizes the results.[52] Consider first the effect of the implementation of a Strong Leniency policy on self-reporting. In theory, self-reporting by both injurers is the unique N.E. under Strong Leniency and no-reporting by both injurers is the unique N.E. under No Leniency. Then, the report rate under Strong Leniency is 100% and the report rate under No Leniency is zero. Our results indicate that Strong Leniency significantly increases the report rate (90 v. 6%, for Strong Leniency and No Leniency, respectively; $p < .001$). Our results are aligned with our theoretical qualitative predictions and provide strong support for Hypothesis 1.[53]

Result 1. *Strong Leniency significantly increases the report rate, with respect to No Leniency.*

Consider now the effect of implementing a Mild Leniency policy on self-reporting. In theory, when the risk-dominance refinement is applied, Mild Leniency increases the report rate, with

[52]See Supplementary Material for pair-level statistical analysis.

[53]Regarding the behavior of pairs (pair-level data), our findings suggest that the (R, R) rate is significantly higher under Strong Leniency (.79 v. .00, for Strong and No Leniency, respectively, $p < .001$, one-sided Fisher's exact test), and the (NR, NR) rate is significantly higher under No Leniency (.89 v. .00, for No Leniency and Strong Leniency, respectively; $p < .001$, one-sided Fisher's exact test). These results are consistent across b-values. See Supplementary Material for details.

respect to No Leniency (100% v. zero, for Mild and No Leniency, respectively). If the Pareto-dominance refinement is applied instead, Mild Leniency and No Leniency exhibit the same zero report rate. Our results indicate that Mild Leniency significantly increases the report rate (76 v. 6%, for Mild Leniency and No Leniency, respectively; $p < .001$). These findings are aligned with our theoretical predictions under the risk-dominance refinement. Importantly, our results suggest that the enforcement agency can elicit injurers' cooperation without relying on large discounts for self-reporting (Strong Leniency), i.e., by implementing a coordination-game environment through Mild Leniency. Our results provide strong support for Hypothesis 1.[54]

Result 2. *Mild Leniency significantly increases the rate of self-reporting, with respect to No Leniency.*

Finally, consider the effect of implementing Mild and Strong Leniency on self-reporting. In theory, when the risk-dominance refinement is applied, both injurers self-report in equilibrium. Then, Strong and Mild Leniency exhibit the same 100% report rates. If the Pareto-dominance refinement is applied instead, neither injurer self-report in equilibrium under Mild Leniency. Then, Strong Leniency increases the report rate (100% v. zero, for Strong and Mild Leniency, respectively). Our results indicate that, although Mild Leniency also exhibits a relatively high report rate, Strong Leniency significantly increases the report rate (90 v. 76%, for Strong and Mild Leniency, respectively $p = .020$). These findings might be explained by the behavior of subjects under Mild Leniency: Although the vast majority of subjects chose the strategy associated with the risk-dominant equilibrium, (R, R), some subjects attempted to coordinate on their preferred outcome and chose the no-report strategy.[55]

Result 3. *Strong Leniency significantly increases the rate of self-reporting, with respect to Mild Leniency.*

[54] Regarding the behavior of pairs (pair-level data), our findings suggest that Mild Leniency significantly increases the (R, R) rate (.58 v. zero, for Mild and No Leniency, respectively; $p < .001$, one-sided Fisher's exact test), and No Leniency significantly increases the (NR, NR) rate (.89 v. .05, for No Leniency and Mild Leniency, respectively; $p < .001$, one-sided Fisher's exact test). These results are consistent across b-values. See Supplementary Material for details.

[55] Regarding the behavior of pairs (pair-level data), our findings suggest that, although Strong Leniency increases the (R, R) rate (79 v. 58%, for Strong and Mild Leniency, respectively; $p = .031$, one-sided Fisher's exact test), Strong and Mild Leniency exhibit the same (NR, NR) rate (zero v. 5%, for Strong and Mild Leniency, respectively; $p = .253$, one-sided Fisher's exact test). These results are consistent across b-values. See Supplementary Material for details.

Table 5 – Effect of Ordered Leniency on Time to Submit Chosen Action in the Self-Reporting Stage[a]

Conditions	Number of Individuals[b]	Time to Submit Chosen Action (R or NR)[c]	p-value[d]
Strong Leniency v. No Leniency	78, 70	3.49 v. 16.24 (7.25) (17.77)	$p < .001$
Mild Leniency v. No Leniency	80, 70	4.99 v. 16.24 (9.51) (17.77)	$p < .001$

Notes: [a]"v." denotes "versus;" [b]number of individuals who committed the act; [c]mean time to submit chosen action (R or NR) in seconds, conditional on committing the act; standard deviation in parentheses; [d]p-value corresponds to Wilcoxon rank-sum (Mann Whitney) test.

5.2.2 Race-to-the-Courthouse Effect

We now investigate whether ordered-leniency policies exhibit a race-to-the-courthouse effect. In theory, Strong Leniency and Mild Leniency when the risk-dominance refinement is applied incentivize the injurers to minimize their reporting times to increase their chances to get the first position in the self-reporting queue. As a result, all injurers self-report immediately.

As discussed in Section 4.2, given our general theoretical framework and robust experimental design, we can assess whether ordered-leniency policies exhibit this "incentivizing" property in the lab by comparing the times to submit the chosen action under Strong Leniency and Mild Leniency with the time to submit the chosen action under No Leniency. In theory, Strong and Mild Leniency when the risk-dominance refinement is applied (weakly) reduce the time to submit the chosen action, with respect to No Leniency. In other words, ordered-leniency policies exhibit a race-to-the-courthouse effect.

Table 5 summarizes our findings. Consider first the effect of implementing Strong Leniency. Our results indicate that the average time to submit the chosen action under Strong Leniency is less than 4 seconds and the average time to submit the chosen action under No Leniency is more than 16 seconds, a highly significant difference ($p < .001$). These findings are mainly explained by the high report rate and the relatively low time to submit the R action under Strong Leniency, and the high no report rate and the relatively high time to submit the NR action under No Leniency. These results suggest that Strong Leniency generates a race-to-the-courthouse effect where self-reporting occurs promptly, and provides strong support to Hypothesis 3.

