Prosecutor plea bargaining and conviction rate structure: evidence from an experiment

Jason Ralston[1] · Jason Aimone[2] · Lucas Rentschler[3] · Charles North[2]

Abstract

We present a model of plea bargaining and vary the value a prosecutor places on a conviction obtained via plea bargain relative to a conviction obtained at trial. We show that increasing the relative value of a plea bargain increases the trial penalty and decreases the severity of the equilibrium plea bargain. We report the results of an exploratory experiment which assesses this prediction in a more realistic setting, in which subjects are incentivized by conviction rates. Our treatment variable is whether convictions obtained via plea bargain are included in conviction rate calculations. Including plea bargains in conviction rates increases the number of plea offers made and increases the trial penalty, which is qualitatively in line with our predictions.

1 Introduction

Prosecutors in the United States have tremendous discretion in charging decisions and can use this discretion to bargain with defendants (see e.g., Alschuler (1979)). Prosecutors may offer a reduced charge, and a corresponding reduced sentence recommendation, in exchange for a guilty plea. In practice, more than 90% of convictions are the result of guilty

---

✉ Jason Ralston
ralstonj@whitman.edu

Jason Aimone
jason_aimone@baylor.edu

Lucas Rentschler
lucas.rentschler@usu.edu

Charles North
charles_north@baylor.edu

[1] Department of Economics, Whitman College, Walla Walla, USA

[2] Department of Economics, Baylor University, Waco, USA

[3] Department of Economics and Center for Growth and Opportunity, Utah State University, Logan, USA

pleas (Barkai, 1977; Bibas, 2004). While prosecutorial discretion with respect to plea bargains can move cases through the criminal justice system more quickly and at a lower monetary cost, understanding the effects, including possible unintended consequences, requires an understanding of how prosecutors determine plea offers.[1]

In this paper we adopt a public choice approach to the study of plea bargaining, in which we assume that prosecutors rationally respond to the incentives they face.[2] An immediate question we must address is: what is the objective function of prosecutors? Standard 3–1.2(b) of the ABA's Criminal Justice Standards states, "The primary duty of the prosecutor is to seek justice within the bounds of the law, not merely to convict." In practice, however, prosecutor incentives often center on convictions. Leonetti (2012) states that "conviction rates and sentence lengths are used both as indicators of success and as grounds for retention or promotion... [P]rosecutors with the highest conviction and sentencing statistics are in the best position for career advancement, but those who exercise their discretion to achieve the most just and beneficial outcomes are not." Barkow (2009) and Silverglate (2011) agree that higher conviction rates are more likely to lead to promotions or high-paying private sector work. Anecdotes circulate of rewards to prosecutors for convictions, such as a Colorado district attorney who offered pay bonuses to prosecutors who tried at least five cases a year and won at least 70 percent of them (Fender, 2011),[3] or a report of Louisiana prosecutors who gave each other informal awards for murder convictions (Balko, 2013). Similarly, in a letter to the editor three decades after convicting an innocent Louisiana man of capital murder, the responsible former prosecutor admitted that, at the time, he had been more concerned with "obtaining a conviction of a person who I believed to be guilty" than with pursuing known leads that might have yielded exculpatory evidence (Stroud, 2015).

In this paper, we thus assume that prosecutorial incentives focus on convictions.[4] We study the effect of varying the value a prosecutor places on a conviction obtained via plea bargain relative to a conviction obtained at trial, all else constant. How does prosecutor behavior in plea bargaining depend on this relative value? To address this question, we

---

[1] It is difficult to effectively study plea bargaining with naturally occurring criminal justice data. For instance, cases that are dropped or not pursued will not show up in trial or plea bargaining records. Likewise, it would be difficult, invasive, and expensive (and likely not even possible) to measure or obtain accurate data on evidence strength and case by case options prosecutors weigh when making their decisions. In our view, theory and experiments provide us with an opportunity to control many of these factors so that incentives and decisions can be cleanly explored.

[2] Mueller (2003) defines choice as "the economic study of nonmarket decision making.".

[3] Joy and McMunigal (2011) acknowledge that such bonus schemes are not technically violations of the ABA's Model Rules of Professional Conduct, though they argue for making prosecutor bonuses unethical.

[4] In the literature, it is common to assume that prosecutors focus on convictions and conviction rates. Shamir and Shamir (2012) develop a model of court congestion built on the assumptions that prosecutors maximize either conviction rates or numbers of convictions, while Ferguson-Gilbert (2001) sets forth a general overview of why prosecutors seek to maximize conviction rates. Rasmusen et al. (2009) note that prosecutors often tout their conviction rates when running for elections, and media accounts provide examples of campaign battles over conviction rates (Dallas County DA 2010; Dujardin 2017; Rodricks 2022). There is theoretical and empirical evidence that when prosecutors are beholden to voters, elections can spur increases in the relative number of cases taken to trial relative to those resolved through plea bargaining (Bandyopadhyay and McCannon 2014; Dyke 2007). Gordon and Huber (2002) show that, when information about specific cases is limited, voters who care about Type I and II errors can best achieve their electoral goals by voting for prosecutors who successfully obtain convictions. Further, when prosecutors are not directly accountable to voters, institutional incentives nevertheless encourage them to maintain a high conviction rate (Bibas 2004; Zacharias 1991).

first present a simple model which focuses on how this relative value affects equilibrium behavior in the context of a single case. In this model, we assume that the prosecutor's goal is to obtain a conviction, and that she prefers a harsher sentence.[5] The prosecutor first decides whether they wish to drop charges. If they opt not to do so, they decide whether to plea bargain or not. If the prosecutor decides to bargain, and an agreement is not reached, the prosecutor chooses a charge to take to trial. If an agreement is made, it is implemented. In equilibrium, the case is always resolved via plea bargain. The key insight of our model is that the equilibrium distance between the charge the prosecutor would take to trial (the threat point) and the charge associated with the plea bargain is lower when a prosecutor values a conviction obtained via plea bargain less than a conviction obtained at trial, because the prosecutor demands a more severe charge to compensate for the reduced value of the conviction.

While this prediction is striking, it is important to note that our model does not capture many features of real-world prosecutorial decision making. With this in mind, we also report the results of a laboratory experiment which more closely resembles actual plea bargaining. Our experiment is framed as plea bargaining, and subjects know that their decisions will affect other subjects in the role of defendants. Prosecutor incentives in our experiment focus on conviction rates, rather than the outcome of a single case. Specifically, a prosecutor will receive a high bonus if they have the highest conviction rate in their group, a low bonus if they have the second highest conviction rate, or no bonus. This reflects the fact that prosecutors deal with a large number of cases, and that relative conviction rates are important for promotion. Our treatment variable is whether or not plea bargains are included in the conviction rate. That is, are bonuses determined by the conviction rate at trial, or by the overall conviction rate?

Our experimental design evaluates whether the prediction of our model extends to a more realistic setting. Most importantly, our experiment allows for the possibility that prosecutor preferences are not narrowly focused on convictions. Subjects' homegrown preferences regarding Type 1 and Type 2 errors, and their notions of justice, are relevant to their decisions.[6] Does the prediction of our model hold in such an environment? Our design also allows for an interesting possibility that our model does not account for. When prosecutors are incentivized to focus on their conviction rate, they may cherry pick cases. This is particularly true when the relevant conviction rate does not include plea bargains. Thus our experimental design allows us to evaluate the prediction of our model in the face of this possible confound.

Interestingly, despite the richer and more realistic setting, the qualitative prediction of our model holds in our experimental environment. We find that the distance between an offered plea bargain and the trial threat is higher when plea bargains are included in the conviction rate. We also find that when convictions obtained via plea bargain are not included in the conviction rate, prosecutors forego plea bargaining more frequently. In addition, we find no difference in case selection across methods of calculating the conviction rate.

---

[5] Importantly, note that if the prosecutor only cared about conviction, they could always obtain one by offering a sufficiently lenient plea offer. To ensure that our model is interesting, and to match the convention in the literature, we assume that the prosecutor also prefers a harsher sentence, all else equal.

[6] Some of our sessions also involved law students at Baylor University's law school, to further address the concern that a focus on conviction alone may not reflect prosecutor preferences.

Our results have important implications for the criminal justice system. If prosecutors value convictions obtained via plea bargain as much as conviction at trial, then the trial penalty (the difference between the plea offer and the trial threat) is larger. One concern with such an outcome is that large trial penalties may induce innocent defendants to plead guilty.[7] Further, lower plea offers may undercut crime deterrence, since the expected penalty of offending may be lowered.

We are not alone though in exploring the decision making of prosecutors. Many such studies explore topics such as expectations of trial outcomes, guesses of jury decisions, and what would be the best plea bargains (see review in Charness & DeAngelo, 2018).[8] Charness and DeAngelo (2018) discuss that criminal justice experiments often (though not exclusively) suffer from being either hypothetical in nature (no actual consequential trials/penalties/plea bargains are being implemented from decisions) or the experiment is conducted in neutral frames (participants are not "role-playing".) Our study differs by placing subjects into a framed role as a prosecutor where their decisions will affect other experimental subjects. Our experiment is non-hypothetical in a framed environment, which puts us in an uncommon position among the extant experimental law and economics literature.[9]

2 Motivational model

To explore prosecutorial incentives regarding the value of convictions by plea bargain, we first develop a simple model. Consider a prosecutor who is assigned a case. Since our focus in the model is on how varying prosecutor incentives affects equilibrium outcomes within a given case, we do not differentiate the case by the type of crime or the severity of the offense. The case is accompanied by some evidence of the defendant's guilt, but the prosecutor does not know for certain whether the defendant is guilty or innocent.

Let the available charges be the interval $(0, \overline{c}]$. The probability of obtaining a conviction at trial with charge $c \in (0, \overline{c}]$ is given by $p(c)$, and this is common knowledge. This probability function summarizes the prosecutor's beliefs about the strength of the available evidence of guilt.[10] For simplicity, we assume that $p(\overline{c}) = 0$, that $p(c)$ is continuously differentiable and that $p'(c) < 0$. Finally, we assume that $\lim_{c \to 0^+} p(c) = 1$.

