The People's Hired Guns?
Experimentally Testing the
Inclination of Prosecutors
to Abuse the Vague
Definition of Crimes

Christoph Engel
Alicja Pluta

MAX PLANCK SOCIETY

The People's Hired Guns?

Experimentally Testing the Inclination of Prosecutors to Abuse the Vague Definition of Crimes

Christoph Engel / Alicja Pluta

July 2011

The People's Hired Guns?
Experimentally Testing the Inclination of Prosecutors to Abuse the
Vague Definition of Crimes[*]

Christoph Engel / Alicja Pluta

Abstract: Legal realists expect prosecutors to be selfish. If they get the defendant convicted, this helps them advance their careers. If the odds of winning on the main charge are low, prosecutors have a second option. They can exploit the ambiguity of legal doctrine and charge the defendant for vaguely defined crimes, like "conspiracy". We model the situation as a signalling game and test it experimentally. If we have participants play the naked game, at least a minority plays the game theoretic equilibrium and use the vague rule if a signal indicates that the defendant is guilty. This becomes even slightly more frequent if a misbehaving defendant imposes harm on a third participant. By contrast if we frame the situation as a court case, almost all prosecutors take the signal at face value and knowingly run the risk of loosing in court if the signal was false. Our experimental prosecutors behave like textbook legal idealists, and follow the urge of duty.

---

1. Introduction

The institution of the prosecutor in U.S. criminal law holds a great deal of esteem. Prosecutors are fighting for a noble cause: the People's desire to see criminals convicted, for the sake of deterring future crime, but also to restore justice and to respect victims' sufferings. Yet for prosecutors, loosing in court is quite likely, given the standard of proof in criminal matters is "beyond a reasonable doubt". The legal order is much more willing to accept that a guilty defendant is acquitted, rather than tolerating that an innocent is convicted. If there is serious doubt, the presumption of innocence trumps society's wish to convict criminals. Consequently the law expects prosecutors to endure frequent failure, given prosecutors are obliged to go to court if it is only "sufficiently likely" that the defendant will eventually be convicted.[1]

Legal orders differ in the way in which they impose respect for the presumption of innocence on prosecutors. Some legal orders, like the U.S., basically content themselves with giving defendants and their attorneys' rights. This explains why, in the U.S., prosecutors as individuals are largely immune to legal action (Brink 2009). Other legal orders go further. For instance, the United Nations, based on the Universal Declaration of Human Rights, stipulate that all member states should guarantee the impartiality of prosecutors.[2] In Germany, prosecutors may even be prosecuted themselves if they are guilty of perversion of justice.[3]

Yet is it realistic to expect that prosecutors will live up to these heroic standards? Criminal procedure is organized as a tournament. Prosecutors have the right of initiative. They select charges and choose which evidence is presented not only during the trial but already during the investigation. Psychologically, prosecutors are most likely committed to their side of the case. The litigant spirit may be exacerbated by the feeling that the counsel for the defense, or the defendant himself, have behaved strategically. Straightforward incentives combine with the psychology of the conflict. Prosecutors have a better chance to be promoted if they "win", i.e. if the defendant is convicted. They may come under pressure from the district attorney who wants to be reelected, or from the press who urges prosecutors "to be tough on crime". If prosecutors give in to such temptations, they become "the People's hired guns." In this study, we investigate empirically to which degree they do, and to which degree, in contrast, their sense of duty and responsibility prevails.

A subtle but potentially very effective strategy for changing the odds of winning relies on the precision of the law. While crimes like fraud, embezzlement, insider trading or forgery are reasonably well defined in legal doctrine, another set of criminal offenses is laid down in very vague terms. Prominent examples include "obstruction of justice" (18 U.S.C.S. § 1503), "conspiracy" (21 U.S.C. § 846) and "false statements" (18 U.S.C. § 1001). Observers have repeatedly suspected prosecution to replace a charge for a more narrowly defined crime by one of these vaguely circumscribed delicts, hoping that the ambiguity of these terms would help them get a defendant convicted who would otherwise escape conviction.

---

[1]   See for instance § 152 II German Code of Criminal Procedure: "Prosecution is […] obliged to accuse the defendant for all crimes provided evidence is sufficiently suggestive", or in the German original „Sie [prosecution] ist, soweit nicht gesetzlich ein anderes bestimmt ist, verpflichtet, wegen aller verfolgbaren Straftaten einzuschreiten, sofern zureichende tatsächliche Anhaltspunkte vorliegen".

[2]   Guidelines on the Role of Prosecutors Adopted by the Eighth United Nations Congress on the Prevention of Crime and the Treatment of Offenders, Havana, Cuba, 27 August to 7 September 1990, http://www2.ohchr.org/english/law/pdf/prosecutors.pdf.

[3]   Under § 339 Criminal Code, see BGHSt 32, 357.

In the field, it is difficult to show whether this suspicion has any merit at all, and next to impossible to measure how frequently this practice is used. One would at best spot a few of the most salient cases, which could just be exceptions that prove the rule of prosecutorial impartiality. At best, one would show correlations between, say, the fact that prosecutors are elected in a jurisdiction and the frequency of convictions for certain crimes, without being able to prove causality. We therefore create a prosecution-like situation in the lab. By our design, we create a tension between incentives and duty. Under tightly controlled conditions, we measure how often mock prosecutors are willing to trade impartiality for success, by exploiting the vague wording of an imprecise rule. Through our treatment manipulations, we disentangle the mere force of incentives, the moderating effect of fighting for a victim, and the additional impact of framing the issue as a court conflict.

To maintain full control, we have designed our experiment as a (sequential) game. We have chosen parameters such that this game has a unique equilibrium in pure strategies. In equilibrium, the prosecutor does not bring the case if she has received a signal indicating that the would-be defendant is likely to be innocent. If she receives the opposite signal, in equilibrium she uses a vaguely defined charge. If she does, her probability of winning in court is 50%, irrespective of the true action of the defendant. By contrast, if she charges the defendant for the actual crime, i.e. selects the clear-cut rule, she loses in court in case this signal was wrong. However, this choice is off the equilibrium path. Anticipating the prosecutor's choice, in equilibrium the defendant always abides by the law.

Our results look very different. Defendants break the law quite frequently. If they receive a signal to the effect that the defendant likely was well-behaved, a minority of prosecutors charges them under the vague rule. If they receive the opposite signal, again only a minority uses the vague rule, while the large majority relies on the precise rule, although this is more risky and off the equilibrium path. We do not find any significant treatment differences for defendants. By contrast, it already is clear from the descriptive statistics that prosecutors only very rarely invoke the vaguely defined crime when we call a spade a spade, i.e. when we label the defendant a manager, call his action honesty or fraud, and single out the other player as a prosecutor. The defendant is most frequently charged for the vaguely defined crime when we keep all labels neutral, but invite a third person whose earnings are reduced if the first agent behaves improperly. Prosecutors use vague terms slightly less frequently if they just play the naked two-person game.

Any experiment in law is the result of a tradeoff. It is set up to generate evidence on a problem of legal policy. Yet to make this evidence valid, it must abstract from many features that are likely to matter in the field. Our experiment is no exception to this rule. In the concluding section, we discuss these limitations. We are, however, convinced that our experiment addresses the key feature of the issue in the field: is a person who is entrusted with prosecuting perpetrators willing to rely on vaguely defined charges if this is optimal for her payoff? We have a clear negative answer to this question. Even if the situation is neutrally framed, only a minority of prosecutors uses the vague rule. This number becomes extremely small if we let participants know that they are in the role of a prosecutor. By our design, including a number of post-experimental tests to be reported below, we can rule out alternative explanations. At least the treatment difference must be attributed to prosecutors' sense of responsibility. If they knowingly hold the public office of a prosecutor, they suppress personal incentives and listen to the call of prosecutorial duty. At least in the lab prosecutors are not the People's hired guns.

In the next section, we develop the legal research question our experiment is meant to answer. We then introduce the design of the experiment (section 3) and derive the hypotheses to be tested (section 4). We report the results (section 5) and conclude with discussion (section 6).

2. Legal Research Question

Historically, criminal procedure had been inquisitorial. The judge not only held power to adjudicate. He also was the investigator. It has been one of the major advances of rule of law to separate these functions. In modern (U.S.) criminal procedure the jury is responsible for deciding guilt or innocence and the judge is responsible for sentencing. There is a separate authority representing the government's interest in convicting criminals. It is the responsibility of the prosecution to apprehend alleged criminals, to spot incriminating evidence, and to fight for the People's cause.

Through separating roles, the law acknowledges the inherently partisan character of prosecution. This is not to say, though, that the law just cares about convictions. The presumption of innocence is the cornerstone of criminal procedure. False positives, i.e. convicting an innocent, carry much more weight than false negatives, i.e. acquitting a guilty defendant (leading case: *Addington v. Texas*, 441 U.S. 418, 422 (1979)). The standard of proof is strict. The defendant may only be convicted if her guilt has been established beyond reasonable doubt (see e.g. Pa. SSJI (Crim) 7.01). This translates into rules about prosecutor impartiality (e.g. Rule 3.8 New Jersey Rules of Professional Conduct) or neutrality (Green and Zacharias 2004).

Observers, and prosecutors themselves, are divided over the question to which degree prosecutors live up to the normative expectation of being "litigant but impartial" (Yaroshefsky 1999). There is casual empiricism of prosecutors being unduly wedded to the goal of conviction (*United States v. Shaygan*, No. 08-20112-CR-GOLD-MCALILEY, 2009 WL 980289 (S.D. Fla. Apr. 9, 2009); *Pottawattamie County, Iowa v. McGhee*, 547 F.3d 922 (8th Cir. 2008) (further see the quotes from interviews with prosecutors in Yaroshefsky 1999:945, 949 f.; and see the case reported by Hoeffel 2004; also see the cases reported in Brink 2009). A considerable number of criminal convictions have been reversed at a later stage, recently in particular due to the availability of DNA evidence (Gross, Jacoby et al. 2004; Garrett 2008). Scholars have wondered to which degree these wrongful convictions were caused by prosecutor zeal (Brink 2009).[4]

Were prosecutors inclined to bend the law in favour of conviction, there would be ample opportunity. Police investigations could be biased (Baldwin 1993; Kassin, Meissner et al. 2005). Prosecutors might choose which cases to litigate (Priest and Klein 1984), what evidence to present (Hoeffel 2004) and when (Podgor 1999), whether to remunerate the defendant for her cooperation by proposing a reduced sentence (Lee 1997; Podgor 1999), which judgements to enforce (Heminway 2002), and which defendants to push into a plea bargain, even if they might well be innocent (Hessick and Saujani 2001).

A particularly important possibility for changing the odds of winning rests in the power of prosecution to determine the charge (Podgor 1999; Gordon and Huber 2002). A prominent case is the one of Martha Stewart. Many observers believe prosecutors actually wanted to get

---

[4] Innocence Project, *Understand the Causes: Government Misconduct*, http://www.innocenceproject.org/understand/Government-Miscond uct.php attributes 44.6 % of DNA based reversals to prosecutor misconduct.

her for insider trading, but shifted to a charge of conspiracy to make sure they would win in court (Heminway 2002; Seigel and Slobogin 2004; Moohr 2006; Szott 2006). In such cases, the power balance is tilted in favour of prosecution by the fact that the offense is of a very general nature, and only vaguely defined in doctrine. This channel is the object of the present study.

Prosecutors might have reason to exploit this channel. Winning their cases helps them advance their careers (Boylan and Long 2005; Rasmusen, Raghav et al. 2009), be re-elected (Brink 2009), respond to media (Brink 2009) or social pressure (Moohr 2006), uphold their self-esteem (Crank, Flaherty et al. 2007), or retaliate against the defendant for exercising rights of defense (*Blackledge v. Perry*, 417 U.S. 21, 28-29 (1974)). These incentives to be excessively partial are not counteracted by the threat of sanctions if prosecutors are largely immune to prosecution for unduly prosecuting defendants, as is the case in the U.S. (Zacharias 2000; Brink 2009).