Result 4. *Strong Leniency significantly reduces the time to submit the chosen action, with respect to No Leniency. There is a race-to-the-courthouse effect.*

Consider now the effect of implementing Mild Leniency. Our findings suggest that the average time to submit the chosen action under Mild Leniency is less than 5 seconds and the average time to submit the chosen action under No Leniency is more than 16 seconds, a highly significant difference ($p < .001$). Our results indicate that Mild Leniency incentivizes the injurers to self-report promptly. In other words, Mild Leniency exhibits a race-to-the-courthouse effect. These findings provide additional evidence of the subjects' choice of the risk-dominant strategy and a strong support to Hypothesis 3.[56]

Result 5. *Mild Leniency significantly reduces the time to submit the chosen action, with respect to No Leniency. There is a race-to-the-courthouse effect.*

5.2.3 Detection Rate

Next, we study the effect of ordered leniency on the detection rate. Table 6 reports our findings. In theory, when Strong Leniency or Mild Leniency (when the risk-dominance refinement is applied) are implemented, both injurers self-report. Hence, an individual who committed a harmful act is detected with certainty in equilibrium. In contrast, when No Leniency for self-reporting or Mild Leniency (when the Pareto-dominance refinement is applied) are implemented, no injurer self-reports the act. Therefore, an individual who committed a harmful act is detected with a 40% likelihood only. Our results indicate that ordered-leniency policies are very effective detection mechanisms. Specifically, Strong Leniency significantly increases the detection rate of an injurer (99 v. 53%, for Strong and No Leniency, respectively; $p < .001$). Similarly, Mild Leniency significantly increases the detection rate of an injurer (96 v. 53%, for Mild and No Leniency, respectively; $p < .001$). Importantly, as predicted, Mild and Strong Leniency exhibit the same likelihood of detection. Our findings are aligned with Hypotheses 4 and 5, and provide strong support for the theoretical predictions under the risk-dominance refinement.[57]

[56] Our results also suggest that the average times to submit the chosen action under Strong and Mild Leniency are only marginally different ($p = .062$, Wilcoxon rank-sum (Mann Whitney) test).

[57] Regarding the likelihood of detection of one or both members of a pair of injurers (pair-level data), our results suggest that the implementation of ordered leniency significantly increases the detection rate of one or both injurers (100 v. 54%, for Strong and No Leniency, respectively; $p < .001$, one-sided Fisher's exact test; and, 98 v. 54%, for Mild and No Leniency, respectively; $p < .001$, one-sided Fisher's exact test).

Table 6 – Effect of Ordered Leniency on Detection[a]

Conditions	Number of Individuals[b]	Detection Rate[c]	p-value[d]
Strong Leniency v. No Leniency	78, 70	.99 v. .53	$p < .001$
Mild Leniency v. No Leniency	80, 70	.96 v. .53	$p < .001$
Mild Leniency v. Strong Leniency	80, 78	.96 v. .99	$p = .320$

Notes: [a]"v." denotes "versus;" [b]number of individuals who committed the act; [c]detection rate conditional on committing the act; [d]p-value corresponds to one-sided Fisher's exact test.

Result 6. *Strong Leniency significantly increases the detection rate, with respect to No Leniency.*

Result 7. *Mild Leniency significantly increases the detection rate, with respect to No Leniency.*

Result 8. *Strong Leniency and Mild Leniency exhibit the same detection rate.*

5.2.4 Individual Fine

We now investigate the effect of implementing ordered-leniency policies on the individual fine. Table 7 reports our results. In theory, when the risk-dominance refinement is applied, ordered-leniency policies increase the expected fine. Mild Leniency exhibits the highest expected fine and No Leniency exhibits the lowest expected fine. These outcomes are explained by the higher detection rates under ordered leniency. Our results indicate that the average individual fines under Strong, Mild and No Leniency are equal to 588.46, 638.25 and 475.71, respectively, and that the pairwise differences are significant (Strong v. No Leniency, $p = .026$; Mild v. No Leniency, $p = .035$; and, Mild v. Strong Leniency, $p = .008$). These results are aligned with our theoretical predictions under the risk-dominance refinement and provide support for Hypothesis 6.

Result 9. *Strong Leniency significantly increases the individual fine, with respect to No Leniency.*

Result 10. *Mild Leniency significantly increases the individual fine, with respect to Strong Leniency and No Leniency.*

Table 7 – Effect of Ordered Leniency on Individual Fine[a]

Conditions	Number of Individuals[b]	Individual Fine[c]	p-value[d]
Strong Leniency v. No Leniency	78, 70	588.46, v. 475.71 (307.47) (452.51)	$p = .026$
Mild Leniency v. No Leniency	80, 70	638.25 v. 475.71 (268.71), (452.51)	$p = .035$
Mild Leniency v. Strong Leniency	80, 78	638.25 v. 588.46 (268.71) (307.47)	$p = .008$

Notes: [a]"v." denotes "versus;" [b]number of individuals who committed the act; [c]mean individual fine conditional on committing the act; standard deviation in parentheses; [d]p-value corresponds to Wilcoxon rank-sum (Mann-Whitney) test.

Result 11. *Mild Leniency significantly increases the individual fine, with respect to Strong Leniency.*

5.2.5 Deterrence Rate

Remember that the deterrence rate is defined as the percentage of groups where one or both potential injurers decided not to participate in the act, i.e., the percentage of harmful acts that are deterred. Then, the analysis in this section uses pair-level data.

We start our analysis of deterrence by studying the effect of the benefit derived from the commission of the harmful act (b-value) on the deterrence rate, a within-condition analysis. Table 8 outlines our findings. For each condition, we consider the theoretical deterrence threshold, and compare the deterrence rates for b-values below and above this threshold. Given that our previous analysis of the effect of ordered-leniency policies on self-reporting suggests that the risk-dominant N.E. (the (R, R) outcome) is chosen by the majority of subjects under Mild Leniency, we considered the deterrence threshold that corresponds to this refinement. Two important insights deserve attention. First, our findings suggest that the benefit from the harmful act incentivizes the potential injurers to commit the act. Consider, for instance, the effect of the b-value on deterrence for the case of Strong Leniency (first three lines of Table 8). The data indicate that when the benefits are above the theoretical threshold (600), the deterrence rate is equal to 6%; and, when the benefits are below this threshold, the deter-

Table 8 – Effect of Benefit on Deterrence[a]
(Within-Condition Analysis)

Condition	Number of Pairs[b]	Deterrence Rate
Strong Leniency		
$b \in [200, 600]$	15	.47
$b \in (600, 1600]$	33	.06
p-value		$p = .002$
Mild Leniency		
$b \in [200, 660]$	26	.31
$b \in (660, 1600]$	22	.00
p-value		$p = .004$
No Leniency		
$b \in [200, 360]$	4	.75
$b \in (360, 1600]$	44	.23
p-value		$p = .055$

Notes: [a]p-value corresponds to one-sided Fisher-exact test; [b]total number of pairs.

rence rate rises to 47% (a statistically significant effect, $p = .002$). More generally, for each leniency environment, the likelihood of deterrence is significantly lower when the benefits are greater than the theoretical deterrence threshold. These results are aligned with our theory and provide strong support for Hypothesis 7.