---

[7] A variety of proposals have been made with the intention of curbing guilty pleas by innocent defendants. Reinganum 1988 finds that restricting prosecutorial discretion such that they must offer uniform plea bargains for a given charge can improve welfare; the proposal works by limiting prosecutors' ability to offer larger sentence differentials when cases against defendants are weaker. Another more extreme proposal is to abolish plea bargaining altogether (Schulhofer 1984). Others suggest screening cases prior to referring them to a prosecutor to ensure the available evidence meets a predefined threshold (Wright and Miller 2002). The purpose of such screening would be to allow prosecutors to bargain only with defendants who are more likely to be guilty.

[8] Ever since Becker (1968), many studies have sought to explore the deterrence effect of policies and criminal justice procedures. Zeiler and Puccetti (2018) outlines many such theoretical models and experimental studies that have done so. Recent examples include Cuellar and Rentschler (2023a), Cuellar and Rentschler (2023b), Cuellar and Rentschler (2023c) and Cuellar (2023).

[9] Adamson and Rentschler (2023) is a notable exception, although they focus on policing incentives, rather than prosecutors.

[10] This probability function implicitly accounts for the distribution of evidence of innocence the defendant may present (which the prosecutor does not observe), and how that distribution of evidence may depend on the true guilt or innocence of the defendant.

The prosecutor must first decide whether to drop the charges altogether. If they decide to do so the game ends, and both the defendant and the prosecutor have payoffs of zero. If the prosecutor decides not to drop the charges, they decide whether to plea bargain or not. If they decide not to plea bargain, they choose a charge, $c_s \in (0, \overline{c}]$, and proceed directly to trial. If the defendant is convicted, then the defendant has a payoff of $-c_s$, and the prosecutor has a payoff of c_s. If the defendant is found not guilty, the defendant has a payoff of 0, while the prosecutor has a payoff of $-c_s$.

If the prosecutor does decide to offer a plea bargain, the prosecutor and defendant simultaneously announce a charge they are willing to accept as part of a plea bargain. Let c_p^p be the demand of the prosecutor, and c_p^d be the demand of the defendant. If $c_p^p > c_p^d$, the prosecutor chooses a charge that they will take to trial, $c_t \in [0, \overline{c}]$. A conviction is obtained with probability $p(c_t)$, the game ends, and payoffs are analogous to those outlined above.

If $c_p^p \leq c_p^d$, there is no trial, and a sentence of c_p^p is imposed, and the game ends. The defendant has a payoff of $-c_p^p$, while the prosecutor has a payoff of bc_p^p. The parameter $b > 0$ indicates the value of obtaining a conviction via plea bargain, relative to obtaining it via trial. If $b = 1$, for instance, then obtaining a conviction via plea bargain is weighted just as heavily as a conviction obtained at trial. If, however, $b < 1$, a conviction obtained via plea bargain is not considered to be as valuable to the prosecutor as the same conviction obtained via a trial.

We first consider the prosecutor's choice of c_t. We assume, for simplicity, that the prosecutor's expected payoff at trial is single peaked. Denote the prosecutor's optimal choice as c_t^*. That is,

$$c_t^* \equiv arg \max_{c_t} p(c_t)c_t - (1 - p(c_t))c_t.$$

Turning attention to the prosecutor's demand during the plea-bargaining process, note that their demand must satisfy

$$2p(c_t^*)c_t^* - c_t^* \leq bc_p^p,$$

since the demand must yield at least as much expected utility as the outside option obtained at trial. Similarly, for the defendant, their demand must satisfy

$$-c_t^* p(c_t^*) \leq -c_p^d.$$

To ensure that a plea bargain that is mutually acceptable for both parties exists, we assume that $p(c_t^*) \leq 1/(2 - b)$. This restriction implies that value to the prosecutor of a sentence obtained via plea bargain is not discounted to the point that the prosecutor would prefer to simply go to trial.[11] The set of plea bargains that are mutually acceptable is

$$\left[\frac{2p(c_t^*)c_t^* - c_t^*}{b}, c_t^* p(c_t^*) \right].$$

Following Nash (1950), we focus on the equilibrium corresponding to the midpoint of this interval, so that the equilibrium plea bargain is given by

---

[11] Since, in practice, prosecutors almost always make a plea offer, this assumption is not particularly restrictive. If this assumption fails, the prosecutor would simply decline to plea bargain, and go directly to trial with a charge of c_t^*.

$$c_p^* \equiv \frac{1}{2}\left(\frac{2p(c_t^*)c_t^* - c_t^*}{b} + c_t^* p(c_t^*)\right).$$

Note that if the prosecutor opts to plea bargain, the case will be resolved via plea bargain, and will not go to trial.

We are primarily interested in determining the effect of comparing the case of $b = 1$ (where the prosecutor is indifferent between obtaining a conviction via plea bargaining or via trial), and the case where $b < 1$ (where the prosecutor, all else equal, prefers to obtain conviction at trial). The results of this comparison are striking. Since c_p^* is decreasing in b, and $c_t^* > c_p^*$, the difference between the equilibrium plea offer and the charge at trial is increasing in b. That is, when a prosecutor values a conviction obtained via plea bargain less than a conviction obtained at trial, all else equal, the prosecutor demands a harsher charge in equilibrium, so that the plea bargain is closer to the charge at trial. This is the key prediction of our model and motivates the experimental we report below.

Now turn attention to the case in which the prosecutor opts not to plea bargain. In this event, the prosecutor would simply choose a charge of c_t^*. Since the equilibrium payoff of plea bargaining is strictly higher than the expected payoff of going to trial with a charge of c_t^*, the prosecutor will always plea bargain.[12] Now turning to the prosecutor's first decision, note that the prosecutor will never drop charges, provided $p(c_t^*) \geq 1/(2+b)$, since equilibrium payoffs of plea bargaining are greater than zero in expectation.

3 Experimental design

Our simple model is far from a realistic setting and may trigger concerns about the potential extension of our predictions to real-world situations. In order to mitigate these concerns, we've designed an experiment that lessens the stringency of our model's assumptions. Though not directly validating our theory, this experiment is exploratory in the spirit of Smith (1982). Crucially, the experiment serves as a robustness check for our theoretical predictions.

One key characteristic of our experiment is the framing of the scenario as a plea-bargaining process. It opens up the possibility that prosecutorial motivations may extend beyond a mere focus on conviction rates. Participants' intrinsic preferences regarding Type 1 and Type 2 errors, as well as their personal sense of justice, play significant roles in their decision-making processes. Our design also acknowledges a fascinating potential outcome that our model overlooks: when prosecutors are encouraged to concentrate on their conviction rates, they might adopt a selective approach to case picking and discard those that seem weak. This likelihood is heightened especially when the measured conviction rate excludes plea bargains. Consequently, our experimental design enables us to gauge the prediction accuracy of our model against this plausible complication.

There are several other ways in which our experimental design is more realistic than our model that are worth highlighting. Our experiment features ambiguity regarding the available evidence, which is more realistic than in our simple model. Similarly, the probability

---

[12] To see this, note that the payoff of plea bargaining is bc_p^*, and the expected payoff of going to trial without plea bargaining is $2p(c_t^*)c_t^* - c_t^*$. Since we assume that $p(c_t^*) \leq 1/(2-b)$, the payoff the former exceeds that of the latter.

of obtaining a conviction at trial is not known by prosecutors. Rather, they know that subjects in the role of jurors will evaluate a case at trial, and they will determine the outcome. In addition, potential defendants will have the opportunity to steal real money from other subjects. This combination of features promises to provide a more comprehensive, and potentially revealing, exploration of our theoretical predictions.

3.1 Background

The current experiment on prosecutor decision making is part of a larger research program on the criminal justice system, which also includes studies of juror and defendant decision making, the details of which can be found in Aimone et al. (2019) and Ralston et al. (2019) respectively.[13] We first provide a broad outline of the overarching program and then provide details of the prosecutor experiment that is presented below. In the criminal justice system, crime decisions come first, followed by a bargaining game between prosecutors and defendants within which decisions on charges/plea bargains and trials are made, and finally juror decision making concludes a trial. Wanting to have a salient experimental economic decision environment (one that can bring data to bear upon the theory of the preceding section), we conducted our experimental sessions in the opposite order. We first conducted juror decision-making sessions, as they could credibly and in an incentive compatible way, be conducted without prosecutor or defendant decisions being in place.

3.1.1 Summary of juror experiment sessions

The necessary role of the juror experiment sessions in the three-part experiment, is to generate real saliently incentivized human juror decisions that could be used in defendant trials to form a peer-jury determined verdict of guilty or not-guilty. In these juror experiment sessions (see Aimone et al., 2019), jurors were told that future sessions would be conducted in which real people (potential defendants) would be given an opportunity to take money from other subjects and some people may be accused of a crime and a trial may be held to determine a defendant's guilt status. In brief, jurors were tasked with determining a verdict for every possible trial that they could have been presented. Jurors were told that for these trials three jurors' decisions for that trial situation (crime level, evidence of innocence, and evidence of guilt), would be randomly chosen from the pool of juror decisions until a unanimous decision was drawn and that would determine the defendant's guilt status. Jurors earned a flat rate for participation (we used same rate as real-world first day jury duty pay in our county), and did not earn outcome based pay (like in the real-world as well). We did not need prosecutors or defendants to be present to conduct these studies. We next conducted the prosecutor experiment sessions, which are the focus of the current paper.

---

[13] We conducted two seed sessions as well, which included both prosecutor and defendant participants that enabled us, along with the jury sessions, to pay the prosecutors in the current study using real participant decisions and standard non-deception techniques. These seed sessions are described in further detail in Sect. 3.3.

3.1.2 Summary of prosecutor experiment sessions

The full details of the prosecutor experiment is found in Sect. 3.2 below. Here, we explain the necessary role of the prosecutor experiment sessions in the three-part experiment. We need from the prosecutor experiment sessions (1) participant decisions whether or not to pursue charges against an accused defendant in a case; (2) participant plea bargain offers for a certain case; and (3) participant choices of what charges would be pursued at trial. Both the prosecutor experiment decisions and juror experiment decisions are needed for defendant sessions which were conducted last.