Experiments have been an established criminological tool for quite some time (for overviews see Farrington 2003; Farrington and Welsh 2005; Farrington 2006; Farrington and Welsh 2006). Yet lab experiments, in particular those designed in the spirit of experimental economics, are still rare (for an example involving an experimental public authority see Engel and Irlenbusch 2010). To the best of our knowledge, our research question has not been tested experimentally, nor in any other rigorous empirical way.

3. Experimental Design

Essentially, we are interested in a tension between incentives and duty. In our experiment we therefore create a situation in which a payoff maximising agent exploits the vagueness of the criminal code and charges the defendant for an auxiliary crime. She does so, despite the fact that, with this charge, the odds of winning in court are unrelated to the defendant's action. In the interest of generating clear evidence, our design thus radicalises the degree of ambiguity, and makes the charge orthogonal to the crime. Casual empiricism suggests that, sadly, not so rarely this is not far off from the situation in the field. To disentangle motives, we have three treatments. In the *baseline*, we just have two neutrally labelled players who are exposed to the naked incentive structure. In the *harm* treatment, we keep the neutral wording, but add a third inactive player. Depending on the first player's action, this passive player is either worse or equally well off. Comparing the *baseline* with the *harm* treatment, we may investigate whether choosing the vaguely defined charge is driven by the desire to avenge a victim. In the final *frame* treatment, we lay the situation in which we are interested open. We now call the first agent a manager, the passive agent a shareholder, and the second mover a prosecutor. The manager may now choose between "honesty" and "fraud". The calculation of payoffs is explained by the respective tension between individual and social benefits. Comparing the *frame* treatment with the other two treatments, we may explore to which degree holding a public office changes behaviour. Specifically, the normative expectation of the law is this: if the signal is good, the prosecutor should not take the defendant to court. If the signal is bad, she should charge the clear-cut rule. If we find that these choices are significantly more frequent in the *frame* treatment, compared with the *baseline*, we see the effect of prosecutorial duty. If we find that these choices are also significantly more frequent than in the *harm* treatment, we can rule out that prosecutors simply desired to help the victim, but were indeed motivated by the specific duties that go with holding the office of a prosecutor.
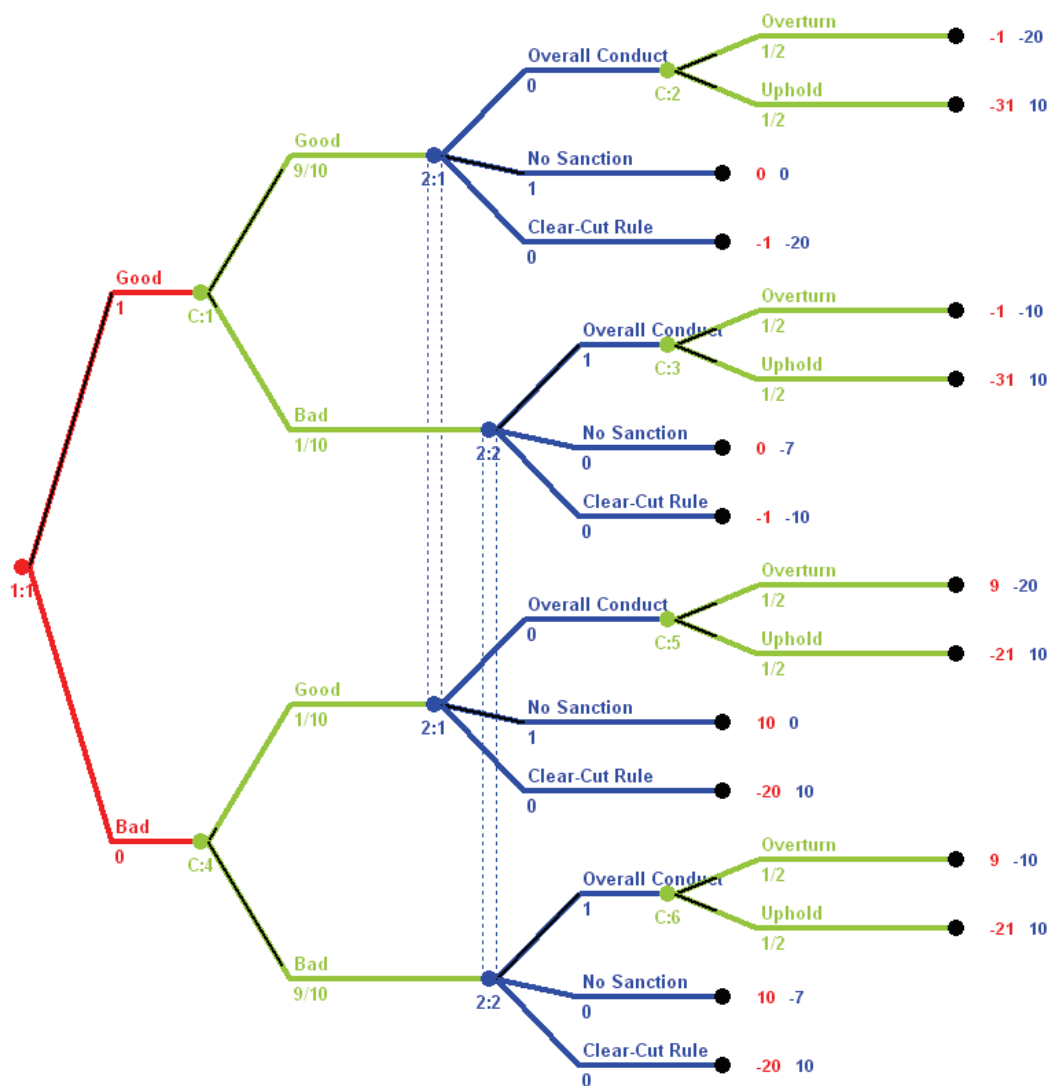
Figure 1
Game Tree

Note: The moves of all three players are described in detail in this tree. In the first move the managers decides whether or not to commit fraud. Then Nature signals to the prosecutor with 90% accuracy if this has taken place. Next, the prosecutor decides between her three options. Finally, Nature once more plays a 50% draw if the overall conduct choice was selected. The first payoff is for the first player, i.e. the manager (and appears in red), the second payoff is for the second player, i.e. the prosecutor (and appears in blue).

We model the prosecutor's choice by a game of two decision nodes and two moves by Nature, Figure 1. In the first stage, the later potential defendant decides whether to break the law or not. If he does, and if the prosecutor does not take him to court, he is better off. Yet in the *harm* and *frame* treatments, if the manager behaves dishonestly, the shareholder incurs a loss. The loss for the shareholder is bigger than the gain for the manager, so that dishonesty is not only individually, but also socially harmful. The prosecutor may not directly observe the manager's action. She only receives a signal. We inform all participants that this signal is positively correlated with the manager's action. In nine out of ten cases, the signal is correct. As is standard in game theory, in our game tree the signal is represented by a draw of Nature.

Knowing the signal, the prosecutor chooses between three options. She may do nothing, or charge the defendant for one of two crimes. In the *baseline* and in the *harm* treatment, we do not label these actions. In the *frame* treatment, we call the first law "clear-cut rule", and the

second "overall conduct". In terms of incentives for the prosecutor, the two crimes differ by the relevance of the manager's action. If the prosecutor charges the clear-cut rule, her payoff is high if the manager has indeed behaved dishonestly. Yet her payoff is low if he actually has been honest. By contrast, the manager's behaviour is irrelevant for payoffs if the prosecutor charges for overall conduct. Then with 50% probability he receives a high payoff, and with 50% probability his payoff is low. In the parlance of game theory, we thus have a second move of Nature that decides about payoffs. Another method through which the doctrinal difference between the clear-cut rule and a charge for overall conduct is reflected is the determinants of managers' payoffs. If the prosecutor uses the clear-cut rule, the court is assumed to successfully investigate the manager's behaviour, and sanction him only if he actually committed fraud. On the other hand, when the overall conduct rule is applied, the manager has a 50% chance of being convicted. He receives a sanction that is unrelated to his behaviour; if he committed fraud , either payoff is increased by 10 tokens.

In the instructions (see the Appendix), we do not only explain the game, but also the calculation of payoffs. The payoff of the manager is increased by 10 tokens if he behaves dishonestly. If the prosecutor chooses no sanction, intermediate are also final payoffs. If the prosecutor chooses the clear-cut rule and if the manager was well-behaved, just one token is subtracted from the manager's payoffs. In the *frame* treatment we explain that this token is meant to reflect trial cost. If the prosecutor chooses the clear-cut rule and the manager was not well-behaved, his first stage earnings are reduced by 30 tokens. In the *frame* treatment, we call this a sanction. If the prosecutor chooses overall conduct, the manager's payoff depends on the second draw of Nature. If this draw is to the manager's benefit, and if he had been well-behaved, he again loses one token. In the *frame* treatment, we once more motivate the reduction by trial cost. If the manager had been dishonest but is lucky at trial, he keeps the extra 10 tokens, but also loses one token for trial cost, resulting in a payoff of 9 tokens. If Nature's draw is to the manager's detriment, he receives 30 tokens less. In the *frame* treatment, we call this reduction a sanction.

The prosecutor's payoff depends on the manager's actual behaviour, the signal, and the prosecutor's choice. If the manager has been well-behaved and this has also been signalled to the prosecutor, she receives a payoff of zero for not taking the manager to court. If the manager has been dishonest and this has been signalled to the prosecutor, she receives an additional 10 tokens if she charges the clear-cut rule. The difference between the two cases is explained as the reward prosecutors receive for correctly prosecuting perpetrators. If the manager has been well-behaved, but the prosecutor has received a wrong signal and charges the clear-cut rule, 20 tokens are subtracted from her payoff. This is called a malus for accusing an innocent. If the manager has been dishonest, but the prosecutor has received a bad signal and does not take action, she loses seven tokens. In the instructions of the *frame* treatment, we explain this as punishment for not accusing a potential offender. If the prosecutor charges overall conduct, neither the manager's true behaviour nor the signal has payoff relevance. With 50% probability the charge is upheld, which gives the prosecutor a premium of 10 tokens. This is the same payoff as if the manager has rightly been charged under the clear-cut rule, and it is as good as the best possible outcome for the prosecutor. Hence irrespective of the manager's true behaviour and of the information the prosecutor has received, she has a 50% chance to receive the maximum payoff. We justify this as a reward for punishing correctly. With the same probability the charge is dismissed, in which case the prosecutor loses 10 tokens if the signal is bad, and 20 tokens if the signal is good. We call this a punishment for abuse of power. This second outcome reflects the situation is which the defendant is successful in convincing the court that the charge for overall conduct is unfounded. The difference between a loss of 10 and of 20 tokens is meant to capture the

condition that prosecutorial behaviour is particularly egregious if a manager is charged with overall conduct, despite the fact that the publicly observable signal was good.

The structure of the game and all payoffs are common knowledge. Yet to make it easier for our participants to understand the structure of the game, we do not show them the original game tree, but use the role-specific representations reported in the appendix. Since negative payoffs are difficult to handle in the lab, we add 60 tokens to all payoffs. In the instructions, we introduce this as an endowment. Through the endowment, the minimum payoff for the manager is 29 tokens, and the minimum payoff for the prosecutor is 40 tokens. The maximum is 70 tokens for both players. In the *harm* and *frame* treatments, if the manager is well-behaved, the shareholder keeps her endowment of 40 tokens. If the manager misbehaves, her payoff is reduced to 28 tokens. We announce 30 repetitions, yet every period with a new, randomly chosen partner. Following the procedure that is standard in the experimental literature (see e.g. Charness 2000; Montero, Sefton et al. 2008), we assign participants to matching groups, but do only tell them they will be rematched every period, not that matching groups have limited size.[5] This procedure is meant to guarantee independent observations, without inducing participants to try to second guess group composition. In each treatment, we had one or two smaller matching groups, due to the fact that invited participants did not show up. This was not communicated to participants, and therefore should not have biased results.