Result 12. *Within each leniency environment, the deterrence rate is lower when the benefit from the harmful act is greater than the deterrence threshold.*

Second, our results indicate that the deterrence rates for b-values below the theoretical thresholds are lower than the rates predicted by the theory for the Strong, Mild and No Leniency policies (47, 31 and 75% instead of 100%). The low deterrence rate under Strong Leniency might suggest that some subjects considered an alternative deterrence threshold. In theory, both injurers report immediately, and hence, they are equally likely to get the first position in the self-reporting queue (i.e., each has a 50% chance to get the first position). In the lab, some subjects might exhibit cognitive biases, such as self-serving bias (Babcock et al., 1995), and believe that their chances to be the first to report are greater than the chances of the other injurers (i.e., that their chances are greater than 50%). In the limiting case, some subjects might believe that they will always be the first to report. As a result, they will consider a fine equal to 300 instead of an expected fine equal to 600 when making their decisions about participating in the harmful activity. Hence, they will also choose to participate in the harmful activity when the benefits are lower than 600.[58]

Regarding Mild Leniency, the low deterrence rate might be explained, in part, by the decisions in Stage 2. Remember that, although the majority of subjects coordinated on the risk-dominant N.E. (R, R), some subjects attempted to coordinate on the Pareto-dominant N.E. (NR, NR). Subjects who expect to coordinate on (NR, NR) will consider a lower deterrence threshold (360 instead of 660). Hence, they will also choose to commit the act when the benefits are lower than 660. The low deterrence rate might be also explained by the presence of self-serving bias on the subjects' beliefs about their chances to be the first to report. Some subjects might believe that they will always be the first to report. As a result, they will consider a fine equal to 420 instead of an expected fine equal to 660. Hence, they will also choose to commit the act when the benefits are lower than 660.[59]

Finally, the relatively high deterrence rate under No Leniency for b-values below the theoretical threshold (75 versus 47 and 31% for Strong and Mild Leniency) might be explained by the degree of ambiguity across strategic environments. As argued by Kunda (1990, 1987), environments characterized by strong ambiguity are more prone to elicit self-serving bias. In our experiment, subjects might perceive a stronger degree of ambiguity under the Strong and Mild Leniency conditions because the expected payoffs depend also on the subjects' beliefs about their likelihoods to get the first position in the self-reporting queue. As a result, we might expect that self-serving bias would be more prominent in the Strong and Mild Leniency conditions.

We now assess the effect of ordered leniency on the deterrence rate. Table 9 summarizes our results. Remember first that, in theory, when the benefit from committing the harmful act is not greater than 600, the deterrence rate under Strong Leniency is higher than the deterrence rate under No Leniency (100 and 27%, for Strong Leniency and No Leniency, respectively);[60] when the benefit is greater than 600, Strong Leniency and No Leniency exhibit the same zero deterrence rate. Second, in theory, when the benefit from committing the harmful act is

[58] See Section 6 for further discussion.

[59] Failure to apply backward induction due to limited computational abilities might explain some of the deviations from the theoretical predictions on deterrence, under Strong, Mild and No Leniency. Camerer and Johnson (2004) state, "motivated intelligent subjects behave sensibly, but do not exhibit the extent of strategic reasoning which is commonly assumed when game theory is applied" (p. 15). See also Johnson et al. (2002).

[60] Given the number of pairs per b-value segment, the deterrence rate for No Leniency is equal to 4/15.

Table 9 – Effect of Ordered Leniency on Deterrence[a]

Conditions	Number of Pairs[b]	Deterrence Rate	p-value[c]
Strong Leniency v. No Leniency			
$b \in [200, 600]$	15, 15	.47 v. .47	$p = .642$
$b \in (600, 1600]$	33, 33	.06 v. .18	$p = .129$
Across b-values	48, 48	.19 v. .27	$p = .233$
Mild Leniency v. No Leniency			
$b \in [200, 660]$	26, 26	.31 v. .38	$p = .386$
$b \in (660, 1600]$	22, 22	.00 v. .14	$p = .116$
Across b-values	48, 48	.17 v. .27	$p = .162$
Mild Leniency v. Strong Leniency			
$b \in [200, 660]$	26, 26	.31 v. .31	$p = .618$
$b \in (660, 1600]$	22, 22	.00 v. .05	$p = .500$
Across b-values	48, 48	.17 v. .19	$p = .500$

Notes: [a]"v." denotes "versus;" [b]total number of pairs; [c]p-value corresponds to one-sided Fisher's exact test.

not greater than 660 and the risk-dominance refinement is applied, the deterrence rate under Mild Leniency is higher than the deterrence rate under Strong Leniency and No Leniency (100, 58 and 15%, for Mild, Strong, and No Leniency, respectively);[61] when the benefit from committing the harmful act is greater than 660 and the risk-dominance refinement is applied, Strong Leniency, Mild Leniency and No Leniency exhibit a zero deterrence rate.[62] Although the behaviors of subjects in the self-reporting stage, and hence, the actual fines, are aligned with the theory, the deterrence rates across conditions are not significantly different. These results might be explained by the relatively low deterrence rates for b-values below the theoretical thresholds under Strong and Mild Leniency due, for instance, to self-serving bias. As a result, acts are committed more frequently than predicted. These findings are aligned with previous literature on crime and deterrence (Nagin and Pogarsky, 2003; Kleck et al., 2005; Kleck and

[61]Given the number of pairs per b-value segment, the deterrence rates under Strong and No Leniency are equal to 15/26 and 4/26, respectively.

[62]In theory, across b-values, the deterrence rates under Strong Leniency, Mild Leniency and No Leniency are equal to 31, 54, and 8%, respectively. Given the number of pairs per b-value segment, across b-values, the deterrence rates for Strong Leniency, Mild Leniency and No-Leniency are equal to 4/48, 15/48, and 26/48, respectively.

Barnes, 2014; Lochner, 2007).[63]

In sum, our experimental results demonstrate the effectiveness of ordered-leniency policies as detection mechanisms. In particular, the implementation of either Strong or Mild Leniency policies significantly increases the likelihood of self-reporting. These findings provide support for the theoretical predictions under the risk-dominance refinement. We provide empirical evidence of a race-to-the-courthouse effect of ordered-leniency policies: Prompt self-reporting is observed when Strong or Mild Leniency policies are implemented. Importantly, ordered-leniency policies significantly raise the likelihood of detection of harmful acts. As a consequence, the average fines significantly increase. As predicted under the risk-dominance refinement, Mild Leniency exhibits the highest average fine. Overall, our results are aligned with the theoretical predictions.

Our findings provide important insights for the design of enforcement policies with ordered leniency. We present strong evidence that the injurers also confront "the dilemma" of choosing self-reporting instead of no reporting (their preferred outcome) in coordination-game environments (Mild Leniency). These results suggest that the enforcement agency does not need to rely on strong penalty reductions (i.e., on the implementation of Strong Leniency) to elicit self-reporting of harmful activities. Mild Leniency might also induce cooperation with the enforcement agency when the injurers are sufficiently distrustful of each other after committing the act.