3.1.3 Summary of defendant experiment sessions

The defendant experiment (see Ralston et al., 2019) gave potential defendants a series of opportunities to steal from another participant in a modified dictator game. In each of these opportunities, the computer randomly allocated $10 between one participant (the potential defendant) and another participant and gave the defendant participant the option to take a portion of the funds allocated to the other participant. Whether the participant stole or not, the computer could generate an accusation[14] of a crime, along with crime and evidence types. The defendant reported for the range of potential plea bargain and trial situations what they would do in each situation. If there was an accusation of a crime, they would be randomly assigned a prosecutor's decision (from the already conducted prosecutor study) for their crime and evidence levels associated with the accusation, and the defendant's own decision in the plea or trial situation chosen by that prosecutor would determine what happened. Ultimately, if the decisions of the prosecutor and defendant indicated a trial occurred, a three-person jury would be randomly chosen from the pool of juror decisions (that were collected in the previously run juror experiment sessions) to determine whether the defendant was found guilty or not guilty at their trial.

3.1.4 Evidence generation

The three types of experiment sessions above all center around the idea of a defendant who may be truly innocent or may be truly guilty of taking money from another participant. All three participant types (jurors, prosecutors, and defendants) are concerned with the evidence of that taking (or "crime"). Since these experiment sessions are linked, they all involve the same evidence generation process that differs based upon whether the defendant is truly innocent or truly guilty (a status known to the defendant participant but unknown to the prosecutor and juror participants).

Figure 1 below shows how likely a given defendant would be to have any evidence of innocence and any evidence of guilt. This information was provided to the experimental participants in their instructions (see Appendix A for the prosecutors' instructions.) As can be seen in Fig. 1, while there was a 70% chance that a truly guilty defendant would have some evidence of guilt there was only a 30% chance for a truly innocent defendant to have some evidence of guilt (and therefore be accused of a crime by the computer). Participants in the experiment were also told that those who were truly guilty were less likely to receive

---

[14] We will refer to the computer-generated claim that a crime has occurred as an "accusation." We will refer to the decision by the prosecutor as a "charge.".

any evidence of innocence (30% chance) than those who were truly innocent (who had an 80% chance of having some evidence of innocence.) For example, the figure shows us that there is only a 6% chance a truly innocent individual has some evidence of guilt and no evidence of innocence (0.3*0.2), while there is a 49% chance a truly guilty individual has some evidence of guilt and no evidence of innocence (0.7*0.7).

Participants were also told that there were three sizes of crime (Small, Medium, and Large) and three levels each of evidence of guilt and innocence (Weak, Medium, and Strong). While the probabilities of having some evidence of innocence or guilt were known for both the truly innocent and the truly guilty, prosecutor participants were not told the values of the individual probabilities of Strong, Medium, and Weak evidence. To induce participants' beliefs about such probabilities to be in the same general direction and similar to what would be expected in the "real world", prosecutors were told in regard to evidence of guilt that:

> As the names suggest, STRONG evidence is harder to get (and implies there is more evidence of guilt) than MEDIUM evidence. MEDIUM evidence is harder to get (and implies there is more evidence of guilt) than WEAK evidence. For each of these three levels, a prosecutor is always more likely to get a particular level of evidence of guilt if the defendant is truly guilty compared to a truly innocent defendant.

And in regard to evidence of innocence that:

> As the names suggest, STRONG evidence is harder to get (and implies there is more evidence of innocence) than MEDIUM evidence. MEDIUM evidence is harder to get (and implies there is more evidence of innocence) than WEAK evidence. For each of these three levels, a defendant is always more likely to get a particular level of evidence of innocence if the defendant is truly innocent compared to a truly guilty defendant.[15]

The experiment was designed to mimic conditions found in the real criminal justice system. The reduced level of detail in Fig. 1 about the evidence generation process reflects the lack of detail available to real-world prosecutors, defendants, and jurors about how evidence arises. In a real trial, no party knows exactly how much more likely it is for a truly innocent person to have strong alibi (e.g., CCTV footage showing them at the time of the crime) compared to a medium alibi (e.g., a friend who says the defendant was with them at the time of the crime).[16]

---

[15] For more details about the exact probabilities used to generate evidence of guilt and innocence the reader can see Figure SI1. Remember though, that participants did not receive this level of detail on the evidence generation process.

[16] Note our study uses non-neutral language rather than a more standard neutral language approach. There are factors in the criminal justice realm that could affect decisions that we think would be lost were instructions neutral in language. A "sense of justice" or "responsibility" that comes when one is a prosecutor facing a theft situation would, for instance, be lost in a neutral language environment. Similarly, if defendants' decisions were framed to prosecutors neutrally as "choice A" or "choice B" rather than as crimes or theft, then there would be no aspect of deservingness of punishment. It would be a bargaining environment still, but factors like inequality aversion, altruism, etc. may play very different roles in determining outcomes. We conclude that while a similar experiment could be designed using neutral language, that we could not generalize to the criminal justice system the results from such a study in as believable a manner.

	Prosecutorial Evidence of *Guilt*				Defense Evidence of *Innocence*	
	If Truly Innocent Defendant	If Truly Guilty Defendant			If Truly Innocent Defendant	If Truly Guilty Defendant
Some Evidence of Guilt	30 of every 100 truly innocent	70 of every 100 truly guilty		**Some Evidence of Innocence**	80 of every 100 truly innocent	30 of every 100 truly guilty
No Evidence of Guilt	70 of every 100 truly innocent	30 of every 100 truly guilty		**No Evidence of Innocence**	20 of every 100 truly innocent	70 of every 100 truly guilty

Fig. 1 Evidence generation tables seen by participants

3.2 Prosecutor's decisions

After the creation of the court case from the decisions of the defendants in the modified dictator game,[17] participants in the role of prosecutors were presented with all possible combinations of generated accusations and evidence of guilt. Prosecutors were made aware of the evidence generation process in the same way as participants were in all other experiments. Participants were given two tasks: a belief elicitation and then a prosecutorial decision. Prior to making prosecutorial decisions, prosecutors were shown possible accusations of crime and evidence levels of guilt and asked to provide beliefs about the likelihood a jury would convict a defendant of the accused level of crime based on the evidence of guilt described. Participants were told their answers on these beliefs would not affect their pay in the experiment. We think this task helps ensure participants are cognizant of the connection between differential evidence levels and differential probabilities of conviction at trial, an important connection for prosecutorial plea/trial decisions. We do not discuss these beliefs further in the paper nor do we attempt to control for these beliefs in regressions or analyses. Such controls may be misleading as the beliefs were not incentivized and participants may update these beliefs in different ways as the experiment progresses.

After completing the belief elicitation, prosecutors were again shown every possible combination of accused crime and level of evidence of guilt. For each possible combination, prosecutors chose how to handle the case: drop the case, proceed straight to trial without offering a plea deal, or offer a plea deal.

If prosecutors chose either to continue to trial or to offer a plea deal, they next had to choose what level of crime to charge the defendant with at trial (Small, Medium, or Large), as well as the punishment for this crime (Low or High). This punishment was only implemented if the jury found the defendant guilty at trial. If prosecutors chose to offer a plea deal, they also had to choose a plea offer, consisting of a charged crime and a proposed punishment level, that the defendant could accept in order to avoid the trial. We imposed the condition that severity of the crime in the plea offer (Small, Medium, or Large) must be less than the severity of the crime charged at trial. For example, prosecutors could offer a

---

[17] See Sect. 3.3 below for a description of the seed experiments used to determine prosecutors' payoffs in this experiment.

plea bargain that allowed a defendant to plead guilty to a small crime to avoid going to trial for a medium crime, but they could not offer a plea bargain such that a defendant could plead guilty to a medium crime to avoid going to trial for a small crime.

In choosing between taking the defendant to trial or offering a plea bargain, prosecutors had to decide what charge to level against a defendant. Prosecutors did not necessarily have to charge the same crime as the initial accusation generated by the computer, and they had some leeway in deciding the severity of the charged crime. If prosecutors wished to increase the severity of the charge above the initial accusation, the strength of the evidence of guilt decreased. Conversely, if a prosecutor wished to decrease the severity from the initial accusation, the strength of evidence of guilt increased. This is summarized in Table 1.

When making these decisions, prosecutors were also made aware that defendants received evidence of innocence of some level (None, Weak, Medium, or Strong) that was unobservable to prosecutors. Further, prosecutors were told that the defendants' evidence of innocence would be adjusted if the prosecutor decided to increase or decrease the severity of the charged crime from the initial accusation. Specifically, if a prosecutor increased the severity of the charge, the defendant's evidence of innocence increased in strength. Conversely, if a prosecutor decreased the severity of the charge, the defendant's evidence of innocence decreased.

Each prosecutor made decisions for all possible combinations of crime severities and levels of evidence of guilt. Once completed, prosecutor decisions were later paired with real defendant decisions on how to respond to the prosecution's charges made against them. The details of defendant sessions can be found in Ralston et al. (2019).[18] Following defendant decisions, the case could be resolved by acceptance of a plea bargain or by dismissal of the case. If the case was not resolved at this stage, the case went to a jury trial that drew on decisions previously made in the project's jury experiment. There, participants acting as jurors made individual decisions on guilty/not guilty for all possible combinations of charged crime levels and strengths of evidence of guilt and innocence. In a trial during this paper's phase of the experiment, the individual decisions of three jurors were randomly drawn for the specific combination of charged crime and evidence levels; if all three agreed, the jury had a verdict. If all three did not agree, then a mistrial was declared, and the computer randomly drew another three-person jury. This process continued until all members of a jury agreed on a judgment of either guilty or not guilty. Note that a guilty finding at trial thus depended on the behavior of subjects acting as jurors, and did not feature exogenous and common knowledge probabilities. This is another way in which our experimental design is more realistic than the our simple model. Further details on the jury experiment can be found in Aimone et al. (2019).

Prosecutor participants were paid based on their within-group performance of 100 simulated cases (drawn from real behavior of human defendants in seed sessions.) That is, the participant with the highest conviction rate received the highest payoff, the second highest

---

[18] Note having our lab prosecutors know that their plea and trial decisions can affect real defendants is a critically important aspect of our study. "Real world" prosecutors face not only personal incentives ranging from salary, bonuses, promotions, and career advancements, but they know that their decisions affect real people and that they are partly responsible for ensuring that "justice" is served. We want that in our study as well, which we do when our prosecutors know their decisions can affect real people in the future. If this "sense of justice" is meaningful to people (in the lab or in the "real world"), it might affect prosecutors' decisions and push them away from some behaviors that in some circumstance may be monetarily profit maximizing, like dropping all charges except where prosecutors have strong evidence of guilt.

conviction rate earned the second highest payoff, and all other players only earned show-up and completion fees. In roughly half of sessions, the conviction rate was the number of convictions resulting from a trial divided by the number of cases taken to trial. In the other half of the sessions, in addition to trial results, every plea deal accepted by a defendant counted as a conviction in the numerator of the conviction rate and as a case in the denominator.