In each round, and before they choose their own action, we ask prosecutors to estimate the number of managers who have been well-behaved. If they get this number exactly right, they receive an extra 3 tokens. At the end of each round, the computer informs participants in all roles about the decision of the manager, the signal, the decision of the prosecutor, and earnings of all participants. Feedback on beliefs is withheld until the end of the experiment. From the second period on, feedback of all previous periods is provided on one screen.

After the end of the main experiment, we have participants take two additional tests. We measure risk aversion, using the incentivized procedure introduced by (Holt and Laury 2002). We also elicit DOSPERT risk tests (Blais and Weber 2006).

4. Hypotheses

If both the manager and the prosecutor hold textbook preferences, and expect the other player to hold such preferences as well, the equilibrium of the stage game predicts their choices. We have defined parameters such that the game has a unique equilibrium in pure strategies. In this equilibrium, the prosecutor reacts to a good signal by not taking the manager to court, and she reacts to a bad signal by charging for overall conduct. Anticipating this reaction, the manager is well-behaved. To see this, it is easiest to represent the game in normal form.

---

[5]   Matching groups are multiples of 2 in the *baseline*, and multiples of 3 in the remaining treatments. In the *baseline* we have 7 matching groups of 6 and, due to the fact that invited participants did not show up, one matching group of 4. In the two treatments we each time have 4 matching groups of 9 and, again since participants did not show up, 2 matching groups of 6.

	11		12		13		21		22		23		31		32		33	
1	-16	$-\frac{8}{5}$	$-\frac{72}{5}$	-26	$-\frac{29}{5}$	-11	$-\frac{8}{5}$	0	0	$-\frac{7}{10}$	$-\frac{1}{10}$	-1	$-\frac{5}{5}$	-18	$-\frac{3}{10}$	$-\frac{187}{10}$	-1	-19
2	-6	$-\frac{2}{5}$	$-\frac{42}{5}$	$-\frac{34}{5}$	$-\frac{83}{5}$	$-\frac{7}{5}$	$-\frac{22}{5}$	0	0	10	$-\frac{83}{10}$	-17	$-\frac{87}{5}$	1	7	$-\frac{53}{10}$	-20	10

Figure 2
Game in Normal Form

Note: The game presented in the game tree in Figure 1 is described here in Normal Form. The manager draws first and has the option to behave honestly (1) or to commit fraud (2). The prosecutor has three options: charging overall conduct (1), no sanction (2) or the clear-cut rule (3). Since the prosecutor only sees the signal, she must condition her choice on the signal. The prosecutor therefore must adopt one of 9 possible strategies, where the first number defines her reaction to a good signal, and the second her reaction to a bad signal. Hence "21" means that she remains inactive if the signal is good, and that she charges overall conduct if the signal is bad.

The manager has just two options. He either is well or badly behaved. By contrast, given she only sees the signal, the prosecutor must choose for every signal which of her three options to take. Figure 2 assembles her expected payoffs, given the signal and its probability to be correct. Off this table, one sees that the prosecutor's best response is (21), i.e. "no action" (2) if the signal is good (0 is better than all the other payoffs, that are negative), and "overall conduct" (1) if the signal is bad, provided the manager is well behaved (1). By contrast, if the manager behaves dishonestly (2), then the prosecutor's best response is using the clear-cut rule, irrespective of the character of the signal (33), since in expectation this gives her a payoff of 10, more than any other strategy. Yet provided the prosecutor chooses this strategy, if he behaves dishonestly the manager expects the worst of all payoffs, -20. If he behaves properly, he still expects a loss of -8/5, and is far away from his best outcome 10. But if both parties are fully rational, this is much better than -20, and therefore the unique equilibrium of the game. Whether there is a shareholder or not and how the situation is framed does not affect players' payoffs. This leads to

H₁: In all treatments, if the manager and the prosecutor are fully prevoyant, risk-neutral, money-maximising agents, and expect the other player to hold the same type of preferences, the prosecutor chooses no sanction if the signal is good, and overall conduct if the signal is bad. The manager never behaves dishonestly.

In equilibrium, deterrence is perfect. The threat of punishment is never executed since there is no crime in the first place. Managers always incur a loss. Managers may be tempted to deviate from this equilibrium. If prosecutors treat such deviations as behaviour off the equilibrium path, we have

H₂: Managers sometimes behave dishonestly. Prosecutors react by charging overall conduct. There are no treatment differences.

In equilibrium, prosecutors do not expect a positive profit. They may be tempted to gamble, meaning that they react to a good signal with charging overall conduct. While this is not an equilibrium, it may seem attractive given that, in expected values, the payoff is the same as if they do not sanction the manager. This leads to

H₃: Prosecutors charge overall conduct if they see a good signal. There are no treatment differences.

Prosecutors may anticipate or observe that managers misbehave and play their best responses rather than the equilibrium, which implies

H₄: Prosecutors react to a good signal by not taking the manager to court, and to a bad signal by using the clear-cut rule, irrespective of treatment.

Managers may be reluctant to impose harm on shareholders. Prosecutors may feel obliged to react more aggressively to manager dishonesty if they are not only putting their own money at risk, but if the shareholder suffers as well. This might follow from indirect reciprocity. This yields

H₅: Managers behave more honestly in the *harm* and *frame* treatments, compared to the *baseline*. In the *harm* and *frame* treatments prosecutors are more likely to charge the clear-cut rule when receiving a bad signal.

Managers may feel morally inhibited from behaving dishonestly if they are assigned the role of a manager entrusted with a shareholder's money, and if their detrimental behaviour is called fraud. Their sense of duty and justice may induce prosecutors to use the clear-cut rule, which conditions the sanction on the manager's actual behaviour, rather than overall-conduct which imposes a random sanction on the manager. It follows

H₆: Managers are behaving more honestly in the *frame* treatment. In this treatment, prosecutors are more likely to use the clear-cut rule when receiving a bad signal.

5. Results

The experiment was conducted at Columbia University in 2011. A total of 142 students of different majors participated in one of 9 sessions, 3 for each treatment. Both participant-participant and participant-experimenter anonymity were guaranteed. Each session lasted approximately one and a half hours. The experiment was programmed in zTree (Fischbacher 2007). All periods were paid out. In the *baseline*, managers on average earned 18.24 $, and prosecutors earned 20.39 $. In the *harm* treatment, managers earned 17.84 $, prosecutors made 19.54 $, and shareholders made 18.85 $. In the *frame* treatment, managers had an average profit of 18.62 $, prosecutors had 21.24 $, and shareholders had 19.57 $.

a) Managers

In clear opposition to the game-theoretic prediction, in all treatments a substantial fraction of managers behaves dishonestly. Actually, in the *baseline* and the *harm* treatments, no more than 2, and in the *frame* treatment no more than 3 participants behave properly in all of the 30 periods of the game, Figure 3. In opposition to **H₁,** managers clearly violate the equilibrium prediction. The most plausible explanation is the larger gain in case prosecutors do not take action, or charge overall conduct but the charge is overturned. Seemingly managers are tempted by the opportunity of a higher payoff. As predicted by **H₂**, almost all managers at least sometimes give in to temptation and behave badly.
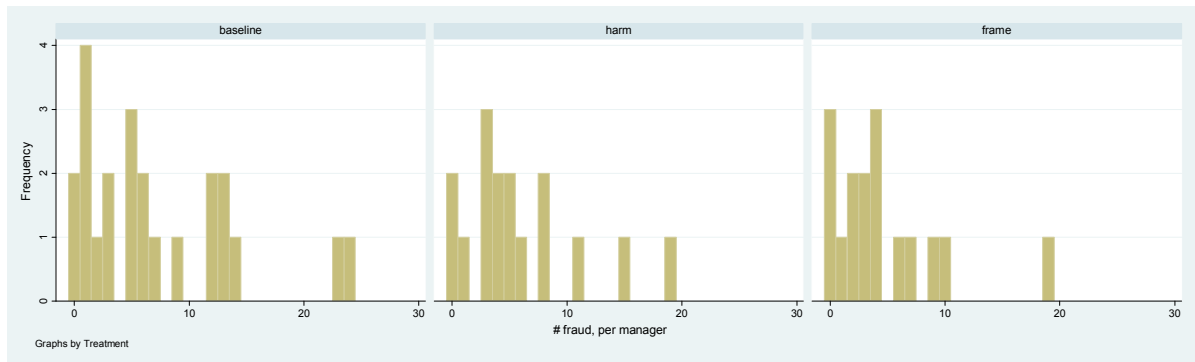
Figure 3
Manager Behaviour
Note: The game is played for 30 periods and in each period the managers select to commit fraud or not. The x-axis indicates how many times, out of a maximum of 30 opportunities, a certain manager committed fraud. The y-axis indicates how many managers chose to commit fraud that frequently...

Descriptively, fraud is more pronounced in the *baseline* (on average in 7.22 of 30 periods) than in the *harm* treatment (mean: 5.94 of 30 periods) than in the *frame* treatment (4.63 of 30 periods). Descriptively we thus have support for H_5 and H_6. Yet neither non-parametrically nor parametrically do we establish significant treatment differences.

b) Prosecutors

Unconditional prosecutor decisions would not be meaningful. Several observations immediately follow from Figure 4. Prosecutors strongly condition their behaviour on the signal they receive. If the signal is good, they usually do not take the manager to court, as predicted by the equilibrium of the game (H_1). Yet in opposition to game theory, in a substantial minority of cases, prosecutors charge overall conduct if the signal is good. If the signal is bad, in the *harm* treatment prosecutors are most likely to play the equilibrium and charge overall conduct. In the *baseline*, this is slightly less frequent. By contrast, in the *frame* treatment, this is very rare. In opposition to the equilibrium, prosecutors predominantly react to a bad signal by using the clear-cut rule.
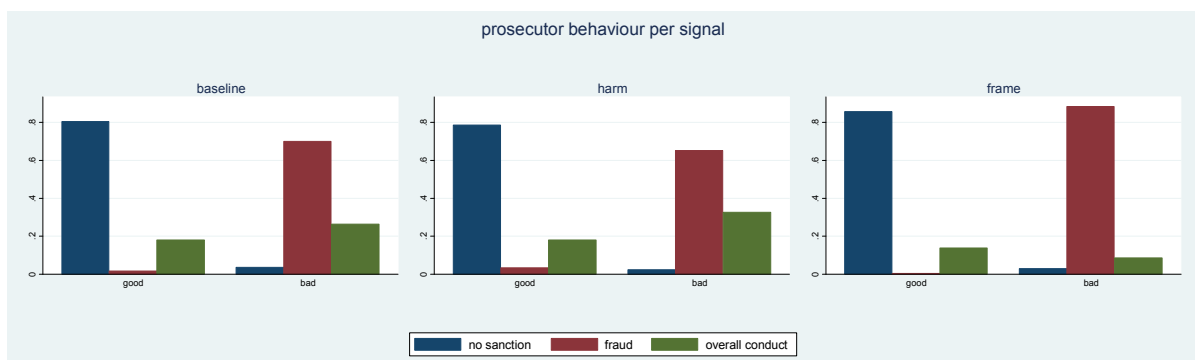

Figure 4
Prosecutor Action Conditional on Signal
Note: In each treatment, the prosecutor chooses between selecting No Action, the Clear-cut rule and Overall Conduct. This figure shows how these choices are split according treatment and signal.