6 Discussion and Conclusions

Can an enforcement authority strengthen detection and deterrence of socially harmful activities by implementing ordered-leniency policies? Landeo and Spier (2018) consider a theoretical model of law enforcement with ordered leniency, where the design of optimal policies depends crucially on the equilibrium refinement applied in case of multiplicity of equilibria. When the risk-dominance refinement is applied, they demonstrate that the enforcement authority can induce all injurers to immediately self-report by offering only small penalty discounts. If the Pareto-dominance refinement is applied instead, the enforcement agency must offer bigger penalty discounts to induce self-reporting. As a result of the high self-reporting rates under

[63]See Chalfin and Mc Crary (2017) for a survey on the literature on crime and deterrence. See Section 6 for further discussion.

ordered leniency, detection of harmful acts is enhanced without increasing investigatory costs. Given the higher detection rates under ordered leniency, the expected sanctions rise, and hence, deterrence and social welfare are enhanced.

Our experimental findings suggest that ordered-leniency policies are effective detection mechanisms. In particular, Strong and Mild Leniency significantly increase the likelihood of self-reporting. These results suggest that the subjects' choices are aligned with the risk-dominance refinement. We provide empirical evidence of a race-to-the-courthouse effect of ordered-leniency policies: Prompt self-reporting is observed when Strong or Mild Leniency policies are implemented. Importantly, ordered-leniency policies raise the likelihood of detection: Harmful acts are detected with (almost) certainty when Strong or Mild Leniency are implemented. Given the high detection rates under ordered leniency, the average fines rise. Our results on deterrence indicate that some subjects systematically underestimate the likelihood and severity of sanctions when making their decisions about participating in the harmful act. These findings might suggest the presence of self-serving bias on subjects' beliefs about getting the first position in the self-reporting queue. As a result, deterrence is weakened, and harmful acts are committed more frequently than predicted.

Importantly, our findings are aligned with previous literature on crime and deterrence. In particular, in a recent *Journal of Economic Literature* survey on the economics and criminology literature on crime and deterrence, Chalfin and McCrary (2017) point out the importance of perceptual deterrence (i.e., the individual's perception of the risk associated with the participation in an illegal activity).[64] This literature provides significant insights for the assessment of law enforcement policies with ordered leniency. First, the findings from the crime and deterrence literature about the association between perceptual deterrence, self-serving bias and crime participation provide support for our claim that the relatively low deterrence rate under ordered leniency might be explained by the individuals' biased perceptions of the risks associated with the participation in the harmful activity. Second, the findings from this literature on the sensitivity of perceptual deterrence to crime detection experiences suggest that ordered-

[64]Apel (2013) and Pogarsky (2009) provide two other excellent surveys on this literature. Three important findings from this literature deserve special attention: (1) The difference between perceptual deterrence (subjective risk) and actual risk (Kleck et al., 2005; Kleck and Barnes, 2014; Lochner, 2007); (2) the malleability of perceptual deterrence and the sensitivity of perceptual deterrence to crime detection experiences (Pogarsky et al, 2005, 2004); and, (3) the association between perceptual deterrence, self-serving bias, and participation in harmful activities (Nagin and Pogarsky, 2003).

leniency policies might have an additional welfare-enhancing effect associated with deterrence: By increasing the likelihood of detection, ordered-leniency policies might help align individuals' subjective perceptions of risk and actual risk. As a result, a higher deterrence of unlawful activities committed by former wrongdoers might be observed.

Significant policy implications are derived from our study. We demonstrate that the enforcement agency might generate optimal detection of harmful activities by implementing a coordination game or a prisoners' dilemma game through Mild and Strong Leniency, respectively. In particular, when Mild Leniency is implemented, although the injurers are jointly better off by refusing to cooperate with the enforcement authority, they are individually induced to cooperate due to strategic uncertainty. As a result, immediate self-reporting and optimal detection are also elicited under Mild Leniency. Our findings challenge the common view among legal practitioners and policy makers that only the prisoners' dilemma game is relevant for the design of enforcement policies and cooperation agreements, and provide relevant insights for the design of enforcement policies with self-reporting in a variety of real-world environments.[65]

Our analysis focuses on environments where the harmful acts are committed by pairs of injurers. In some real-world applications, the groups of injurers might involve more than two members. It might be interesting to experimentally study environments with more than two potential injurers, and to assess how the group size affects self-reporting, detection and deterrence. In theory, immediate self-reporting by all injurers is robust to group size (Landeo and Spier, 2018). In the lab, however, larger groups may find even harder to coordinate their actions on their preferred outcome. As a result, the detection power of Mild Leniency policies might be stronger than observed here.

In our experimental setting, both injurers receive the same benefits from participating in the harmful act. A possible extension might experimentally assess the performance of ordered-leniency policies in environments with asymmetric private benefits from committing the act. Landeo and Spier (2018) demonstrate that, in theory, the results derived from the benchmark model are robust to asymmetric benefits as long as the injurers can write side contracts with each other and negotiate transfer payments. In experimental settings, however, including a transfer negotiation stage might also affect the participation and self-reporting decisions of the injurers. These, and other extensions, may be fruitful topics for future research.

[65]See Section 1 for further discussion.

References

Andreoni, James. 1991. "The Desirability of a Permanent Tax Amnesty." *Journal of Public Economics*, 45: 143–159.

Apel, Robert. 2013. "Sanctions, Perceptions, and Crime: Implications for Criminal Deterrence." *Journal of Quantitative Criminology,* 29: 67–101.

Apesteguia, Jose, Martin Dufwenberg, and Reinhard Selten. 2007. "Blowing the Whistle." *Economic Theory,* 31: 143–166.

Arlen, Jennifer and Reinier Kraakman. 1997. "Controlling Corporate Misconduct: An Analysis of Corporate Liability Regimes." *New York University Law Review,* 72: 687–779.

Aubert, Cécile, Patrick Rey, and William E. Kovacic. 2006. "The Impact of Leniency and Whistle-Blowing Programs on Cartels." *International Journal of Industrial Organization,* 24: 1241–1266.

Babcock, Linda, George Loewenstein, Samuel Issacharoff, and Colin Camerer. 1995. "Biased Judgments of Fairness in Bargaining." *American Economic Review*, 11: 109–26.

Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy*, 76: 169–217.

Bigoni, Maria, Sven-Olof Fridolfsson, Chloe Le Coq, and Giancarlo Spagnolo. 2015. "Trust, Leniency, and Deterrence." *Journal of Law, Economics, and Organization*, 31: 663–689.

Bigoni, Maria, Sven-Olof Fridolfsson, Chloe Le Coq, and Giancarlo Spagnolo. 2012. "Fines, Leniency, and Rewards in Antitrust." *RAND Journal of Economics*, 43: 368–90.

Buccirossi, Paolo and Giancarlo Spagnolo. 2006. "Leniency Policies and Illegal Transactions." *Journal of Public Economics*, 90: 1281–1297.

Burton, Anthony and Martin Sefton. 2004. "Risk, Pre-Play Communication and Equilibrium." *Games and Economic Behavior*, 46: 23–40.

Camerer, Colin and Eric J. Johnson. 2004. "Thinking about Attention in Games: Backward and Forward Induction." In Isabelle Brocas and Juan D. Carrillo, eds., *The Psychology of Economic Decisions. Volume Two: Reasons and Choices.* New York: Oxford University Press Inc.