3.3 Seed experiments

A natural question while running the prosecutor experiments is how to know whether a defendant would accept a plea bargain, since there were no participants present and playing the role of defendants. For this, we used two seed experiment sessions, which were run before the prosecutor sessions analyzed in this paper. These seed experiments were pen-and-paper versions of the prosecutor and defendant experiments described above. The seed experiments utilized a strategy-elicitation method, so that participant decisions for each possible contingency were recorded and able to be used to determine participant payoffs. The prosecutor decisions in the seed study and the defendant decisions (for both the truly innocent and the truly guilty of taking money from a real victim participant in the seed study) were used to determine each other's payments (and the payments of the victim of theft). The data collected in these seed experiments were used to identify plea acceptance rates for each possible situation that the prosecutors in the current experiment could face. Thus, this study used real, salient data for both prosecutors and defendants when calculating conviction rates. For jury outcome decisions, calculations used the real, human participant juror judgment decision data collected for each possible trial combination of crime level, prosecution evidence of guilt, and defense evidence of innocence. See Aimone et al. (2019) for a detailed description of the juror experiment process and data.

3.4 Procedures/parameters

Data comes from 10 sessions made up of a total of 60 prosecutors (32 in the INCLUDED treatment and 28 in the EXCLUDED treatments). Participants completed the experiment in the computer program ZTREE (Fischbacher, 2007) which included a series of comprehension questions prior to making decisions to help ensure participants understand the instructions. Sessions lasted for about 90 min.[19] Potential participant payoffs varied on the basis of whether the participant was a law student or an undergraduate student. Law students were included to compare the behavior of those with some legal training to the behavior of those with no legal training. 11 participants were law students (8 in the EXCLUDED treatment and 3 in the INCLUDED treatment). The number of law students was ultimately smaller than we hoped and thus while we control for being a law student (see regressions in Table 2 below), we do not analyze the data separately or draw conclusions based off of this participant type. All undergraduate participants were paid a $5 show-up payment and $5 for completing the experiment. In addition, there was a $10 bonus for having the highest

---

[19] After finishing the prosecutor game described here, all participants also participated as prosecutors in a second game identical to the game they played the first time, but where bonuses were paid out randomly rather than based upon the conviction rate (this game always occurred second). We do not discuss that data here. Participants were paid cash for one of the two games, randomly selected.

Table 1 Result of changing charge on final evidence of guilt level

Initial crime level	Initial evidence	Change charge to small	Change charge to medium	Change charge to large
Large	Weak	Strong	Medium	x
	Medium	Strong	Strong	x
	Strong	Strong	Strong	x
Medium	Weak	Medium	x	Not Allowed
	Medium	Strong	x	Weak
	Strong	Strong	x	Medium
Small	Weak	x	Not Allowed	Not Allowed
	Medium	x	Weak	Not Allowed
	Strong	x	Medium	Weak

conviction rate within their group. There was a lower bonus of $7.5 for having the second highest conviction rate within their group. There was no bonus for other prosecutors in a group. All law students were paid a $5 show-up payment and $15 for completing the experiment. In addition, there was a $25 bonus for having highest conviction rate within their group. There was a lower bonus of $15 for having the second highest conviction rate within their group. There was no bonus for other prosecutors in a group.

4 Results

We first report summary statistics, broken down by whether conviction rates included plea bargains, which are contained in Table 2. The first three rows of Table 2 report the frequency with which prosecutors drop the case, go straight to trial, or offer a plea bargain. The second column provides results for the entire sample, the third column (INCLUDED) provides results for the prosecutors whose plea bargains counted toward the conviction rate, and the fourth column (EXCLUDED) provides results for the prosecutors whose plea bargains did not count toward the conviction rate.

Recall that in a given scenario a prosecutor was presented with a crime level, and an evidence of guilt. The first decision a prosecutor made was whether or not to drop the charge altogether. While a prosecutor illustrated in our motivating theory would, in equilibrium, never drop charges, our experiment features a richer and more realistic environment in which conviction rate structure has the potential to affect the initial screening of cases. In our experiment, when conviction rates do not include plea bargains, a prosecutor has little to gain by proceeding with a case they perceive to be weak. If they proceed to trial, they are likely to lose. If they offer a plea bargain, and it is rejected, they are likely to lose at trial (although they could mitigate this risk to some extent by reducing the charge). If a plea bargain is accepted, it does not improve their conviction rate. However, when plea bargains are included in the conviction rate, some cases that would otherwise be dropped are likely to proceed, provided the prosecutor is able to offer terms they expect the defendant to accept.

Interestingly, we find no evidence that the conviction rate structure affects the initial screening of cases. That is, there is no statistically significant difference in the proportion of cases that are dropped when the conviction rate includes plea bargains relative to when it

Table 2 Summary statistics of prosecutor participants by treatment

	Entire Sample	INCLUDED	EXCLUDED
Dropped Case	17.2%	17.9%	16.7%
Straight to Trial	42.2%	32.5%	50.7%
Offered Plea	40.6%	49.6%	32.6%
Trial Penalty	1.34	1.43	1.23
Female	63.3%	50.0%	75%
Conviction Rate[a]	57.6%	51.1%	65.0%

We include conviction rates as part of our summary statistics. However, further analyses on conviction rate are not performed. This paper is concerned with prosecutor behavior, not the final outcomes of the justice system which depend on many confounding factors and are less likely to be externally valid

excludes them ($p = 0.6207$).[20] We further illustrate this result in Fig. 2, which contains bar graphs of this proportion for each crime level and level of evidence.[21] While the proportion of dropped cases does not meaningfully differ by conviction rate structure, the crime and evidence level associated with a scenario are important. In particular, cases are typically only dropped when the level of evidence is weak. Further, the most frequently dropped case involves a small crime with weak evidence, as would be expected.

When prosecutors decide not to drop charges, they have two options: proceed directly to trial or offer a plea bargain. If plea bargains are not included in conviction rates, prosecutors can only ensure that a case will be included in their conviction rate by going to trial, making this option more attractive in cases they perceive to be strong. However, if plea bargains are included in conviction rates, going to trial is less appealing, leading to an expected increase in cases going to trial when plea bargains are not included in conviction rates.

This is confirmed by our findings in the third row of Table 2, where 50.7% of cases go directly to trial when plea bargains are not included in conviction rates. This percentage drops to 32.5% when plea bargains are included in conviction rates, and the difference is statistically significant ($p = 0.0032$). Figure 3 illustrates how this varies based on the level of crime and evidence, and it should be noted that this result is primarily driven by cases with high levels of evidence.[22]

As the proportion of cases that are dropped remains constant regardless of the conviction rate structure, the decrease in the proportion of cases going to trial when plea bargains are included in conviction rates is due to a corresponding increase in the proportion of cases in which a plea bargain is offered. Specifically, the percentage of cases with a plea bargain increases from 32.6 to 49.6%, and this finding is statistically significant ($p = 0.0018$).[23] Figure 4 further demonstrates how the proportion of cases with a plea bargain varies based

---

[20] All tests reported are two-tailed Mann–Whitney tests, unless otherwise noted.

[21] Table 4 in Appendix B contains summary statistics and the results of Mann–Whitney tests comparing the proportion of dropped cases for each possible combination of crime level and level of evidence.

[22] Table 5 in Appendix B contains summary statistics and the results of Mann–Whitney tests comparing the proportion of cases that proceed directly to trial for each possible combination of crime level and level of evidence.

[23] Table 6 in Appendix B breaks this result down by all possible crime and evidence levels.
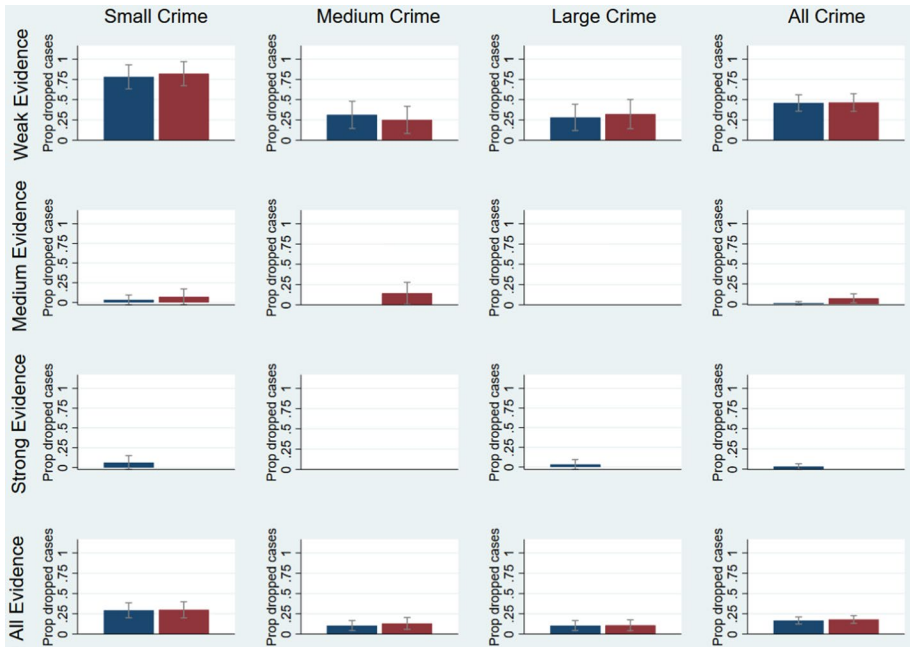
Fig. 2 Proportion of dropped cases. Blue bars indicate the EXCLUDED treatment, while red bars indicate the INCLUDED treatment

on crime level and evidence level. It is important to note that this result is primarily driven by cases with strong evidence of guilt.

Thus far, we have found that the incentivization of plea bargains does not significantly change decisions to drop cases, and that prosecutors in our experiment largely substitute between going straight to trial and offering pleas. We now turn to comparative static from our theoretical framework that primarily motivated our experiment: that incentivizing plea bargains is predicted to result a larger differential between the plea offer and the charge that will be taken to trial if the plea offer is rejected. We define this differential as the trial penalty.