Prosecutors have three options. If we were to analyse each option in isolation, and would compare them with choosing either of the remaining options, we would neglect that error terms of these separate regressions are likely to be correlated, given prosecutors choose from their portfolio of three options. We can also not treat the three options as ordered. A sanction is certainly more onerous than no sanction. But there are equally sound reasons to claim that the clear-cut rule is more severe (since the maximum effect is stronger, and since the sanction is applied with certainty if the manager was indeed dishonest), and that the overall conduct charge is even more severe (since it affects the manager randomly with 50 % probability, irrespective of the gravity of the offense). We must therefore treat the three options as categorical. We have no reason to explain any of these choices with different independent variables. Therefore multinomial logit is the appropriate functional form. Unfortunately there is no generally acknowledged mixed effects multinomial logit estimator. We therefore estimate ordinary multinomial logit models, but cluster standard errors at the highest possible level of dependence, namely matching groups. Note that this procedure is conservative. We do not exploit the fact that we know our data generating process to be more structured, in that individuals are nested in matching groups.[6]

Our main interest is in prosecutors using the overall conduct charge. We therefore define this option to be the baseline category. Our model predicts how likely prosecutors are to deviate to either no sanction or the clear-cut rule. The regressions in Table 1 show that adding the time trend and its interaction with treatments clears the picture. Model 2 provides the statistical test for Figure 4: If the signal is bad, in the *frame* treatment prosecutors are much more likely to charge the clear-cut rule. Interestingly, we also find a significant main effect of the *harm* treatment. In this treatment, clear-cut rule charges are somewhat more likely than in the *baseline*, irrespective of the signal. Prosecutors might feel the urge to use a sanction of the clear-cut rule as a signal to the manager that they care about the effects for the third player.

	model 1		model 2	
	no sanction	**clear-cut rule**	**no sanction**	**clear-cut rule**
Harm	-.023	.751[+]	-.070	.792*
Frame	.323	-.849	-.197	-.149
Signal	-3.486***	3.390***	-3.498***	3.370***
harm*signal	-.630	-1.039	-.577	-.943
frame*signal	.564	2.192	.553	3.352*
Period			-.029*	-.048**
harm*period			.002	-.005
frame*period			.032	-.087
Cons	1.500***	-2.409***	1.974***	-1.652***
N		1650		1650
p model		<.001		<.001
pseudo R^2		.3985		.4092

Table 1
Treatment Effects for Prosecutors Conditional on Signal
Note: A multinomial logit, with standard errors clustered at the level of matching groups, is run to model the behaviour of prosecutors. The constants determine how likely prosecutors are to deviate from overall conduct to either no charge or using the clear-cut rule. The coefficients of the explanatory variable show how a one unit change in these variables alters the likelihood of choosing the alternative to overall conduct. Model 2 introduces period effects. Signal is a dummy variable that is 1 if the signal is bad. ***, **, *, + indicates significance at the .001, .01, .05 and 0.1 levels, respectively.

We first have to explain why so many prosecutors violate the equilibrium if the signal is bad, and charge the clear-cut rule. To understand this finding, recall why, in equilibrium, prosecutors react to a bad signal with a charge of overall conduct. In equilibrium, managers

---

[6]   If we analyse individual options with mixed effects logit models, results look similar.

never behave badly. Therefore the bad signal must result from the fact that Nature has sent out a wrong signal. Yet if the prosecutor does not charge the manager when the signal is bad, she is punished with -7 tokens (Figure 1). Reacting with a charge of overall conduct to a bad signal is necessary to deter a money maximising, risk neutral manager. One may also say the payoff losses both players incur when the signal is bad is the price they are paying for the fact that the prosecutor can only imperfectly supervise the manager. Now the Nash equilibrium is a normative, not an empirical prediction. Were the prosecutor to know that a bad signal actually stands for the fact that the manager truly misbehaved, it would be her best response to charge the clear-cut rule. She then expects 10 tokens, instead of a gamble with a 50% chance to also get 10 tokens, but a 50% chance to loose 10 tokens. The data suggests that prosecutors take a bad signal as information about the likely true action of the manager, and play their best response. Given they know the signal to be correct with 90% probability, this interpretation seems reasonable, despite the fact that it violates game theory.

We are now in a position to address the main issue of this paper. The law expects prosecutors to stay inactive if the would-be defendant is very likely innocent (and no additional means of evidence to resolve the remaining uncertainty are available). In our experiment, if the signal is good, with 90% probability the manager is innocent, and the prosecutor has no technology for further reducing the remaining uncertainty. As both models of Table 1 show, in all treatments prosecutors are very likely to fulfil this normative expectation. In the *frame* treatment (and in the first round), the predicted probability of no charge if the signal is good is 83.82%. A charge of the clear-cut rule only has a 2.04% probability, so that the predicted probability of a charge for overall conduct is 14.14%. The prediction is even clearer if the signal is bad. In this case, the law would want the prosecutor to take action. To the degree the prosecutor is able to investigate the case, there is a 90% probability of guilt. The prosecutor knows she runs a small risk of loosing in court. But the law would want her to take this risk. In the *frame* treatment, almost all prosecutors do. The predicted probability of this charge if the signal is bad is as high as 98.92%, with only 0.78% left for a charge of overall conduct, and only 0.3% left for taking no action.

Figure 5 investigates this further. It collects average marginal effects[7] of the effect of main interest from Table 1 model 2: Right from the start, in the *frame* treatment prosecutors are about 20% more likely to use the clear-cut rule if they receive a bad signal. The treatment effect, conditional on the signal being bad, remains significant until period 21.

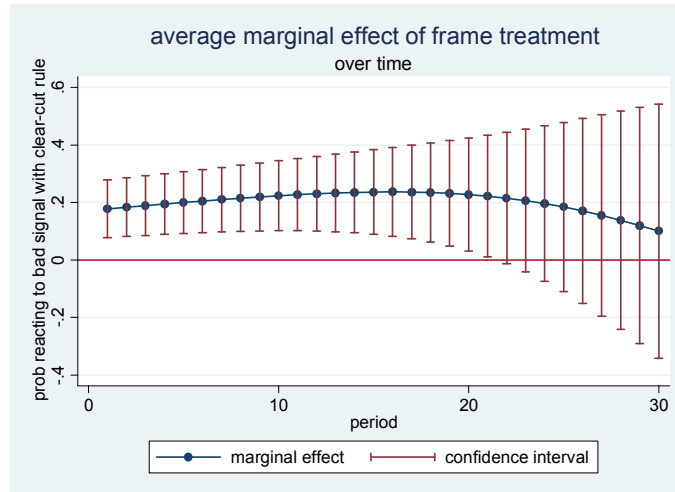[7]   Of course taking the multiplicative character of the interaction terms into account.

Figure 5
Marginal Effect of Frame Treatment on Using Clear-Cut Rule if Signal is Bad

By contrast, in the *harm* treatment, the predicted probability of a charge of the clear-cut rule when the signal is bad is only 80.32%, more than 18% less than in the *frame* treatment. In the *baseline*, the predicted probability is very similar, namely 81.43%. Our data suggests that we have two effects. We have already explained the first effect. In all treatments, most prosecutors believe a bad signal to be true, and play their best response. Yet in the *frame* treatment, and only in this treatment, there is an additional effect, which accounts for another 20% of choices. Through our design we know this additional effect to result from the call of duty and justice.

To cast further light on this effect, we test how prosecutors react to two types of experiences: false acquittals and false charges.

Recall that, after the end of each period, all participants get complete feedback. Through feedback, prosecutors learn the true action of managers, and they consequently know whether they have taken an innocent manager to court, or whether a guilty manager has escaped their scrutiny. We count the number of times prosecutors have let escape a guilty manager, and the number of times they have wrongly charged an innocent manager with the clear-cut rule. As Figure 6 shows, in all treatments prosecutors are more likely to make the latter than the former mistake. Overly harsh reactions are particularly likely in the *harm* treatment.
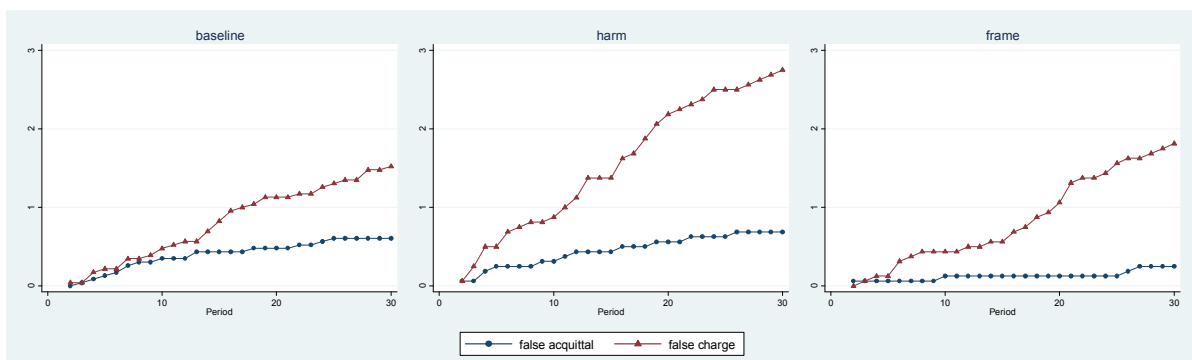

Figure 6
Prosecutor Failure
Note: Separately for each treatment, these graphs show how frequently, in the respective period, a prosecutor has falsely acquitted or falsely charged a manager in the entire past.

14

In Table 2 we use the number of times a prosecutor has failed in either direction as an additional explanation for their choices in the current period. All main effects, and all interactions with the *harm* treatment, are insignificant. The one significant three-way interaction with the *harm* treatment even points into the "wrong" direction. The more often prosecutors have wrongly acquitted a defendant, the more frequently they acquit him again if the signal is bad. By contrast, the two three-way interactions with the *frame* treatment significantly explain the decision to abstain from sanctioning, and all four interactions significantly explain the decision to charge the clear-cut rule. If participants are told they are holding the position of a prosecutor, they care about minimising mistakes. Actually they do so in both directions. They not only want to make sure that guilty managers are apprehended. They are equally zealous preventing innocent managers from being sentenced. Actually, if we calculate average marginal effects, it turns out that we only find a weakly significant, yet sizeable effect of having once falsely acquitted the defendant. In that case, the probability of using the clear-cut rule increases by 20.7 % (p = .064) in the *frame* treatment, compared with the remaining treatments. By contrast, if the prosecutor knows the *frame*, we find a strongly significant negative effect in reaction to one (-.019, p < .001) and to two false convictions (-.028, p = .028). These reaction patterns show quite clearly that prosecutors feel the urge of duty and justice.

	no sanction	clear-cut rule
Harm	.701	1.701^{+}
Frame	-.017	-.201
Signal	-3.517***	3.754***
harm*signal	-1.570	-2.234*
frame*signal	1.604^{+}	1.555
# past false acquittals	-.297	.323
# past false charges	-.194	.020
signal*# past false acquittals	-.598	-.686
signal*# past false charges	.228	.125
harm*false acquit	-.608	-.642
harm*false charge	-.089	-.246
harm*signal*false acquit	1.589*	1.171
harm*signal*false charge	-.491	.248
frame*false acquit	.131	2.834*
frame*false charge	.255	-14.388***
frame*signal*false acquit	-16.186***	-2.934^{+}
frame*signal*false charge	-18.227***	14.295***
Cons	1.789**	-2.776***
N		1595
p model		<.001
pseudo R^{2}		.4190

Table 2
The Effect of Past False Charges
Note: A multinomial logit, with standard errors clustered at the level of matching groups, is run to model the effect prosecutors' past false charges have on their selection of punishment. Interaction terms include treatments as well as signals..***, **, *, + indicates significance at the .001, .01, .05 and 0.1 levels, respectively.