Chalfin, Aaron and Justin McCrary. 2017. "Criminal Deterrence: A Review of the Literature." *Journal of Economic Literature*, 55: 5–48.

Chen, Zhijun and Patrick Rey. 2013. "On the Design of Leniency Programs." *Journal of Law and Economics*, 56: 917–957.

Cooper, Russell W., Douglas V. DeJong, Robert Forsythe, and Thomas W. Ross. 1992. "Communication in Coordination Games." *Quarterly Journal of Economics*, 107: 739–771.

Feess, Eberhardt and Markus Walzl. 2004. "Self-Reporting in Optimal Law Enforcement When There Are Criminal Teams." *Economica*, 71: 333-348.

Feltovich, Nick and Yasuyo Hamaguchi. 2018. "The Effect of Leniency Programmes on Anti-Competitive Behaviour: An Experimental Study." *Southern Economic Journal,* 84: 1024–1049.

Greenberg, Michael D. 2011. *For Whom the Whistle Blows. Advancing Corporate Compliance and Integrity Efforts in the Era of Dodd-Frank.* California: The RAND Corporation.

Grossman, Gene M. and Michael L. Katz. 1983. "Plea Bargaining and Social Welfare." *American Economic Review*, 73: 749–757.

Harsanyi, John C. and Reinhard Selten. 1988. *A General Theory of Equilibrium Selection in Games*. Cambridge: MIT Press.

Hinloopen, Jeroen and Adriaan R. Soetevent. 2008. "Laboratory Evidence on the Effectiveness of Corporate Leniency Programs." *RAND Journal of Economics*, 39: 607–616.

Innes, Robert. 1999. "Remediation and Self-reporting in Optimal Law Enforcement." *Journal of Public Economics*, 72: 379-393.

Johnson, Eric J., Colin Camerer, Sankar Sen, and Talya Rymon. 2002. "Detecting Failures of Backward Induction: Monitoring Information Search in Sequential Bargaining." *Journal of Economic Theory*, 104: 16–47.

Kaplow, Louis and Steven Shavell. 1994. "Optimal Law Enforcement with Self-Reporting of Behavior." *Journal of Political Economy*, 102: 583–606.

Kobayashi, Bruce. 1992. "Deterrence with Multiple Defendants: An Explanation for "Unfair" Plea Bargains." *RAND Journal of Economics*, 23: 507–517.

Kraakman, Reinier H. 1986. "Gatekeepers: The Anatomy of a Third-Party Enforcement Strategy." *Journal of Law, Economics & Organization*, 2: 53–104.

Kleck, Gary, and J. C. Barnes. 2014. "Do More Police Lead to More Crime Deterrence?" *Crime and Delinquency*, 60: 716–38.

Kleck, Gary, Brion Sever, Spencer Li, and Marc Gertz. 2005. "The Missing Link in General Deterrence Research." *Criminology*, 43: 623–60.

Kunda, Ziva. 1990. "The Case of Motivated Reasoning." *Psychological Bulletin,* 108: 480–98.

Kunda, Ziva. 1987. "Motivated Inference: Self-Serving Generation and Evaluation of Causal Theories." *Journal of Personality and Social Psychology,* 53: 636–47.

Landeo, Claudia M. 2018. "Law and Economics and Tort Litigation Institutions: Theory and Experiments." Chapter 9. In Kathryn Zeiler and Joshua Teitelbaum, eds., *Research Handbook on Behavioral Law and Economics.* Massachusetts: Edward Elgar Publishing.

Landeo, Claudia M. 2009. "Cognitive Coherence and Tort Reform." *Journal of Economic Psychology*, 6: 898–912.

Landeo, Claudia M. and Kathryn E. Spier. 2018. "Optimal Law Enforcement with Ordered Leniency." Mimeo. University of Alberta and Harvard University.

Landeo, Claudia M. and Kathryn E. Spier. 2015. "Incentive Contracts for Teams: Experimental Evidence." *Journal of Economic Behavior and Organization*, 119: 496–511.

Landeo, Claudia M. and Kathryn E. Spier. 2012. "Exclusive Dealing and Market Foreclosure: Further Experimental Results." *Journal of Institutional and Theoretical Economics*, 168: 150–170.

Landeo, Claudia M. and Kathryn E. Spier. 2009. "Naked Exclusion: An Experimental Study of Contracts with Externalities." *American Economic Review*, 99: 1850–1877.

Landeo, Claudia M., Maxim Nikitin, and Sergei Izmalkov. 2013. "Incentives for Care, Litigation, and Tort Reform under Self-Serving Bias." In Thomas J. Miceli and Matthew J. Baker, eds., *Research Handbook on Economic Models of Law*. Massachusetts: Edward Elgar Publishing.

Landes, William M. 1971. "An Economic Analysis of the Courts." *Journal of Law and Economics*, 14: 61–108.

Livernois, John and C.J. McKenna. 1999. "Truth or Consequences: Enforcing Pollution Standards with Self-Reporting." *Journal of Public Economics*, 71: 415–440.

Lochner, Lance. 2007. "Individual Perceptions of the Criminal Justice System." *American Economic Review*, 97: 444–6 0.

Malik, Arun S. 1993. "Self-Reporting and the Design of Policies for Regulating Stochastic Pollution." *Journal of Environmental Economics and Management*, 24: 241–257.

Malik, Arun S. and Robert M. Schwab. 1991. "The Economics of Tax Amnesties." *Journal of Public Economics*, 46: 29–49.

Motta, Massimo and Michele Polo. 2003. "Leniency Programs and Cartel Prosecution." *International Journal of Industrial Organization*, 21: 347–379.

Nagin, Daniel S. and Greg Pogarsky. 2003. "An Experimental Investigation of Deterrence: Cheating, Self-Serving Bias, and Impulsivity." *Criminology*, 41: 167–194.

Ochs, Jack. 1995. "Coordination Problems." In J.H. Kagel and A.E. Roth, eds., *Handbook of Experimental Economics*. New Jersey: Princeton University Press Inc.

Pogarsky, Greg. 2009. "Deterrence and Decision Making: Research Questions and Theoretical Refinement." In Krohn, Marvin D., Alan J. Lizzote and Gina Penly Hall, eds., *Handbook on Crime and Deviance*. New York: Springer.

Pogarsky, Greg, KiDeuk Kim, and Ray Paternoster. 2005. "Perceptual Change in the National Youth Survey: Lessons for Deterrence Theory and Offender Decision-Making." *Justice Quarterly*, 22: 1–29.

Pogarsky, Greg, Alex R. Piquero, and Ray Paternoster. 2004. "Modeling Change in Perceptions about Sanction Threats: The Neglected Linkage in Deterrence Theory." *Journal of Quantitative Criminology*, 20: 343–69.