To evaluate the trial penalty in our study, we take into account that all plea bargains consist of a proposed crime level and punishment that the defendant will receive if they accept the plea deal. Additionally, the prosecutor issues a credible threat of a higher crime level and punishment if the defendant rejects the plea bargain and proceeds to trial. To analyze this in our experimental data, we establish an ordinal ranking of the crime level and punishment combinations. As our experiment features three crime levels (small, medium, and large) and two corresponding punishment levels (low and high), there are six possible pairs of crime and punishment levels. We assign the lowest rank, 1, to the least severe combination (small crime and low punishment), and the highest rank, 6, to the most severe combination (large crime and high punishment). This ranking is summarized in Table 3. The trial penalty is then the difference between the ordinal rank associated with the trial threat, and the ordinal rank associated with the plea offer.
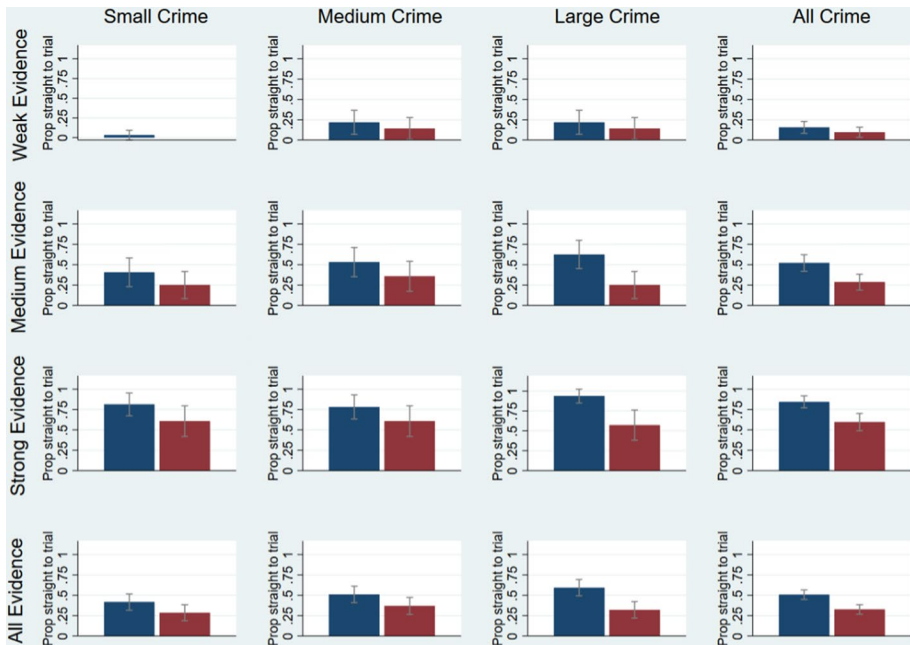
Fig. 3 Proportion of cases taken straight to trial. Blue bars indicate the EXCLUDED treatment, while red bars indicate the INCLUDED treatment

Note that the fourth row of Table 2 demonstrates that, as predicted, the trial penalty increases when plea bargains are incentivized ($p = 0.05$).[24] This is despite the relative complexity of our experimental framework. Given this complexity, it is striking that the predicted increase in the trial penalty is present. That is to say, our experimental results demonstrate that the key comparative static of our model is not a knife-edge result that is sensitive to relaxing the admittedly strong assumptions of our model. Rather, in an environment that is much more general, we observe behavior in line with this prediction. This result is broken down by crime level and evidence level in Fig. 5. It is worth noting that the magnitude of the trial penalty is most affected by the inclusion of plea bargains into the conviction rate when it is a large or medium crime accompanied with strong evidence of guilt.

Our experimental results show that when prosecutors are incentivized to offer plea bargains, prosecutors not only make more plea offers, but they also make plea offers with a larger trial penalty. Is the trial penalty increased due to an increase in the trial threat, a decrease in the plea offer, or some combination of the two?

While there is considerable heterogeneity across subjects in our experiment, we find that the increased trial threat is the primary driver of the increased trial penalty: including plea bargains in the conviction rate increases the severity of the trial threat from 3.14 to 3.58, although this result is only marginally significant using a two-tailed Mann–Whitney test ($p = 0.0593$). Interestingly, including plea bargains in the conviction rate results in slightly harsher plea offers, increasing from 1.91 to 2.15, although this result is not significant

---

[24] Table 7 in Appendix B reports summary statistics and Mann–Whitney tests of this difference for each possible level of crime and evidence.

Table 3 Ordinal ranking of crime punishment levels

| Crime level	Punishment level	Ordinal rank
Small	Low	1
Small	High	2
Medium	Low	3
Medium	High	4
Large	Low	5
Large	High	6

Fig. 4 Proportion of cases where pleas were offered. Blue bars indicate the EXCLUDED treatment, while red bars indicate the INCLUDED treatment

(p = 0.2332). Tables 8 and 9 in Appendix B and Figs. 6 and 7 in Appendix C break these results down by crime level and evidence level.

5 Discussion

Plea bargains are beneficial to the criminal justice system for many reasons. From a voters' perspective, a criminal trial costs time and money. An early study found that criminal trials in nine state courts consumed an average of 11 h of court time, with actual times varying widely by offense and jurisdiction (Sipes et al., 1988). Criminal trials in federal courts lasted four-and-a-half days in the early 1990s (Cook et al., 1995). In 1998, the Los Angeles Times reported that the average criminal trial lasted two weeks, and that the daily cost of
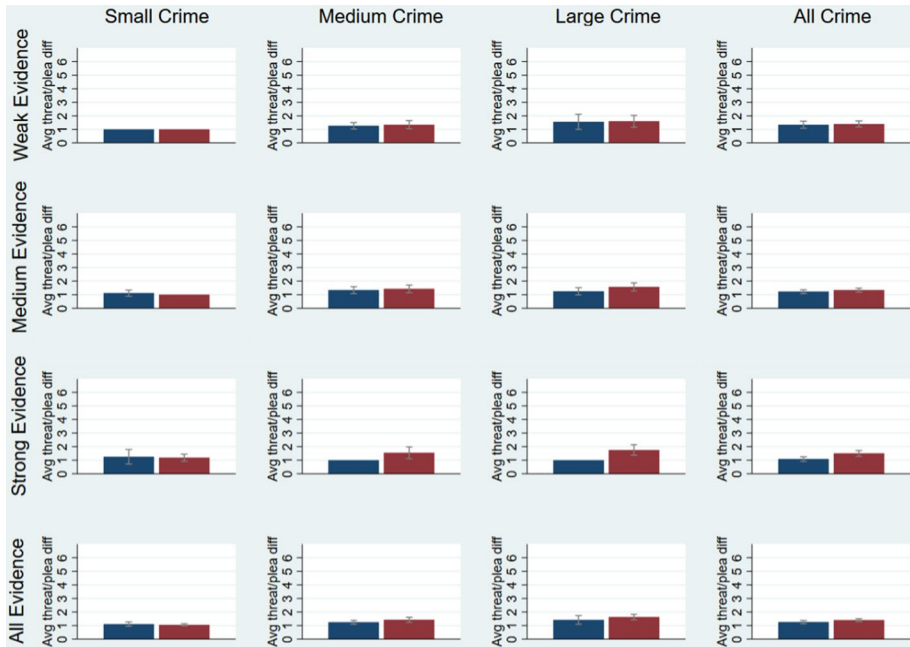
Fig. 5 Average plea/threat differential. Blue bars indicate the EXCLUDED treatment, while red bars indicate the INCLUDED treatment

a Los Angeles Superior Court was just under $9500 ("The Cost of Justice, 1998). A 1999 study of nine state courts found that the average duration of a criminal case was 161 days when resolved via a guilty plea, compared to 272 days when resolved with a trial (Ostrom & Hanson, 1999). Plea bargains reduce the costs of deciding cases, and they place fewer demands on citizens who would otherwise be called for jury duty.

Plea bargains introduce costs into the criminal justice system as well.[25] If prosecutors can pursue more marginal cases (i.e., weaker evidence) due to plea bargaining, there may be a corresponding increase in innocent people embroiled in the criminal justice system. There is a large literature arguing that the plea bargaining process may contribute to the innocence problem, in part because of altered prosecutorial incentives (see. e.g. Bibas, 2004).

---

[25] These costs may be particularly salient for poorer defendants who cannot afford bail/bond (and have to make plea bargain decisions from behind bars) or who rely upon public defenders or court-appointed attorneys for legal defense. Court-appointed attorneys are often paid a flat rate per case whether it is decided by plea or by trial, which may incentivize them to encourage clients to take plea bargains to reduce the time spent on the case. In contrast, an hourly-paid private defense attorney may prefer more time-consuming trials and therefore advise against plea bargains that are bad for the defendant. Indeed, many criminal defense attorneys handle both hourly and court-appointed cases, so any given attorney may face conflicting incentives across clients. See Agan et al. (2018) for a discussion of criminal defense lawyer incentives and citations to additional literature. Many additional factors, such as poverty, race, fear, gender, risk preferences, etc., could influence a person's willingness to take a plea bargain, even though these factors are irrelevant to proving the elements of the charged crimes. Experimental studies show that bargaining behavior depends on individual risk preferences, on the stake size, and on the bargaining power the individual possesses (McCannon et al. 2016).

This paper aims to improve our understanding of the impacts of prosecutorial incentives on plea bargaining. Adopting a public choice perspective, we focus on how changing the relative value a prosecutor places on a conviction obtained via plea bargain influences outcomes in plea bargains.

We present a simple model that features a prosecutor handling a single case. Our findings suggest that if plea bargains are considered less valuable relative to trial convictions, the gap—referred to as the "trial penalty"—between the plea bargain offer and the trial threat decreases.

We experimentally investigate this prediction in a richer and more realistic environment than our model allows, in order to assess the robustness of our theoretical predictions. Our experiment provides extrinsic incentives to focus on conviction rates, and compares prosecutor decisions when plea bargains are and are not included in the calculation of the conviction rate.