We next turn to the (differential) effect of risk aversion. We first use the standard measure developed by (Holt and Laury 2002). Figure 7 reports switching points from the lottery with a smaller to the lottery with a larger spread, by the degree of risk involved in this choice. The later a participant switches, the more she is risk averse. 49 of 55 prosecutors were consistent on this test, meaning that they did at most switch once. Visibly our treatments did not induce different risk aversion patterns. Treatment differences are not significant, neither non-parametrically nor parametrically.
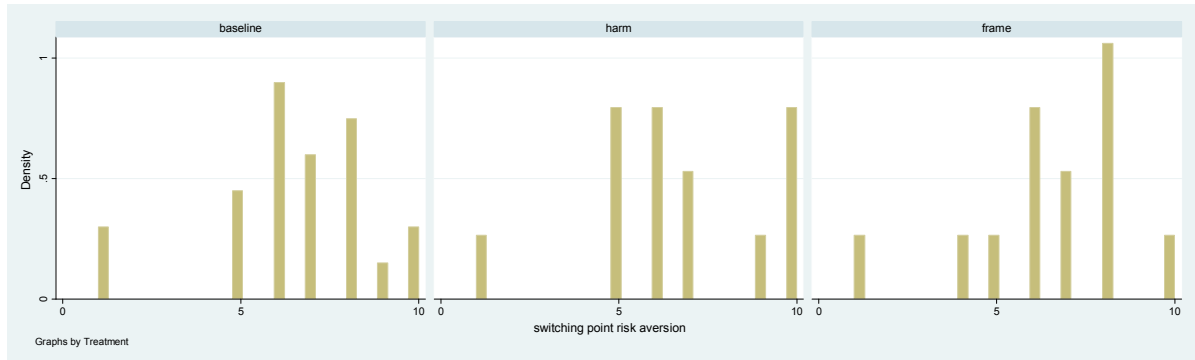
Figure 7
Risk Aversion
Note: Switching Points of prosecutors in the test developed by Holt/Laury, per treatment.

Table 3. however shows that risk aversion matters differently if prosecutors know the frame (two-way interaction with the switching point) and if, on top, the signal is bad (three-way interaction).

	no sanction	clear-cut rule
Harm	-3.106***	-.070
Frame	2.906$^+$	8.447***
Signal	-2.360	4.085***
harm*signal	2.668	-.682
frame*signal	-4.678*	-10.583***
switching point in risk aversion test	-.100	.067
harm*switch	.495***	.158
frame*switch	-.393	-5.091***
signal*switch	-.181	-.106
harm*signal* switch	-.506	-.096
frame*signal* switch	.773*	5.632***
Cons	2.169**	-1.955**
N		1470
p model		<.001
pseudo R^2		.4239

Table 3
Effect of Risk Aversion on Prosecutor Choice
Note: A multinomial logit with standard errors clustered at the level of matching groups is run with overall conduct as the reference point for prosecutorial decisions. The explanatory variables include treatments,
signal and interaction terms. The risk aversion switching points are also included. ***, **, *, + indicates significance at the .001, .01, .05 and 0.1 levels, respectively.

Since models with several interaction terms are difficult to read, we exploit the fact that we only have a single continuous explanatory variable. It therefore is straightforward to graph model predictions, which we do in Figure 8.[8] The graphs show that our treatments induce strikingly different reaction patterns in prosecutors. In the *baseline,* risk seeking prosecutors play the Nash equilibrium if the signal is good: they do not sanction the manager. By contrast, if the signal is bad, the predicted probability of a maximally risk seeking prosecutor to play Nash, i.e. to sanction the manager for overall conduct, is as low as 11.49%. The probability of sanctioning using the clear-cut rule is as high as 81.34%. The higher the degree of risk

---

[8]   We use predictions, rather than average marginal effects, since our result then is easier to understand. For all levels of risk aversion (except for maximal risk seeking) we find a significant negative average marginal effect of being in the *frame* treatment on the choice to use the clear-cut rule if the signal is good. We find a significant positive average marginal effect of being in the *frame* treatment on the choice to use the clear-cut rule if the signal is bad and risk aversion is substantial (switching points 7 and above).

aversion, the more the reaction approaches the equilibrium. But even a maximally risk averse prosecutor is still more likely to charge the clear-cut rule (58.80%) than overall conduct (39.23%).

In the lab, prosecutors do not know whether the manager they happen to be paired with in the current period is actually playing Nash. Figure 8 suggests that, in the *baseline*, prosecutors construct their reaction to a bad signal as a gamble. The more they are risk averse, the more they play it safe and charge overall conduct, despite the fact that a charge of overall conduct is itself a lottery. In expected values, this gives them 0 tokens. In the worst case they loose 10 tokens. This is what they loose with certainty if they charge the clear-cut rule and the manager was well behaved. We thus refute **H₂** and support **H₄**: when they see a bad signal, prosecutors predominantly do not play the equilibrium, but their best response.
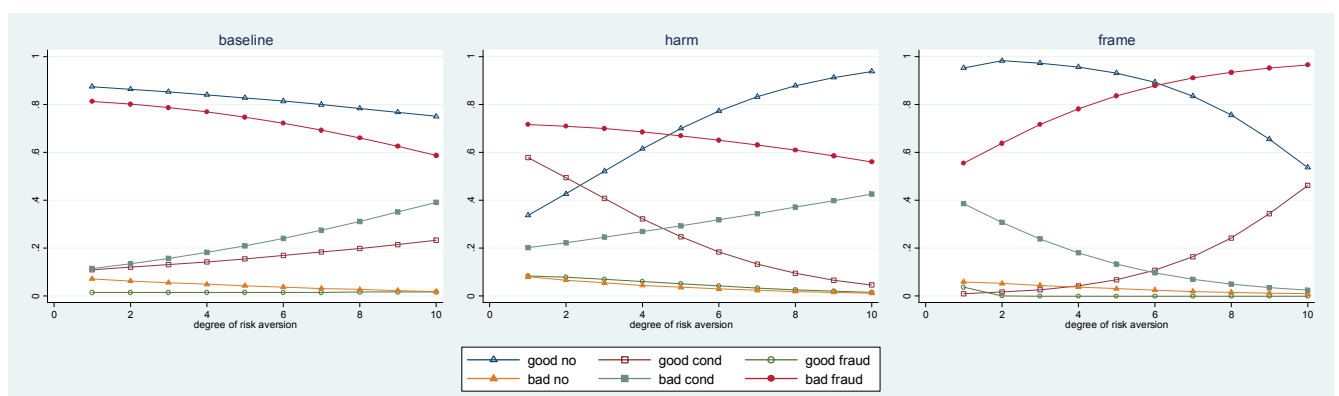

Figure 8
Predicted Prosecutor Choices per Treatment, Signal, and Degree of Risk Aversion
Note: Risk aversion using the Holt/Laury test is graphed on a ten point scale. For each treatment lines indicate the predicted probability of each of the three prosecutor choices (no sanction, overall conduct, the clear-cut rule), conditional on the signal (good or bad).

In the *harm* treatment, the reaction to a bad signal is basically the same as in the *baseline*, only that even risk seeking prosecutors are somewhat more likely to charge overall conduct, and less likely to charge the clear-cut rule. In the regression (Table 3 model 3), neither the main effect of the *harm* treatment, nor the two-way interaction with the signal, or the three-way interaction with the signal and the switching point in the risk aversion test are significant. The reaction pattern looks almost identical to the *baseline* (Figure 8). We thus find no support for **H₅**.

Yet in the *harm* treatment, the reaction to a good signal is strikingly different from the *baseline*. Now the more prosecutors are risk averse, the more they are likely to not sanction if the signal is good. The more they are risk seeking, the more they switch to a sanction for overall conduct.[9] Note that charging overall conduct is not even a best response if the signal is true. In expectation with this choice the prosecutor looses 5 tokens, while she has 0 tokens if she refrains from taking action. However, there is a 50% chance to get 10 tokens. Had we seen the same pattern in the *baseline*, we could simply have thought that risk seeking prosecutors are attracted by the large gain. Yet in the *baseline*, we find the opposite effect of risk preferences. Since payoffs are the same in both treatments, payoffs alone cannot be the explanation. The explanation must rest in the fact that managers' action also affects

---

[9]   Average marginal effects of being in the *harm* treatment, on the choice to charge overall conduct, are significantly positive for switching points in the test for risk aversion from 1-4, and significantly negative for switching points 8-10, while they are insignificant for intermediate levels of risk aversion.

shareholder payoffs. Now for the period in question, prosecutors' choices are irrelevant for shareholder payoffs. Yet this is different in the long run. Arguably, if this manager has not been properly deterred by today's prosecutor, he will be more likely to impose harm on the shareholder tomorrow. Risk seeking prosecutors can use this argument as an excuse for behaviour that is actually driven by their desire to land the large gain. This explanation is in line with a finding from dictator games. If dictators are given a chance to hide selfish behaviour, for instance by making the recipient believe that their choice had no impact on the recipient's earnings, many dictators seize the opportunity of having "moral wiggle room" (Dana, Cain et al. 2006; Andreoni and Bernheim 2009; Lazear, Malmendier et al. 2009).

This explanation fits the pattern in the *frame* treatment. The reaction to a good signal is similar to the *baseline*, not to the *harm* treatment, only that the effect of risk aversion is even more pronounced. Risk seeking prosecutors overwhelmingly react to a good signal by refraining from sanctioning, while maximally risk averse prosecutors are almost as likely to charge overall conduct as taking no action.[10] The reversal of the direction of the effect of risk aversion suggests that risk seeking prosecutors can no longer use victims as an excuse for selfishness and feel bound by their duty to exercise prosecutorial powers properly.

The reaction to a bad signal is diametrically opposed to the pattern in both other treatments. As already Figure 4 showed, and as is supported by the significant frame*signal interaction in model 2 of Table 1, in the *frame* treatment, prosecutors are very likely to use the clear-cut rule when receiving a bad signal.[11] We thus support **H₆**. Actually the regression even provides us with a more fine-grained picture. Risk averse prosecutors almost perfectly react to a bad signal with charging the clear-cut rule, while risk seeking prosecutors are more attracted by the option to charge for overall conduct. Now we have already shown that prosecutors tend to take the signal at face value. Then in expectation charging the clear-cut rule yields the highest payoff. It is the best response. Yet in the *frame* treatment, we have not only framed the prosecutor's role. The manager's choice is also framed as honest behaviour versus fraud. This might induce prosecutors to believe that managers are less likely to behave badly, and therefore a bad signal is less likely to be true.

This explanation is supported by another of our post-experimental tests. In each period, before prosecutors receive the signal, we ask them how many managers they believe have acted dishonestly. While prosecutors believe 38.23 % of all managers to be dishonest in the *baseline*, they only believe 27.38 % of them to misbehave in the *frame* treatment.[12] If this is their belief, bad signals are more likely to be wrong. In that case, in expectation overall conduct gives prosecutors a higher payoff. Yet this higher payoff comes at a 50% risk of only getting a small payoff. Risk averse prosecutors might dread this risk. It of course is also possible that the sense of prosecutorial duty is correlated with risk aversion.

Our risk aversion measure is standard in the experimental economics literature. Since it is fully incentivized, it is particularly reliable. Yet it is a one size fits all approach. Psychologists have wondered whether this is overly general, and have developed an alternative (non-incentived) scale. Through a series of survey items, the DOSPERT elicits risk attitudes in five domains (ethical, financial, health, recreational, social) and in three dimensions (would you take the risk? how risky is this activity? how much benefit would you derive from engaging in

---

[10] Average marginal effects of being in the *frame* treatment, on the decision not to take action if the signal is good are significant and positive for risk seeking prosecutors (switching points in the Holt/Laury test 2-6), and insignificant otherwise.