Polinsky, A. Mitchell and Steven Shavell. 1984. "The Optimal Use of Fines and Imprisonment." *Journal of Public Economics*, 24: 89–99.

Smith, Angela S. . 2011. "An Experimental Study of Exclusive Contracts." *International Journal of Industrial Organization*, 29: 4–13.

Spagnolo, Giancarlo. 2005. "Divide et Impera: Optimal Leniency Programs." Mimeo, Stockholm School of Economics.

Spagnolo, Giancarlo and Catarina Marvão. 2016. "Cartels and Leniency: Taking Stock of What We Learnt." In Luis C. Corchón and Marco A. Marini, eds., *Handbook of Game Theory and Industrial Organization*. Massachusetts: Edward Elgar Publishing.

Appendix: Sample Instructions (Mild-Leniency Condition)

PLEASE GIVE THIS MATERIAL TO THE EXPERIMENTER AT THE END OF THE EXPERIMENT

INSTRUCTIONS

This is an experiment in the economics of decision-making. The National Science Foundation has provided the funds for this research.

In this experiment you will be asked to play an economic decision-making computer game. The experiment currency is the "token." The instructions are simple. If you follow them closely and make appropriate decisions, you may make a large amount of money. At the end of the session you will be paid your game earnings in CASH. If you have any questions at any time, please raise your hand and the experimenter will go to your desk.

PROBABILITY OR CHANCE

The concept of probability or chance will be used in this experiment. **PROBABILITY OR CHANCE (EXPRESSED IN PERCENTAGES)** indicates the likelihood of occurrence of uncertain events. The concept of probability or chance can be illustrated as follows. Suppose that an urn contains 100 balls: 20 out of 100 balls are white, and 80 out of 100 balls are black. Suppose that you randomly extract one ball from the urn. The chance that the ball will be white is equal to 20% because 20 out of 100 balls in the urn are white. Similarly, the chance that the ball will be black is equal to 80%, because 80 out of 100 balls in the urn are black.

SESSION AND PLAYERS

The session is made up of 6 matches. The first 5 matches are practice matches. After the last practice match, **ONE ACTUAL MATCH** will be played.

1) At the beginning of the session, every participant will be randomly assigned a role. The equally likely roles are: **Player B1** and **Player B2**.

   The **ROLES WILL REMAIN THE SAME until the end of the session**.

2) Before the beginning of **EACH PRACTICE MATCH**, the computer will randomly form pairs of **TWO PEOPLE**: **Player B1** and **Player B2**.

   YOU WILL NOT KNOW THE IDENTITY OF YOUR PARTNER.
   YOU WILL PLAY WITH A DIFFERENT PARTNER IN EVERY PRACTICE MATCH.

3) Before the beginning of the **ACTUAL MATCH**, the computer will randomly form pairs of **TWO PEOPLE**: **Player B1** and **Player B2**.

   YOU WILL NOT KNOW THE IDENTITY OF YOUR PARTNER.
   YOU WILL NOT PLAY WITH ANY OF YOUR PREVIOUS PARTNERS.

MATCH STAGES

STAGE 1: DECISION WHETHER TO JOINTLY COMMIT THE ACT

1) Each player has an **initial endowment** equal to **700 tokens**.

2) **THE DECISION TO JOINTLY COMMIT THE ACT** refers to an economic decision involving potential economic benefits and potential economic losses.

- **ECONOMIC BENEFITS** might occur in **STAGE 1**.
- **ECONOMIC LOSSES** might occur in **STAGE 2**.

3) **THE COMPUTER** randomly determines the **NUMBER OF TOKENS X** that each player will get **IF BOTH PLAYERS AGREE TO JOINTLY COMMIT THE ACT.** Both players will receive the same number of tokens X.

- The number of tokens X can be equal to **200, …, 1598, 1599, 1600 tokens**.
- The number of tokens X will be revealed to both players.

4) **Player B1** and **Player B2** decide whether to **<u>JOINTLY COMMIT THE ACT</u>**

- Each player will have **<u>90 SECONDS</u> TO DECIDE WHETHER TO AGREE TO JOINTLY COMMIT THE ACT OR NOT AGREE TO JOINTLY COMMIT THE ACT AND <u>PRESS THE NEXT BUTTON.</u>**
    - If a player **<u>FAILS TO MAKE A CHOICE AND TO PRESS THE NEXT BUTTON WITHIN THE 90-SECOND PERIOD,</u>** it will be implied that he/she decided **<u>NOT TO AGREE TO JOINTLY COMMIT</u>** the act.

5) The possible outcomes are as follows.

- **<u>BOTH PLAYERS AGREE TO JOINTLY COMMIT THE ACT:</u>** Each player gets **X TOKENS** (in addition to the initial endowment of 700 tokens) and **STAGE 2 BEGINS**.

- **<u>ONLY ONE PLAYER AGREES TO JOINTLY COMMIT THE ACT</u>:** Each player gets **ZERO TOKENS** and the **MATCH ENDS**. The match payoff for each player will be 700 tokens (initial endowment).

- **<u>NEITHER PLAYER AGREES TO JOINTLY COMMIT THE ACT</u>:** Each player gets **ZERO TOKENS** and the **MATCH ENDS**. The match payoff for each player will be 700 tokens (initial endowment).

STAGE 2: DECISION WHETHER TO REPORT THE ACT

1) If **both payers agreed to jointly commit the act**, then Stage 2 begins.

2) **A FINE EQUAL TO 900 TOKENS MIGHT BE DEDUCTED FROM A PLAYER'S TOKEN BALANCE AS A CONSEQUENCE OF JOINTLY COMMITTING THE ACT.**

    - A player's decision to report the act **MIGHT DECREASE THE FINE HE/SHE WILL PAY** from 900 tokens to 420 tokens.

    - A player's decision to report the act **MIGHT INCREASE HIS/HER PARTNER'S CHANCE TO PAY A FINE** from 40% to 90% or from 40% to 100%.

    THE SPECIFIC FINE AND THE CHANCE OF PAYING THAT FINE DEPEND ON THE DECISIONS OF BOTH PLAYERS ABOUT REPORTING THE ACT.

3) Each player will have **90 SECONDS TO DECIDE WHETHER TO REPORT OR NOT TO REPORT THE ACT AND PRESS THE NEXT BUTTON.**

    - If a player **FAILS TO MAKE A CHOICE AND TO PRESS THE NEXT BUTTON WITHIN THE 90-SECOND PERIOD,** it will be implied that he/she decided **NOT REPORT** the act.

4) The possible outcomes and match payoffs are presented below.

POSSIBLE OUTCOME 1: BOTH PLAYERS DECIDE NOT TO REPORT THE ACT

- <u>**NEITHER PLAYER GETS A FINE REDUCTION**</u>
- <u>**EACH PLAYER'S CHANCE OF PAYING A FINE EQUAL TO 900 TOKENS IS 40%**</u>: Each player pays a fine equal to **900 tokens** with **40% chance** and **does not pay any fine** with **60% chance**.