Our findings provide compelling evidence that incorporating plea bargains into the calculation of conviction rates spurs a greater frequency of plea offers. Moreover, consistent with the prediction of our model, it expands the trial penalty. These results potentially shed light on the role of prosecutorial incentives in plea bargaining and suggest that the metrics used to evaluate prosecutorial performance can significantly influence plea bargaining outcomes. Future work may find it beneficial to explore whether conviction rates vary with the incentivization of plea bargains using naturally occurring empirical data.

Appendix A: Experimental instructions

We include sample instructions used in a typical experimental session. The particular instructions included are those used for sessions run with law students using accepted guilty pleas and guilty verdicts in the calculation of the conviction rate. In a session where only guilty verdicts were used to calculate the conviction rate, the instructions would have clearly stated so.

The only within experiment differences between law student and undergraduate student instructions were the rewards for performance. Instead of being rewarded with $25 for obtaining the highest conviction rate in a group, undergraduates were rewarded with $10. Similarly, a second place finish was rewarded with $7.5 instead of $15.

Group sizes were variable, starting at 3 at a minimum and being as large as 5. The size depended on how many participants showed up for a particular session. Again, the language used in the instructions would change to specify the correct group size.

Instructions

Thank you for participating in today's experiment. You will receive a $5 payment for coming to the lab today. In addition, you will be paid $15 for participating and completing today's experiment. You may also obtain an additional bonus of either $25 or $15.

You will be in the role of a prosecutor in today's experiment. You are in a group of participants who all have this role. One of your group will receive the bonus of $25, and another participant in your group will receive the bonus of $15.

Your decisions will be used, together with the decisions of other participants in the same role as you, to determine outcomes for participants in a separate set of future experiments.

The nature of the future experiments, which will be explained to you in detail in a few moments, will involve situations in which participants have the opportunity to take money (real US$) which does not belong to them from another participant (from now on, we will refer to this as a theft).

You will provide decisions that determine the monetary punishment, if any, for a participant who has been accused of a theft (from now on, we will refer to a participant who has been accused as a defendant).

Each scenario for which you will provide decisions will involve a defendant. You will have some level of evidence of his or her guilt. The defendant may also have evidence of his or her innocence.

You will decide whether to drop the charges against the defendant or to prosecute the defendant. If you decide to prosecute the defendant you will determine the severity of the charge as well as the possible monetary punishments he or she would face if found guilty. You will also be able to offer a plea bargain. A plea bargain is a charge and a level of monetary punishment that is offered to the defendant if he or she pleads guilty.

If the accused participant decides to plead not-guilty when charged with a theft, there is a trial. In a trial, the evidence of guilt you have, as well as any evidence of innocence the defendant provides, will result in either a guilty or not-guilty finding. The decisions of participants in the role of jurors will be used to make such findings. Participants in the role of jurors have been instructed to make a finding of guilty only if the evidence presented proves guilt beyond a reasonable doubt.

Please take these tasks seriously as real people will be affected by your decisions.

As mentioned earlier you will be paid in cash for your decisions as a prosecutor today. After you finish making your decisions we will ask you to fill out a short survey and you will be paid in cash as you leave the lab.

The decisions you make today will be used in the event that a participant in a future experiment is accused of a theft in the exact manner described to you today. That is, your decisions will only be used in the exact situations that you will evaluate in this experiment.

In this future experiment, $10 will be divided between two participants who are partners. The computer will randomly choose a division of the $10 which allocates a portion of the money to each of the two participants. This preset division is revealed to one of the participants. This participant then must report the preset division chosen by the computer. The division reported by the participant will be implemented. However this participant can choose to take some of the money that is supposed to go to his or her partner by misrepresenting the preset division chosen by the computer.

After the participant reports the division of money (which may or may not correspond to the preset division chosen by the computer) the computer will generate either evidence of a theft or no evidence of a theft. If there is evidence of a theft, the computer will make an accusation. An accusation can be of a SMALL theft (a theft between $0.10 and $1.00), a MEDIUM theft (a theft between $1.10 and $2.00) or a LARGE theft (a theft between $2.10 and $3.00).

Note that an accusation may or may not be true. While a defendant will know for certain whether or not they are guilty of a theft, neither the prosecutor nor the jurors know for certain.

The computer is more likely to make an accusation if the defendant is truly guilty. In particular, 70% of truly guilty defendants will receive an accusation, and 30% of truly innocent defendants will receive an accusation.

If the computer makes an accusation, then there is some evidence of guilt to support this accusation. This evidence of guilt can be WEAK, MEDIUM or STRONG.

As the names suggest, STRONG evidence is harder to get (and implies there is more evidence of guilt) than MEDIUM evidence. MEDIUM evidence is harder to get (and implies there is more evidence of guilt) than WEAK evidence. For each of these three levels, a prosecutor is always more likely to get a particular level of evidence of guilt if the defendant is truly guilty compared to a truly innocent defendant.

If the computer makes an accusation, then the defendant may or may not have some evidence of innocence. Evidence of innocence can be WEAK, MEDIUM or STRONG. As the names suggest, STRONG evidence is harder to get (and implies there is more evidence of innocence) than MEDIUM evidence. MEDIUM evidence is harder to get (and implies there is more evidence of innocence) than WEAK evidence. For each of these three levels, a defendant is always more likely to get a particular level of evidence of innocence if the defendant is truly innocent compared to a truly guilty defendant.

A truly guilty defendant will have no evidence of innocence 70% of the time (note that this means that a truly guilty defendant will have some evidence of innocence 30% of the time). A truly innocent defendant will have some evidence of innocence 80% of the time (so that a truly innocent defendant will have no evidence of innocence 20% of the time).

This information is summarized in the tables below.

	Prosecutorial Evidence of *Guilt*			Defense Evidence of *Innocence*	
	If Truly Innocent Defendant	If Truly Guilty Defendant		If Truly Innocent Defendant	If Truly Guilty Defendant
Some Evidence of Guilt	30 of every 100 truly innocent	70 of every 100 truly guilty	**Some Evidence of Innocence**	80 of every 100 truly innocent	30 of every 100 truly guilty
No Evidence of Guilt	70 of every 100 truly innocent	30 of every 100 truly guilty	**No Evidence of Innocence**	20 of every 100 truly innocent	70 of every 100 truly guilty

If, after an accusation, a defendant goes to trial for a given level of theft (SMALL, MEDIUM or LARGE) then the evidence of guilt available to support that charge (WEAK, MEDIUM or STRONG), as well as any evidence of innocence the defendant decides to provide (WEAK, MEDIUM or STRONG) is shown to three jurors. These jurors independently evaluate all evidence presented, and are instructed to make a recommendation of a guilty finding only if they are convinced of the defendant's guilt "beyond a reasonable doubt." Jurors are told that: "Proof beyond a reasonable doubt is proof of such a convincing character that you would be willing to rely and act upon it without hesitation in making the most important decisions of your own affairs." If all three juror recommendations are the same, then the unanimous decision determines whether the defendant is found guilty (if all three jurors chose "Guilty") or not-guilty (if all three jurors chose "Not-Guilty"). If all three jurors' decisions are not the same, then the trial is called a "mistrial" and the computer will randomly select another alternative

set of three jurors. This process continues until the first time the decisions of all three selected juror decisions match.

In an earlier experiment, jurors have already made recommendations regarding guilty findings in all the possible trials that could arise in today's experiment. It is the decisions from this earlier experiment that will be used to determine the outcome of trials in today's experiment.

If, after an accusation, a defendant is found guilty, then a monetary punishment is subtracted from his or her earnings in the experiment. This monetary punishment is higher for larger thefts. For each level of thefts, there is a low interval of possible punishments, and a high interval of possible punishments. The possible punishments for each level of theft are summarized in the table below.

	Low interval of possible punishment	High interval of possible punishment
SMALL theft	$0.10–$0.30	$0.40–$0.60
MEDIUM theft	$0.60–$0.80	$0.90–$1.10
LARGE theft	$1.10–$1.30	$1.40–$1.60

If a defendant is accused of a certain level of theft, and there is a given level of evidence of guilt to support this accusation, the defendant may be charged with a different level of theft. If a defendant is charged with level of theft that is different than the initial accusation, then the evidence of guilt in support of this different theft is reduced accordingly.

For example, suppose a defendant is accused of a SMALL theft, and there is STRONG evidence of guilt in support of this accusation. This defendant could be charged with a MEDIUM theft, and there would be MEDIUM evidence of guilt to support this charge. This defendant could also be charged with a LARGE theft, and there would be WEAK evidence of guilt in support of this charge.

If a defendant is accused of a SMALL theft, and there is MEDIUM evidence of guilt in support of this accusation, this defendant could be charged with a MEDIUM theft, and there would be WEAK evidence of guilt to support this charge. However, this defendant could not be charged with a LARGE theft, because there is insufficient evidence of guilt to support this charge.

If a defendant is accused of a SMALL theft, and there is WEAK evidence of guilt in support of this accusation, then this defendant cannot be charged with any other level of theft. This is because there is insufficient evidence of guilt to support any other charge.

To provide another example, suppose a defendant is accused of a MEDIUM theft, and there is STRONG evidence of guilt in support of this accusation. This defendant could be charged with a LARGE theft, and there would be MEDIUM evidence of guilt to support this charge. This defendant could also be charged with a SMALL theft, and there would be STRONG evidence of guilt in support of this charge (there is no higher level of evidence of guilt than STRONG).

If a defendant is accused of a MEDIUM theft, and there is MEDIUM evidence of guilt in support of this accusation, this defendant could be charged with a SMALL theft, and there would be STRONG evidence of guilt to support this charge. This defendant could also be charged with a LARGE theft, and there is WEAK evidence of guilt to support this charge.

If a defendant is accused of a MEDIUM theft, and there is WEAK evidence of guilt in support of this accusation, then this defendant could be charged with a SMALL theft, and there is MEDIUM evidence of guilt to support this charge. However, this defendant

cannot be charged with a LARGE theft, because there is insufficient evidence to support the charge.

Lastly, suppose a defendant is accused of a LARGE theft, and there is STRONG evidence of guilt in support of this accusation. This defendant could be charged with either a SMALL or MEDIUM theft, and there would be STRONG evidence of guilt to support either of these charges (there is no higher level of evidence of guilt than STRONG).

Similarly, if a defendant is accused of a LARGE theft, and there is MEDIUM evidence of guilt in support of this accusation, this defendant could be charged with either a SMALL or MEDIUM theft, and there would be STRONG evidence of guilt to support either of these charges.