[11] For average marginal effects see footnote 8, above.

[12] Mixed effects model, random effects for prosecutors nested in matching groups, coef p < .001, N = 1650.

it?). For the subsequent analysis, we focus on the two domains that are most related to our issue: the financial domain (capturing the incentive side of the conflict our mock prosecutors are facing) and the ethical domain (representing the call of duty), and on the dimension that is most relevant, namely the willingness to actually take one of these risks.

Figure 9 shows that there is considerable variance of risk attitudes across domains and across treatments. In the financial domain, the majority of prosecutors are more or less risk neutral (as indicated by scores converging to the midpoint of the scale). In the ethical domain, we have many more extreme scores, indicating either risk seeking or pronounced risk aversion.
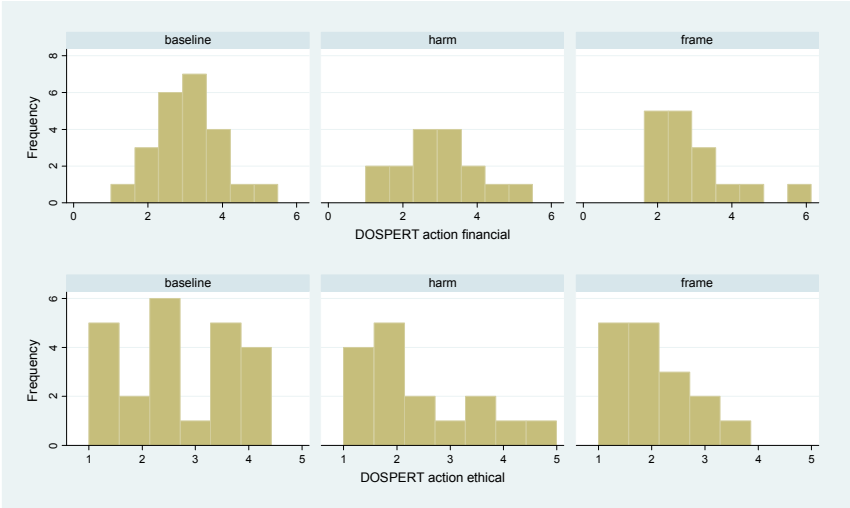
Figure 9
Domain Specific Risk Attitudes of Prosecutors
Note: For each treatment, results regarding risk scales for the DOSPERT survey are displayed. The focus is on the ethical and financial dimensions. For each dimension, the action, benefit and risk choices are determined..

As such, the DOSPERT scores have little explanatory power, Table 4. This should not come as a surprise. One should expect prosecutors to react to the signal and the treatment differences. Since the time trend helped clear the picture (Table 1), we also include it here. We do not find significant effects of these scores on the decision to abstain from action, also not if we calculate average marginal effects. By contrast we do find differential effects on the decision to use the clear-cut rule.

sanction	model 1	
	no sanction	**clear-cut rule**
Harm	.032	1.676*
Frame	-.851	-1.163
Signal	-1.854*	3.776**
harm*signal	-.609	-1.209
frame*signal	.029	1.511
Period	.020	-.048
harm*period	-.002	-.031
frame*period	.023	-.475***
signal*period	-.138*	-.016
harm*signal*period	.016	.019
frame*signal*period	.080	.408***
DOA ethical	.195	-.271
DOA financial	-.573*	-.203
harm*DOA ethical	.051	1.081[+]
harm*DOA financial	-.031	-1.068**
frame*DOA ethical	-.816	.736
frame*DOA financial	.849	.278
Cons	3.099**	-.397
N		1650
p model		<.001
pseudo R^2		.4335

Table 4
Explaining Prosecutor Choices with Domain Specific Risk Aversion Scores
Note: multinomial logit, standard errors clustered at the level of matching groups; DOA: DOSPERT scale for willingness to engage in risky activities; Signal = 1 if signal is bad. *** $p <. 001$, ** $p < .01$, * $p < .05$, [+] $p < .1$

To better understand these effects, in Figure 10 we collect average marginal effects. It turns out that the DOSPERT measures of risk aversion are immaterial for the decision to abstain from action (not reported) and for the decision to choose the clear-cut rule if the signal is good, while they matter greatly if the signal is bad. Interestingly, it takes a more than minimal willingness to take either ethical or financial risk for prosecutors to follow the call of duty and to react to a bad signal with a charge under the clear-cut rule. It is the prosecution frame, and only this frame, that makes prosecutors willing to take these risks.
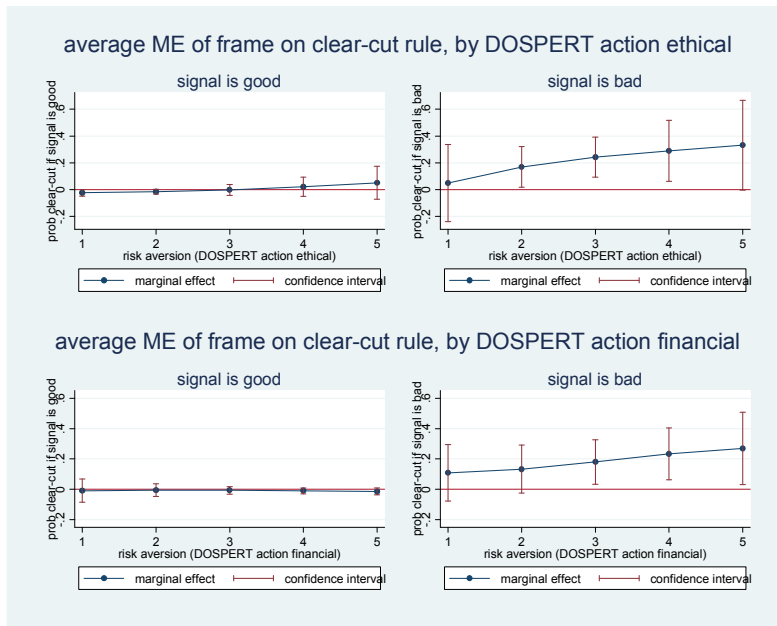
Figure 10
Average Marginal Effects of being in *Frame* Treatment
on Willingness to Take Domain Specific Risks, by DOSPERT action scores

6. Discussion

When payoff maximising agents interact with other agents who do the same, and when all actors believe all others to hold such standard preferences, the situation lends itself to game-theoretic modelling. The rational choice theory of crime believes criminals to maximise (not necessarily pecuniary) utility. This expectation seems particularly plausible with white collar crime. Observers frequently wonder whether the same holds for prosecutors. In their case, utility would most likely not be exclusively pecuniary. But it does not seem far fetched to assume an incentive of prosecutors to win in court. Utility maximising agents exploit opportunities as they present themselves. Prosecutors have the power to select the charge. If the odds of winning are unclear, in expectation prosecutors may be better off when shifting the charge to crimes that are only vaguely defined in doctrine, like conspiracy or false statements. Utility maximising prosecutors should seize this opportunity.

To test this supposition, we have had experimental defendants and prosecutors play a stylized game. The game has a unique equilibrium in pure strategies. In equilibrium, prosecutors do not take the defendant to court if they receive a signal to the effect that the defendant is likely innocent. If they receive the opposite signal, they do not charge her for a well-defined crime. Instead they shift to a charge of "overall conduct". In that case, the odds of winning in court are unrelated to the actual behaviour of the defendant. In anticipation, would-be defendants never commit the crime.

In the *baseline* we have participants play the naked game. Results fully refute the game-theoretic prediction. Defendants quite frequently misbehave. Seemingly defendants are tempted by the possibility of a large gain. If they receive a good signal, prosecutors predominantly do not take action. Yet a substantial fraction charges overall conduct, which is not predicted by game theory. If the signal is bad, even a large majority of prosecutors

violates game theory and charges the specific crime. This suggests that prosecutors believe the bad signal to be true, and play their best response, instead of the equilibrium.

Adding a third participant who suffers if the defendant misbehaves has practically no effect on defendants' behaviour. They do not shy away from imposing harm on a passive outsider. In this treatment, descriptively prosecutors are most likely to react to a bad signal with a charge of overall conduct. Yet this is not statistically significant from the *baseline*.

By contrast, behaviour looks substantially and significantly different if we reveal our research question and speak of a manager, a shareholder and a prosecutor. This polarizes prosecutor behaviour. If they see a good signal, they are strongly inclined to refrain from a charge. If they see a bad signal, they almost unanimously use the clear-cut rule. Through a series of post-experimental tests we corroborate that the treatment effect predominantly results from the fact that prosecutors feel the urge of duty. In our experiment, the suspicion that prosecutors exploit the vagueness of doctrine to their selfish advantage does not hold true.

As in any experiment in law, one has reason to carefully consider external validity. In the courtroom, much more is at stake for the defendant. If at all, this qualification should have worked against us. Since they know that defendants at most lose a couple of dollars, prosecutors might have been less hesitant to ignore the frame and to treat the situation just as a game. This is clearly not what has happened.

In the courtroom, for prosecutors more is also at stake. They lose an opportunity to advance their careers, or they dread pressure from their superiors and the public if they lose a case. By contrast, in our experiment they at most put a couple of dollars at risk. We cannot exclude that higher stakes change behaviour. But it is remarkable that we find a highly significant treatment effect, in particular in reaction to a bad signal. Since the fraction of prosecutors who respond to a bad signal with a charge of the clear-cut rule is already high in the *baseline*, for the treatment effect to be significant we need a very strong and very clear effect.

We have had our participants to act repeatedly. The time trend is significant and negative, and it pushes behaviour closer to the game theoretic equilibrium. In their professional lives, prosecutors not only meet thirty defendants, but hundreds of them. While they may be willing to live up to normative expectations in the beginning of their careers, over time the litigant spirit might gain the upper hand. Again we cannot negate this qualification altogether. But we note that, within the thirty periods of our experiment, the time trend is very flat and far from reversing the treatment effect even at the very end of the experiment.

In our design, the signal is correct with 90% probability. In the field it happens that prosecutors know with near certainty whether the defendant is guilty or innocent. Yet frequently at the end of police investigations there remains a higher degree of uncertainty. The law does not prevent prosecution from accusing such defendants as long as the charge is not frivolous. But in such cases prosecutors might be more tempted to play it safe and exploit the ambiguity of legal doctrine. We acknowledge this but note that we have constructed the game such that the unique equilibrium had prosecutors bring the vague charge. In the field, the ambiguity of the law is not only a panacea for prosecutors, it also is a risk. Precisely because the law is not clear, prosecutors also have a hard time predicting the outcome. We have entirely removed this source of uncertainty, which should have made this choice even more attractive.

In the experiment, if the prosecutor charges overall conduct the odds of winning are totally unrelated to the guilt of defendants. In the courtroom, the difference between clear-cut and vague rules is less extreme. While vague rules are considerably more ambiguous, the defendant still stands a better chance to be acquitted if he is actually innocent. Also, doctrine is never fully settled, nor is its application to the case at hand, so that there is inevitably a dose of ambiguity even in the application of apparently clear-cut rules. Prosecutors who are hesitant to expose defendants to a true gamble might be more willing to exploit the vagueness of the law if they can assuage their conscience with the excuse that, to a degree, all law is ambiguous. We cannot exclude that this might matter, but we note that this objection presupposes our main result to be true. The possibility to assuage one's conscience only matters if conscience guides behaviour in the first place. This is what our experiment was meant to show.