Hence, the match payoffs are as follows.

With a **40% CHANCE,** the match payoffs will be:

Player B1's match payoff = 700 tokens + X tokens – 900 tokens
Player B2's match payoff = 700 tokens + X tokens – 900 tokens

With a **60% CHANCE,** the match payoffs will be:

Player B1's match payoff = 700 tokens + X tokens – 0 tokens
Player B2's match payoff = 700 tokens + X tokens – 0 tokens

POSSIBLE OUTCOME 2: ONLY PLAYER B1 DECIDES TO REPORT THE ACT

- **ONLY PLAYER B1 GETS A FINE REDUCTION**: Instead of paying a fine equal to 900 tokens, Player B1 **always** pays only **420 tokens**.
- **PLAYER B2'S CHANCE OF PAYING A FINE EQUAL TO 900 TOKENS IS 90%**: Player B2 pays a fine equal to **900 tokens** with **90% chance** and **does not pay any fine** with **10% chance**.

Hence, the match payoffs are as follows.

With a **90% CHANCE,** the match payoffs will be:

Player B1's match payoff = 700 tokens + X tokens – 420 tokens
Player B2's match payoff = 700 tokens + X tokens – 900 tokens

With a **10% CHANCE,** the match payoffs will be:

Player B1's match payoff = 700 tokens + X tokens – 420 tokens
Player B2's match payoff = 700 tokens + X tokens – 0 tokens

POSSIBLE OUTCOME 3: ONLY PLAYER B2 DECIDES TO REPORT THE ACT

- **ONLY PLAYER B2 WILL GET A FINE REDUCTION**: Instead of paying a fine equal to 900 tokens, Player B2 **always** pays only **420 tokens**.
- **PLAYER B1'S CHANCE OF PAYING A FINE EQUAL TO 900 TOKENS IS 90%**: Player B1 pays a fine equal to **900 tokens** with **90% chance** and **does not pay any fine** with **10% chance**.

Hence, the match payoffs are as follows.

With a **90% CHANCE,** the match payoffs will be:

Player B1's match payoff = 700 tokens + X tokens – 900 tokens Player B2's match payoff = 700 tokens + X tokens – 420 tokens

With a **10% CHANCE,** the match payoffs will be:

Player B1's match payoff = 700 tokens + X tokens – 0 tokens Player B2's match payoff = 700 tokens + X tokens – 420 tokens

POSSIBLE OUTCOME 4: BOTH PLAYERS DECIDE TO REPORT THE ACT

- **IF PLAYER B1 REPORTS FIRST**
  - **ONLY PLAYER B1 GETS A FINE REDUCTION**: Instead of paying a fine equal to 900 tokens, Player B1 **always** pays only **420 tokens**.
  - **PLAYER B2'S CHANCE OF PAYING A FINE EQUAL TO 900 TOKENS IS 100%**: Player B2 **always** pays a fine equal to **900 tokens.**

Hence, the match payoffs are as follows.

With a **100% CHANCE,** the match payoffs will be:

Player B1's match payoff = 700 tokens + X tokens – 420 tokens
Player B2's match payoff = 700 tokens + X tokens – 900 tokens

- **IF PLAYER B2 REPORTS FIRST**
  - **ONLY PLAYER B2 GETS A FINE REDUCTION**: Instead of paying a fine equal to 900 tokens, Player B2 **always** pays only **420 tokens**.
  - **PLAYER B1'S CHANCE OF PAYING A FINE EQUAL TO 900 TOKENS IS 100%**: Player B1 **always** pays a fine equal to **900 tokens**.

Hence, the match payoffs are as follows.

With a **100% CHANCE,** the match payoffs will be:

Player B1's match payoff = 700 tokens + X tokens – 900 tokens
Player B2's match payoff = 700 tokens + X tokens – 420 tokens

- **IF BOTH PLAYERS REPORT AT THE SAME TIME**
  - **<u>EACH PLAYER GETS A FINE REDUCTION WITH 50% CHANCE</u>**: Instead of paying a fine equal to 900 tokens, each player pays only **420 tokens** with **50% chance**.
  - **<u>EACH PLAYER'S CHANCE OF PAYING A FINE EQUAL TO 900 TOKENS IS 50%:</u>** Each player pays a fine equal to **900 tokens** with **50% chance**.

Hence, the match payoffs are as follows.

With a **50% CHANCE,** the match payoffs will be:

Player B1's match payoff = 700 tokens + X tokens – 420 tokens
Player B2's match payoff = 700 tokens + X tokens – 900 tokens

With a **50% CHANCE,** the match payoffs will be:

Player B1's match payoff = 700 tokens + X tokens – 900 tokens
Player B2's match payoff = 700 tokens + X tokens – 420 tokens

EXERCISES

Suppose that the number of tokens that each player gets **IF BOTH PLAYERS AGREE TO JOINTLY COMMIT THE ACT** is equal **to X tokens.**

Nine exercises, based on the possible outcomes, are presented below. Please fill the blanks.

Exercise 1

Suppose that **BOTH PLAYERS AGREE TO JOINTLY COMMIT THE ACT**. Then, each player gets ____ tokens. Suppose also that **Player B1** decides **NOT TO REPORT** the act and **Player B2** decides **NOT TO REPORT** the act.

The **MATCH PAYOFFS (IN TOKENS)** are as follows.

Player B1:

Chance	Payoff	Chance	Payoff

Player B2:

Chance	Payoff	Chance	Payoff

Exercise 2

Suppose that **BOTH PLAYERS AGREE TO JOINTLY COMMIT THE ACT**. Then, each player gets ____ tokens. Suppose also that **Player B1** decides **TO REPORT** the act and **Player B2** decides **NOT TO REPORT** the act.

The **<u>MATCH PAYOFFS</u> (IN TOKENS)** are as follows.

Player B1:

Chance	**Payoff**	**Chance**	**Payoff**

Player B2:

Chance	**Payoff**	**Chance**	**Payoff**

Exercise 3

Suppose that **BOTH PLAYERS AGREE TO JOINTLY COMMIT THE ACT**. Then, each player gets ____ tokens. Suppose also that **Player B1** decides **NOT TO REPORT** the act and **Player B2** decides **TO REPORT** the act.

The **<u>MATCH PAYOFFS</u> (IN TOKENS)** are as follows.

Player B1:

Chance	**Payoff**	**Chance**	**Payoff**

Player B2:

Chance	**Payoff**	**Chance**	**Payoff**

Exercise 4

Suppose that **BOTH PLAYERS AGREE TO JOINTLY COMMIT THE ACT**. Then, each player gets ____ tokens. Suppose also that **Player B1** decides **TO REPORT** the act and **Player B2** decides **TO REPORT** the act, and that **Player B1 IS THE FIRST TO REPORT**.

The **MATCH PAYOFFS (IN TOKENS)** are as follows.