If a defendant is accused of a LARGE theft, and there is WEAK evidence of guilt in support of this accusation, then this defendant could be charged with a MEDIUM theft, and there would be MEDIUM evidence of guilt to support the charge. This defendant could also be charged with a SMALL theft, and there would be STRONG evidence of guilt to support this charge.

It is important to note that if a defendant is charged with a level of theft that differs from the initial accusation, then his or her level of evidence of innocence also adjusts accordingly.

For example, suppose a defendant is accused of a SMALL theft, and has STRONG evidence of innocence against this accusation. If he or she were charged with either a MEDIUM or LARGE theft, there would be STRONG evidence of innocence against either of these charges.

Suppose a defendant is accused of a SMALL theft, and has MEDIUM evidence of innocence against this accusation. If he or she were charged with either a MEDIUM or LARGE theft, there would be STRONG evidence of innocence against either of these charges.

Suppose a defendant is accused of a SMALL theft, and there is WEAK evidence of innocence against this accusation. If the defendant were charged with a MEDIUM theft, then there would be MEDIUM evidence of innocence against this charge. If this defendant were charged with a LARGE theft, there would be STRONG evidence of innocence against this charge.

Suppose a defendant is accused of a SMALL theft, and there is NO evidence of innocence against this accusation. If the defendant were charged with a MEDIUM theft, then there would be WEAK evidence of innocence against this charge. If this defendant were charged with a LARGE theft, there would be MEDIUM evidence of innocence against this charge.

Suppose a defendant is accused of a MEDIUM theft, and there is STRONG evidence of innocence against this accusation. If this defendant were charged with a LARGE theft, there would be STRONG evidence of innocence against this charge. If this defendant were charged with a SMALL theft, there would be MEDIUM evidence of innocence against this charge.

Suppose a defendant is accused of a MEDIUM theft, and there is MEDIUM evidence of innocence against this accusation. If this defendant were charged with a SMALL theft, and there would be WEAK evidence of innocence against this charge. If this defendant were charged with a LARGE theft, there would be STRONG evidence of innocence against this charge.

Suppose a defendant is accused of a MEDIUM theft, and there is WEAK evidence of innocence against this accusation. If this defendant were charged with a SMALL theft, there would be NO evidence of innocence against this charge. If this defendant were

charged with LARGE theft, there would be MEDIUM evidence of innocence against this charge.

Suppose a defendant is accused of a MEDIUM theft, and there is NO evidence of innocence against this accusation. If this defendant were charged with a SMALL theft, there would be NO evidence of innocence against this charge. If this defendant were charged with LARGE theft, there would be WEAK evidence of innocence against this charge.

Suppose a defendant is accused of a LARGE theft, and there is STRONG evidence of innocence against this accusation. If this defendant were charged with a MEDIUM theft, there would be MEDIUM evidence of innocence against this charge. If this defendant were charged with a SMALL theft, there would be WEAK evidence of innocence against this charge.

Suppose a defendant is accused of a LARGE theft, and there is MEDIUM evidence of innocence against this accusation. If this defendant were charged with a MEDIUM theft, there would be WEAK evidence of innocence against this charge. If this defendant were charged with a SMALL theft, there would be NO evidence of innocence against this charge.

Suppose a defendant is accused of a LARGE theft, and there is WEAK evidence of innocence against this accusation. If this defendant were charged with either a MEDIUM or SMALL theft, there would be NO evidence of innocence against either of these charges.

Suppose a defendant is accused of a LARGE theft, and there is NO evidence of innocence against this accusation. If this defendant were charged with either a MEDIUM or SMALL theft, there would be NO evidence of innocence against either of these charges.

When a participant in the future experiment is accused of a theft (either SMALL, MEDIUM or LARGE), we will use your decisions today to determine what monetary punishment, if any, they will face.

Acting in the role of a prosecutor, you will observe two pieces of information before you make any decisions: the accused level of theft (SMALL, MEDIUM or LARGE) and the level of evidence of guilt in support of the accusation. Remember that you will not know for certain if the defendant is guilty or not. You will not know the level of evidence of innocence, if any, that the defendant has at the time that you make your decisions. However, the defendant will be aware of the level of evidence of guilt that you, as the prosecutor have. Defendants may also choose to exercise their right to not testify on their own behalf. Note: all jurors were instructed that:

> "A defendant has the option to make their evidence of innocence available or not. This right to not "testify" is guaranteed by the United States Constitution and should not be taken as implying guilt."

After observing the accused level of theft (SMALL, MEDIUM or LARGE) and the level of evidence of guilt in support of the accusation, you will decide between three options:

Drop the charge In this case, there is no trial, and the defendant does not face any monetary punishment.

Proceed to trial without offering a plea bargain In this option, you choose a level of theft to charge the defendant with (SMALL, MEDIUM or LARGE theft). In addition, you will choose whether the defendant will pay a monetary punishment from the low interval or the high interval corresponding to the level of theft you are charging him or her with. Of course, the defendant will not pay any monetary penalty if he or she is found not-guilty by the jury. Also, remember that if the charge differs from the accusation, then that the evidence of guilt and the (unobserved) evidence of innocence is adjusted accordingly.

Offer a plea bargain A plea bargain is an offer that you make to a defendant in order to avoid a trial. It consists of a charge (SMALL, MEDIUM or LARGE theft) and associated monetary punishment that you are willing to offer to the defendant in exchange for a guilty plea. If you choose to offer a plea bargain you also choose a charge (SMALL, MEDIUM or LARGE theft) and associated interval of monetary punishment that you will take to trial if the defendant does not accept your offered plea bargain. When choosing these things, you cannot opt for a plea bargain that compares unfavorably with the charge and interval of punishment that go to trial if the plea bargain is rejected.

You will make decisions in every possible scenario a prosecutor could face. Since there are three possible levels of theft, and three possible levels of evidence of guilt, that means you will make decisions in nine different scenarios.

After you have made decisions in each of the nine scenarios, you will go through all nine scenarios for a second time. We will refer to first set of decisions as cycle one, and the second set of decisions (in the same nine scenarios) as cycle two. At the end of the experiment one of these two cycles will be chosen to determine payments.

We will use your decisions today to determine outcomes in future experiments. To help determine payments in today's experiment we will use the decisions of participants in the roles of defendants and jurors collected in a past experiment. We will compare your decisions with instances in which past defendants made decisions for when they were accused of a theft (SMALL, MEDIUM or LARGE), there was some evidence of guilt in support of the accusation (WEAK, MEDIUM or STRONG), and there was either no evidence of innocence or the level of evidence of innocence was WEAK, MEDIUM or STRONG.

We will use your decisions in that exact scenario to determine what the outcome would have been if you had been matched with this past defendant. If there would be a trial, the decisions of jurors from past experiments will be used to determine whether or not the trial results in a finding of guilty or not-guilty.

We will match your decisions in cycle one with one hundred instances in which past defendants have been accused, and determine the outcome as described above. We will then match your decisions in cycle two against these same one hundred instances.

For each of these two cycles, we will calculate your conviction rate, which is the percentage of instances in which you opted to charge a defendant with a crime, and there was a guilty finding. A guilty finding occurs when a defendant pleads guilty or is found guilty in a trial. Note that accepted plea bargains are included in the calculation of this conviction rate.

If cycle one is randomly chosen for payment, then the $25 bonus goes to the prosecutor (in your group of 3) who had the highest conviction rate. The bonus of $15 goes to the prosecutor (in your group of 3) who had the second highest conviction rate. The remaining prosecutor in your group do not get a bonus.

If cycle two is randomly chosen for payment, then the recipients of the $25 bonus and the $15 bonus (among your group of 3) is determined randomly, with each person in your group of 3 having an equal probability of receiving either bonus. No one person can receive both bonuses.

Before you make decisions in cycle one, we will also ask you to tell us how likely a prosecutor is to obtain a verdict of guilty at trial in several situations. Your answers to these questions will not affect your payoffs in any way.

Summary

1. In today's experiment you will be in the role of a prosecutor.
2. You will be asked to make decisions in nine scenarios. In a given scenario you will observe two pieces of information: the level of theft a defendant is accused of (SMALL, MEDIUM or LARGE), and the level of evidence of guilt to support that accusation (WEAK, MEDIUM or STRONG).
3. You will choose to either: (1) drop the charge, (2) proceed to trial without offering a plea bargain, (3) offer a plea bargain.
4. The decisions of the prosecutors, in today's experiment, will be used, in future experiments to affect the real monetary pay of real participants who are the role of defendants in those future experiments.
5. Your decisions as a prosecutor will be matched with the decisions of real participants of past experiments in the roles of defendants and jurors to determine the outcomes of accusations. We will compare your decisions with those of defendants in 100 accusations.
6. You will provide decisions for all nine scenarios twice. The first time you provide answers is cycle one. The second time you provide answers is cycle two.
7. For each of these two cycles, we will calculate your conviction rate, which is the percentage of instances in which you opted to charge a defendant with a crime, and there was a guilty finding. A guilty finding occurs when a defendant pleads guilty or is found guilty in a trial. Note that accepted plea bargains are included in the calculation of this conviction rate.
8. If cycle one is randomly chosen for payment, then the $25 bonus goes to the prosecutor (in your group of 3) who had the highest conviction rate. The bonus of $15 goes to the prosecutor (in your group of 3) who had the second highest conviction rate. The remaining prosecutor in your group does not get a bonus. If cycle two is randomly chosen for payment, then the recipients of the $25 bonus and the $15 bonus (among your group of 3) is determined randomly, with each person in your group of 3 having an equal probability of receiving either bonus. No one person can receive both bonuses.