Whenever there is an opportunity, it is good policy to be vigilant. The stronger the incentives, the more likely it is that agents seize the opportunity, even if the law expects them to ignore it. In the introduction of this paper we list the reasons why prosecutors might be tempted to exploit the vagueness of criminal law to their selfish benefit. Against this backdrop it is remarkable that mock prosecutors next to never seize the opportunity in our lab experiment once we make them aware of their virtual public office. Policymakers, and the watchdogs of the public, should not stop being vigilant. But our experiment justifies giving prosecutors the benefit of the doubt. As long as there are no signs to the contrary, they should not be suspected to be reckless hired guns. The call of prosecutorial duty is stronger than one might have thought.

ANDREONI, JAMES and B. DOUGLAS BERNHEIM (2009). "Social Image and the 50-50 Norm. A Theoretical and Experimental Analysis of Audience Effects." Econometrica **77**: 1607-1636.

BALDWIN, JOHN (1993). "Police Interview Techniques." British Journal of Criminology **33**(3): 325-352.

BLAIS, ANN-RENEE and ELKE U. WEBER (2006). "A Domain-Specific Risk-Taking (DOSPERT) Scale for Adult Populations." Judgement and Decision Making **1**: 33-47.

BOYLAN, RICHARD T. and CHERYL X. LONG (2005). "Salaries, Plea Rates, and the Career Objectives of Federal Prosecutors." Journal of Law and Economics **48**: 627-709.

BRINK, MALIA N. (2009). "A Pendulum Swung Too Far. Why the Supreme Court Must Place Limits on Prosecutorial Immunity." Charleston Law Review **4**: 2-222.

CHARNESS, GARY (2000). "Self-Serving Cheap Talk. A Test of Aumann's Conjecture." Games and Economic Behavior **33**: 177-194.

CRANK, JOHN, DAN FLAHERTY and ANDREW GIACOMAZZI (2007). "The Noble Cause. An Empirical Assessment." Journal of Criminal Justice **35**(1): 103-116.

DANA, JASON, DAYLIAN M. CAIN and ROBYN M. DAWES (2006). "What You Don't Know Won't Hurt Me. Costly (But Quiet) Exit in Dictator Games." Organizational Behavior and Human Decision Processes **100**: 193-201.

ENGEL, CHRISTOPH and BERND IRLENBUSCH (2010). Turning the Lab into Jeremy Bentham's Panopticon. The Effect of Punishment on Offenders and Non-Offenders http://ssrn.com/abstract=1555589.

FARRINGTON, DAVID P. (2003). "A Short History of Randomized Experiments in Criminology. A Meager Feast." Evaluation Review **27**: 218-227.

FARRINGTON, DAVID P. (2006). "Key Longitudinal-Experimental Studies in Criminology." Journal of Experimental Criminology **2**: 121-141.

FARRINGTON, DAVID P. and BRANDON C. WELSH (2005). "Randomized Experiments in Criminology. What Have We Learned in the Last Two Decades?" Journal of Experimental Criminology **1**: 9-38.

FARRINGTON, DAVID P. and BRANDON C. WELSH (2006). "A Half Century of Randomized Experiments on Crime and Justice." Crime and Justice **34**: 55-132.

FISCHBACHER, URS (2007). "z-Tree. Zurich Toolbox for Ready-made Economic Experiments." Experimental Economics **10**: 171-178.

GARRETT, BRANDON J. (2008). "Judging Innocence." Columbia Law Review **108**: 55-142.

GORDON, SANFORD C. and GREGORY A. HUBER (2002). "Citizen Oversight and the Electoral Incentives of Criminal Prosecutors." American Journal of Political Science **46**(2): 334-351.

GREEN, BRUCE A. and FRED C. ZACHARIAS (2004). "Prosecutorial Neutrality." Wisconsin Law Review: 837-904.

GROSS, SAMUEL R., KIRSTEN JACOBY, DANIEL J. MATHESON, NICHOLAS MONTGOMERY and SUJATA PATIL (2004). "Exonerations in the United States 1989 through 2003." Journal of Criminal Law and Criminology **95**: 523-560.

HEMINWAY, JOAN MACLEOD (2002). "Save Martha Stewart. Observations about Equal Justice in US Insider Trading Regulation." Texas Journal of Women and Law **12**: 247-286.

HESSICK, F.ANDREW and RESHMA M. SAUJANI (2001). "Plea Bargaining and Convicting the Innocent: the Role of the Prosecutor, the Defence Counsel, and the Judge." Brigham Young University Journal of Public Law **16**: 189-242.

HOEFFEL, JANET (2004). "Prosecutorial Discretion at the Core. The Good Prosecutor Meets Brady." Pennsylvania State Law Review **109**: 1133-1154.

HOLT, CHARLES A. and SUSAN K. LAURY (2002). "Risk Aversion and Incentive Effects." American Economic Review **92**: 1644-1655.

KASSIN, SAUL M., CHRISTIAN A. MEISSNER and REBECCA J. NORWICK (2005). "“I'd Know a False Confession if I Saw One”: A Comparative Study of College Students and Police Investigators." Law and Human Behavior **29**(2): 211-227.

LAZEAR, EDWARD P., ULRIKE MALMENDIER and ROBERTO A. WEBER (2009). Sorting and Social Preferences http://www.econ.berkeley.edu/~ulrike/Papers/Sorting_Paper_2009-06-19.pdf.

LEE, CYNTHIA K.Y. (1997). "From Gatekeeper to Concierge. Reigning in the Federal Prosecutor's Expanding Power over Substantial Assistance Departures." Rutgers Law Review **50**: 199-251.

MONTERO, MARIA, MARTIN SEFTON and PING ZHANG (2008). "Enlargement and the Balance of Power. An Experimental Study." Social Choice & Welfare **30**: 69-87.

MOOHR, GERALDINE SZOTT (2006). "What the Martha Stewart Case Tells Us About White Collar Criminal Law." Houston Law Review **43**: 591-619.

PODGOR, ELLEN S. (1999). "The Ethics and Professionalism of Prosecutors in Discretionary Decisions." Fordham Law Review **68**: 1511-1535.

PRIEST, GEORGE L. and BENJAMIN KLEIN (1984). "The Selection of Disputes for Litigation." The Journal of Legal Studies **13**(1): 1-55.

RASMUSEN, ERIC, MANU RAGHAV and MARK RAMSEYER (2009). "Convictions versus Conviction Rates: The Prosecutor's Choice." American Law and Economics Review **11**(1): 47-78.

SEIGEL, MICHAEL L. and CHRISTOPHER SLOBOGIN (2004). "Prosecuting Martha. Federal Prosecutorial Power and the Need for a Law of Counts." Pennsylvania State Law Review **109**: 1107-1131.

Szott, Geraldine (2006). "What the Martha Stewart Case Tells us about White Collar Criminal Law." <u>Houston Law Review</u>(43): 30.

Yaroshefsky, Ellen (1999). "Cooperation with Federal Prosecutors: Experiences of Truth Telling and Embellishment." <u>Fordham Law Review</u> **68**: 917-964.

Zacharias, Fred C. (2000). "The Professional Discipline of Prosecutors." <u>North Carolina Law Review</u> **79**: 721.

Instructions

You are participating in an experiment on decision making and will be asked to make a number of decisions. During the experiment you will be able to earn money. Your exact earnings will depend on your and other players' decisions. These instructions describe the decisions you and other participants will make and how they determine your earnings. It is therefore important that you read them carefully.

During the experiment, all the interaction between you and other participants will take place through the computers. Please do not talk with other participants. If you have any questions, please raise your hand; one of us will come to answer your question. Note that the experiment is anonymous: that is, your identity will not be revealed to others and the identity of others will not be revealed to you.

During the experiment your earnings will be calculated in points. At the end of the experiment points will be converted to dollars at the following rate:

<div align="center">

100 points = 1 dollar

</div>

After the experiment your total earnings will be paid out to you in cash.

a) Instructions

In the study, participants are randomly assigned a role. There are three roles: **Manager, Prosecutor, and Shareholder.** Each participant will be assigned to one role and the role will not change for the duration of the experiment.

The experiment is divided into **30 rounds**. In each round, each player receives an endowment of **40** points.

In each round, you will be randomly assigned to groups of 3. Each group consists of a Manager, a Prosecutor, and a Shareholder. At the beginning of the game, one of the three roles will be randomly assigned to you. Roles will stay fixed for the entire experiment. You will be matched with **different people in each of the 30 rounds**, though.

b) Your decision

The Manager will first make a decision. She or he chooses between **Honesty** and **Fraud**. Next, the Prosecutor chooses whether or not to sanction the Manager. She or he chooses between **No Sanction**, sanction through a **Clear-cut Rule (Fraud)** and sanction through an **Overall Conduct Rule**. Shareholders do not make a decision. The earnings of all participants will be affected by all of these decisions. These effects are described below.

    i. Decisions and earnings of the Manager

If you are a Manager, you will decide whether to choose Honesty or Fraud.

If you choose Honesty:

- your earnings and the earnings of the Shareholder will not be impacted by your decision. The Shareholder keeps her endowment of 40 points with certainty
- Your own earnings depend on the decision of the Prosecutor: Particularly,
  - If the Prosecutor chooses No Sanction, you keep your endowment of 40 points
  - If the Prosecutor chooses to charge Fraud, you will be acquitted, but your earnings will decrease by 1 point because of trial costs

o If the Prosecutor charges Overall Conduct, your earnings will depend on chance. Chance selects to overturn or uphold the overall conduct conviction. 5 times out of 10, because the conviction is overturned and you are not sanctioned but you do incur trial costs, your earnings will be decreased by 1 point. 5 times out of 10, because the decision is upheld, you receive a sanction of 30 in addition to bearing the trial costs and hence your earnings will decrease by 31 points.

If you choose Fraud:
- you will increase your earnings by **10** points, but you will decrease the earnings of the Shareholder by **12** points. Hence the Shareholder receives 28 points with certainty.
- Your earnings will also be affected by the decision of the Prosecutor: Particularly,
  o If the Prosecutor chooses No Sanction, no change will be made to your earnings
  o If the Prosecutor charges Fraud, you will be convicted and receive a sanction of **30** points
  o If the Prosecutor charges Overall Conduct, your earnings will depend on chance. Chance selects to overturn or uphold the overall conduct conviction. 5 times out of 10, because the conviction is overturned and you are not sanctioned but you do incur trial costs, your earnings will be decreased by 1 point. 5 times out of 10, because the decision is upheld, you receive a sanction of 30 in addition to bearing the trial costs and hence your earnings will decrease by 31 points.