Player B1:

Chance	Payoff	Chance	Payoff

Player B2:

Chance	Payoff	Chance	Payoff

Exercise 5

Suppose that **BOTH PLAYERS AGREE TO JOINTLY COMMIT THE ACT**. Then, each player gets ____ tokens. Suppose also that **Player B1** decides **TO REPORT** the act and **Player B2** decides **TO REPORT** the act, and that **Player B2 IS THE FIRST TO REPORT**.

The **MATCH PAYOFFS (IN TOKENS)** are as follows.

Player B1:

Chance	Payoff	Chance	Payoff

Player B2:

Chance	Payoff	Chance	Payoff

Exercise 6

Suppose that **BOTH PLAYERS AGREE TO JOINTLY COMMIT THE ACT**. Then, each player gets ____ tokens. Suppose also that **Player B1** decides **TO REPORT** the act and **Player B2** decides **TO REPORT** the act, and that **BOTH PLAYERS REPORT AT THE SAME TIME**.

The **<u>MATCH PAYOFFS</u> (IN TOKENS)** are as follows.

Player B1:

Chance	**Payoff**	**Chance**	**Payoff**

Player B2:

Chance	**Payoff**	**Chance**	**Payoff**

Exercise 7

Suppose that **Player B1 AGREES TO JOINTLY COMMIT THE ACT** and **Player B2 DOES NOT AGREE TO JOINTLY COMMIT THE ACT**.

The **<u>MATCH PAYOFFS</u> (IN TOKENS)** are as follows.

Player B1:

Chance	**Payoff**	**Chance**	**Payoff**

Player B2:

Chance	**Payoff**	**Chance**	**Payoff**

Exercise 8

Suppose that **Player B1 DOES NOT AGREE TO JOINTLY COMMIT THE ACT** and **Player B2 AGREES TO JOINTLY COMMIT THE ACT**.

The **<u>MATCH PAYOFFS</u> (IN TOKENS)** are as follows.

Player B1:

Chance	**Payoff**	**Chance**	**Payoff**

Player B2:

Chance	**Payoff**	**Chance**	**Payoff**

Exercise 9

Suppose that **NEITHER PLAYER AGREES TO JOINTLY COMMIT THE ACT**.

The **<u>MATCH PAYOFFS</u> (IN TOKENS)** are as follows.

Player B1:

Chance	**Payoff**	**Chance**	**Payoff**

Player B2:

Chance	**Payoff**	**Chance**	**Payoff**

SESSION PAYOFF

The game earnings in tokens will be equal to the **PAYOFF FOR THE ACTUAL MATCH**. The game earnings in dollars will be equal to: (game earnings in tokens)/29 (29 tokens = 1 dollar). The session payoff will be equal to the game earnings in dollars plus the $10 participation fee.

GAME SOFTWARE

The game will be played using a computer terminal. You will need to enter your decisions by using the mouse. In some instances, you will need to wait until the other players make their decisions before moving to the next screen. Please **BE PATIENT**. There will be a box, displayed in the upper right-hand side of your screen, which indicates the "Match Number," "Your Role," and "Your Balance."

Please press the **NEXT >>** button to move to the next screen. **<u>DO NOT TRY TO GO BACK TO THE PREVIOUS SCREEN AND DO NOT CLOSE THE BROWSER</u>**: The software will stop working.

Next, the **5 PRACTICE MATCHES** will begin. After that, the **ACTUAL MATCH** will be played. **YOU CAN CONSULT THESE INSTRUCTIONS AT ANY TIME DURING THE EXPERIMENT**.

THANKS FOR YOUR PARTICIPATION IN THIS STUDY!!

<u>PLEASE GIVE THIS MATERIAL TO THE EXPERIMENTER AT THE END OF THE EXPERIMENT</u>

Department of Economics, University of Alberta
Working Paper Series

2018-12: Imperfect Competition in Electricity Markets with Renewable Generation: The Role of Renewable Compensation Policies – **Brown, D., Eckert, A.**
2018-11: The Extensive Margin of Trade and Monetary Policy – Imura, Y., **Shukayev, M.**
2018-10: Macroeconomic Conditions and Child Schooling in Turkey – **Gunes, P., Ural Marchand, B.**
2018-09: Employing Simple Cost-Sharing Policies to Motivate the Efficient Implementation of Distributed Energy Resources – **Brown, D.,** Sappington, D.
2018-08: Sequential Majoritarian Blotto Games – **Klumpp, T.,** Konrad, K.
2018-07: Why are Refugee Children Shorter than the Hosting Population? Evidence from Camps Residents in Jordan – Rashad, A., **Sharaf, M.,** Mansour, E.
2018-06: Optimal Law Enforcement with Ordered Leniency – **Landeo, C.,** Spier, K.
2018-05: Price-Quality Competition in a Mixed Duopoly – **Klumpp, T., Su, X.**
2018-04: “Causes of Sprawl”: A (Further) Public Finance Extension – **McMillan, M.**
2018-03: Financially-Constrained Lawyers: An Economic Theory of Legal Disputes – **Landeo, C.,** Nikitin, M.
2018-02: Information and Transparency in Wholesale Electricity Markets: Evidence from Alberta – **Brown, D., Eckert, A.,** Lin, J.
2018-01: Does a Discount Rate Rule Ensure a Pension Plan Can Pay Promised Benefits without Excessive Asset Accumulation? – **Landon, S., Smith, C.**
2017-13: Income Inequality and Violence Against Women: Evidence from India – Rashad, A., **Sharaf, M.**
2017-12: The Local Effects of the Texas Shale Boom on Schools, Students, and Teachers – **Marchand, J.,** Weber, J.
2017-11: Self-Sabotage in the Procurement of Distributed Energy Resources – **Brown, D.,** Sappington, D.
2017-10: Public Private Competition – **Klumpp, T., Su, X.**
2017-09: Testing for State-Dependent Predictive Ability – **Fossati, S.**
2017-08: Default Risk, Productivity, and the Environment: Theory and Evidence from U.S. Manufacturing – **Andersen, D.**
2017-07: Does Maternal Employment Affect Child Nutrition Status? New Evidence from Egypt – Rashad, A., **Sharaf, M.**
2017-06: The Effect of Default Rates on Retail Competition and Pricing Decisions of Competitive Retailers: The Case of Alberta – **Brown, D., Eckert, A.**
2017-05: Optimal Procurement of Distributed Energy Resources – **Brown, D.,** Sappington, D.
2017-04: The Impact of Schooling Intensity on Student Learning: Evidence from a Quasi-Experiment – Andrietti, V., **Su, X.**
2017-03: The Voting Rights of Ex-Felons and Election Outcomes in the United States – **Klumpp, T.,** Mialon, H., Williams, M.
2017-02: Does the Design of a Fiscal Rule Matter for Welfare? – **Landon, S., Smith, C.**
2017-01: Carbon Pricing with an Output Subsidy under Imperfect Competition: The Case of Alberta's Restructured Electricity Market – **Brown, D., Eckert, A., Eckert, H.**