The following diagram illustrates the final earnings for the Manager:
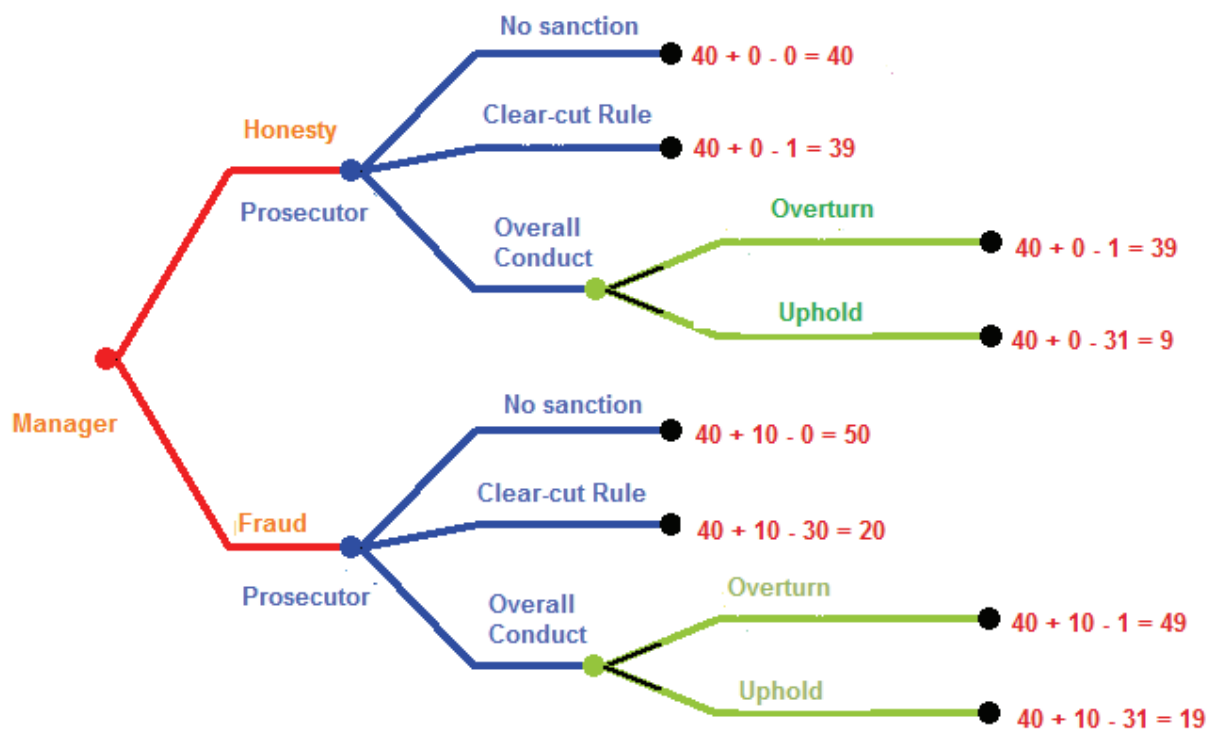
Please follow the diagram as you go through the below example.

<u>Example 1</u>
- You begin with an endowment of **40** points

Suppose that:.
- You select Honesty and hence receive no additional points. (+0 points)
- The Prosecutor selects No Sanction, where no points are subtracted from you (-0 points)
- In this case, your earnings equal 40 +0 + 0 = 40.

<u>Example 2</u>:
- You begin with an endowment of **40** points.

Suppose that
- You select Fraud and hence receive **10** points (+10 points)
- The Prosecutor selects Clear-cut Rule, where you receive a sanction of **30** points (-30 points)
- Therefore, your earnings equal 40 + 10 – 30 = 20

<u>Example 3</u>:
- You begin with an endowment of **40** points.

Suppose that:
- You select Fraud and hence receive **10** points (+10 points)
- The Prosecutor selects Overall Conduct, where chance dictates whether the decision is upheld or overturned
  - Chance selects that the decision is upheld and you therefore have a 30 point sanction and a 1 point trial cost subtracted from you (-31 points)
- Therefore your earnings equal 40 + 10 – 31 = 19

ii.   Decisions and earnings of Prosecutor

If you are a Prosecutor, you will decide whether to choose No Sanction, the Clear-cut Rule (Fraud) or the Overall Conduct Rule.

When making your decision you will **not know with certainty** the choice of the Manager. Instead, you will be informed about the decision through a **signal**. Whether the signal correctly transmits the actual choice is determined by chance. Specifically, 9 out of 10 times the signal will match the actual decision of the Manager and 1 out of 10 times it will not match the decision. In other words, when the Manager chooses Honesty, the signal will indicate Honesty 9 out of 10 times and Fraud 1 out of 10 times. Conversely, if the Manager chooses Fraud the signal will indicate Fraud 9 out of 10 times and Honesty 1 out of 10 times.

Your decisions result in the following earnings:
iii. If you observe the signal of Honesty, you have three choices:
- If you choose No Sanction, your earnings do not change.
- If you charge Fraud, two things can happen:
  - if the action of the Manager was actually Honesty (i.e. the signal was correct; 9 times out of 10), you receive a punishment of **20** points for accusing someone innocent
  - if the action of the Manager was actually Fraud (i.e. the signal was incorrect; 1 time out of 10), you receive a reward of **10** points for punishing correctly.
- If you choose the Overall Conduct Rule, chance selects to overturn or uphold the overall conduct conviction. 5 out of 10 times Overall Conduct will be upheld and you will receive a reward for punishing correctly of **10** points and 5 out of 10 times Overall Conduct will be overturned and you will be punished for abuse of power by **20** points.

The below diagram illustrates the decisions of the Prosecutor when she or he receives a signal of Honesty:
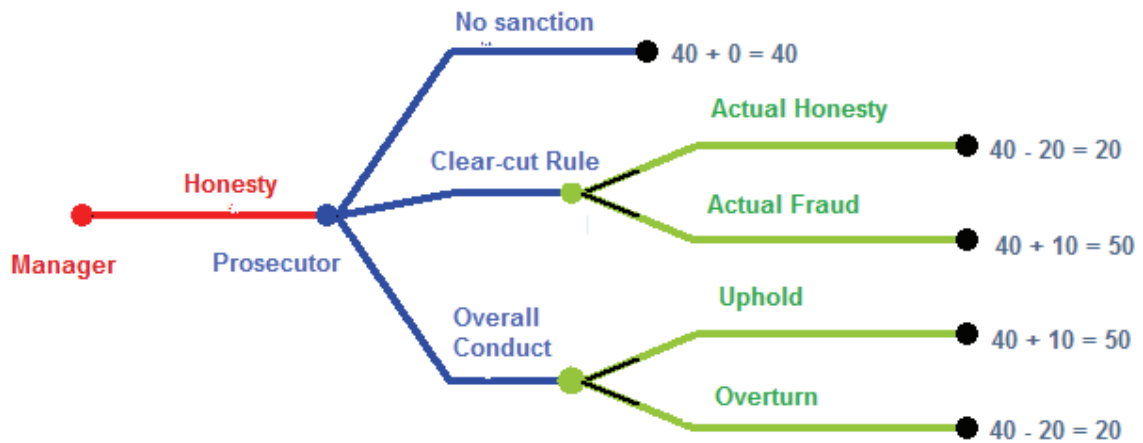
Please follow the diagram as you go through the below example.

Example 1
- You begin with an endowment of **40** points.

Suppose that:
- You select No Sanction, hence there is no change in your earnings (+0 points)
- Therefore, your earnings equal 40 + 0 = 40

Example 2:
- You begin with an endowment of **40** points.

Suppose that:
- You select the Clear-cut Rule, hence your earnings depend on the actual choice of the Manager
- If the signal indicated correctly that the Manager had chosen Honesty, you are punished for accusing someone innocent by **20** points (-20 points)
- Therefore, your earnings equal 40 – 20 = 20
- You receive earnings of **20** points

Example 3:
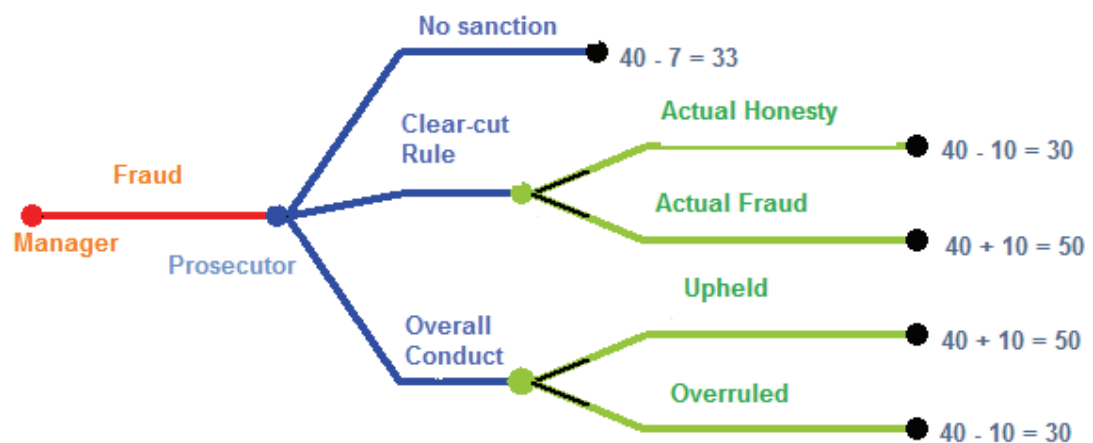- You begin with an endowment of **40** points.

Suppose that:
- You select Overall Conduct, where the number of points subtracted is dictated by chance
    - o Chance selects that your charge is upheld and you have **10** points added (+10)
- Therefore, your earnings equal 40 +10 = 50

iv. If you observe the signal of Fraud, you have three choices:
- If you choose No Sanction, you are punished **7** points for not accusing a potential offender.
- If you choose the Clear-cut Rule, two things can happen:

- If the Manager committed Fraud (i.e. the signal was correct; 9 times out of 10), you receive a reward of **10** points for convicting correctly (+10 points)
- If the Manager was Honest (i.e. the signal was incorrect; 1 time out of 10), you receive a punishment of **10** points for convicting an innocent person. (-10 points)
- If you choose Overall Conduct, your earnings will depend on chance. 5 out of 10 times Overall Conduct will be upheld and you will receive a reward for punishing correctly of **10** points and 5 out of 10 times Overall Conduct will be overturned and you will be punished for abuse of power by **10** points.

The below diagram illustrates the decisions of the Prosecutor when she or he receives a signal of Fraud:

Please follow the diagram as you go through the below example.

Example 1
- You begin with an endowment of **40** points.

Suppose that:
- You select No Sanction, hence you have 7 points subtracted from your earnings for not accusing a potential offender (-7 points)
- Therefore your earnings equal 40 – 7 = 33

Example 2:
- You begin with an endowment of **40** points.

Suppose that:
- You select the Clear-cut Rule, hence your earnings depend on the actual choice of the Manager
- If the signal indicated correctly that the Manager chose Fraud, you receive **10** points for convicting correctly (+10 points)
- Therefore, your earnings equal 40 + 10 = 50

Example 3:
- You begin with an endowment of **40** points.

Suppose that:
- You select Overall Conduct, where the number of points subtracted is dictated by chance

- o Chance selects that your charge is upheld and you have **10** points added (+10 points)
- Therefore, your earnings equal 40 + 10 = 50

c) Earnings of Shareholders

In the experiment, shareholders do not make a decision. If you have this role, your earnings will depend on the action of the Manager. Each round you have an endowment of **40** points. If the Manager chooses Honesty, you keep these 40 points. If the Manager chooses Fraud, your endowment is reduced by **12** points.

At the end of each round, you will be informed of the decision of the Manager, the signal, the decision of the Prosecutor, and each participant's payoff.

d) Screens and order in which decisions are made

The first step of the experiment consists of the Managers selecting an action:

In this screen you decide whether to choose Honesty or Fraud. To make your decision, select from the drop box and click on the "Submit" button.

The second step of the experiment consists of Prosecutors selecting an action.

In this screen, you are informed of the signal that indicates correctly 9 out of 10 times whether the Manager has chosen Honesty or Fraud. Based on this, you decide whether to select Overall conduct, No sanction or the Clear-cut Rule (Fraud). To make your decision, select from the drop box and click on the "Submit" button.

Once both selections have been made, earnings are calculated and all three participants will be informed of the outcome on the following screen and receive their share. However earnings from estimating the number of Managers who have chosen Honesty are calculated in the background. You will be informed about the actual number per period, and your earnings from this task, only after the experiment is completed.

Results for Round 1

The Manager chose **Fraud**, the signal was **Fraud**, and you chose **the Overall Conduct Rule**.
Your earnings in this round were determined by chance and equal: **50 points**

Round	Manager	Signal	Prosecutor	Earnings of Manager	Earnings of Prosecutor	Earnings of Shareholder
1	Fraud	Fraud	Overall	69	50	48

Continue

Raise your hand if you have any questions. Otherwise please click on the "Continue" button.