Appendix B

Table 4 Proportion of dropped cases by crime and evidence level

		Low evidence	Medium evidence	High evidence
Small crime	INCLUDED	0.821	0.071	0.000
	EXCLUDED	0.781	0.031	0.063
	p-value	*0.952*	*0.898*	*0.561*
Medium crime	INCLUDED	0.250	0.143	0.000
	EXCLUDED	0.313	0.000	0.000
	p-value	*0.806*	*0.084*	*1.000*
Large crime	INCLUDED	0.321	0.000	0.000
	EXCLUDED	0.281	0.000	0.031
	p-value	*0.953*	*1.000*	*1.000*

p-values are for two-tailed Mann–Whitney tests

Marginal significance is defined as $0.05 < p < 0.1$

Table 5 Proportion of cases that proceeded directly to trial

		Low evidence	Medium evidence	High evidence
Small crime	INCLUDED	0.000	0.250	0.607
	EXCLUDED	0.031	0.406	0.813
	p-value	*1.000*	*0.314*	*0.140*
Medium crime	INCLUDED	0.143	0.357	0.607
	EXCLUDED	0.219	0.531	0.781
	p-value	*0.677*	*0.275*	*0.236*
Large crime	INCLUDED	0.143	0.250	0.571
	EXCLUDED	0.219	0.625	0.938
	p-value	*0.677*	*0.007*	*0.002*

p-values are for two-tailed Mann–Whitney tests

Marginal significance is defined as $0.05 < p < 0.1$

Table 6 Proportion of cases that received a plea offer

		Low evidence	Medium evidence	High evidence
Small crime	INCLUDED	0.179	0.679	0.393
	EXCLUDED	0.188	0.563	0.125
	p-value	*1.000*	*0.513*	*0.036*
Medium crime	INCLUDED	0.607	0.500	0.393
	EXCLUDED	0.469	0.469	0.219
	p-value	*0.417*	*1.000*	*0.236*
Large crime	INCLUDED	0.536	0.750	0.429
	EXCLUDED	0.500	0.375	0.031
	p-value	*0.986*	*0.007*	*0.000*

p-values are for two-tailed Mann–Whitney tests

Marginal significance is defined as $0.05 < p < 0.1$

Table 7 Trial penalty

		Low evidence	Medium evidence	High evidence
Small crime	INCLUDED	1.000	1.000	1.182
	EXCLUDED	1.000	1.111	1.250
	p-value	*1.000*	*0.973*	*1.000*
Medium crime	INCLUDED	1.353	1.429	1.545
	EXCLUDED	1.267	1.333	1.000
	p-value	*0.910*	*0.884*	*0.108*
Large crime	INCLUDED	1.600	1.571	1.750
	EXCLUDED	1.563	1.250	1.000
	p-value	*0.691*	*0.251*	*0.769*

p-values are for two-tailed Mann–Whitney tests

Marginal significance is defined as 0.05 < p < 0.1

Table 8 Plea offers

		Low evidence	Medium evidence	High evidence
Small crime	INCLUDED	1.000	1.158	1.455
	EXCLUDED	1.000	1.056	1.500
	p-value	*1.000*	*0.778*	*0.813*
Medium crime	INCLUDED	1.471	1.929	2.636
	EXCLUDED	1.333	2.400	3.429
	p-value	*0.784*	*0.194*	*0.064*
Large crime	INCLUDED	2.200	2.762	3.667
	EXCLUDED	2.125	3.583	5.000
	p-value	*0.960*	*0.015*	*0.308*

p-values are for two-tailed Mann–Whitney tests

Marginal significance is defined as 0.05 < p < 0.1

Table 9 Trial threat

|		Low evidence	Medium evidence	High evidence
Small crime	INCLUDED	2.000	2.158	2.636
	EXCLUDED	2.000	2.167	2.750
	p-value	*1.000*	*1.000*	*0.462*
Medium crime	INCLUDED	2.824	3.357	4.182
	EXCLUDED	2.600	3.733	4.429
	p-value	*0.532*	*0.312*	*0.648*
Large crime	INCLUDED	3.800	4.333	5.417
	EXCLUDED	3.688	4.833	6.000
	p-value	*0.876*	*0.109*	*1.000*

p-values are for two-tailed Mann–Whitney tests

Marginal significance is defined as 0.05 < p < 0.1

Appendix C

See Figs. 6 and 7.
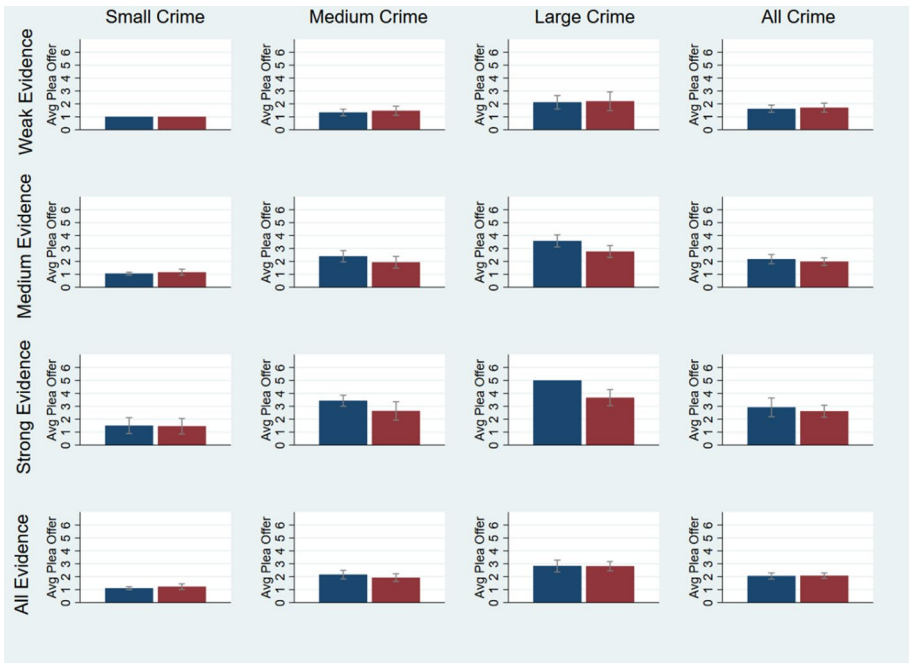
Fig. 6 Average plea bargain. Blue bars indicate the EXCLUDED treatment, while red bars indicate the INCLUDED treatment
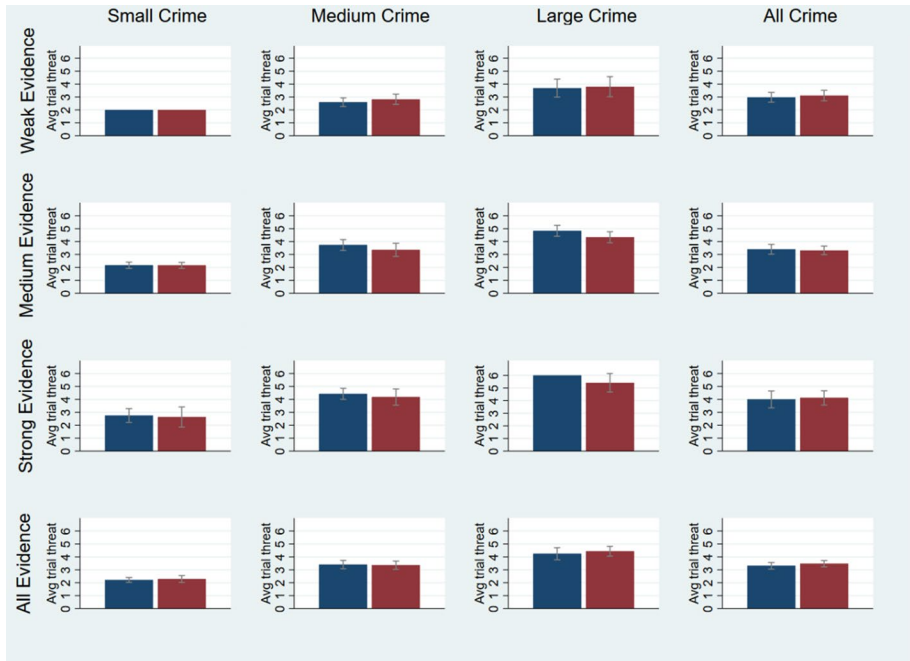
Fig. 7 Average trial threat. Blue bars indicate the EXCLUDED treatment, while red bars indicate the INCLUDED treatment

References

Adamson, J., & Rentschler, L. (2023). How policing incentives affect crime, measurement, and justice. Working Paper.

Agan, A., Freedman, M., & Owens, E. (2018). Is your lawyer a lemon? Incentives and selection in the public provision of criminal Defense. *The Review of Economics and Statistics, 103*(2), 294–309.

Aimone, J., Rentschler, L., & North, C. (2019). Priming the jury by asking for donations: An empirical and experimental study. *Journal of Economic Behavior and Organization, 160*, 158–167.

Alschuler, A. W. (1979). Plea bargaining and its history. *Columbia Law Review, 79*, 1–43.

Balko, R. (2013). The untouchables: America's misbehaving prosecutors, and the system that protects them. Huffington Post.

Bandyopadhyay, S., & McCannon, B. C. (2014). The effect of the election of prosecutors on criminal trials. *Public Choice, 161*, 141–156.

Barkai, J. L. (1977). Accuracy inquiries for all felony and misdemeanor pleas: Voluntary pleas but innocent defendants? *University of Pennsylvania Law Review, 126*, 88–146.

Barkow, R. E. (2009). Institutional design and the policing of prosecutors: Lessons from administrative law. *Stanford Law Review, 61*(4), 869–922.

Becker, G. S. (1968). *Crime and punishment: An economic approach* (pp. 13–68). The Economic Dimensions of Crime.

Bibas, S. (2004). Plea bargaining outside the shadow of trial. *Harvard Law Review, 117*, 2463–2547.

Charness, G., & DeAngelo, G. (2018). *Law and economics in the laboratory*. Edward Elgar Publishing.

Cook, D. L., Schlesinger, S. R., Bak, T. J., & William, T. (1995). Rule II. Criminal caseload in U.S. district courts: more than meets the eye. *American University Law Review, 44*, 1579–1598.

Cuellar, P., & Rentschler, L. (2023c). Plea bargaining with prosecutorial commitment. Working Paper.

Cuellar, P. (2023). Voluntary disclosure of evidence in plea bargaining. *Working Paper*.

Cuellar, P., & Rentschler, L. (2023b). Adversarial Bargaining with Exogenous Shocks. Working Paper.

Cuellar, P., & Rentschler, L. (2023a). Threat, commitment, and brinkmanship in adversarial bargaining. Working Paper.

Dallas County DA touts conviction rate that includes guilty pleas. (2010). Dallas Morning News.

Dujardin, P. (2017). *Hampton prosecutor, challenger spar over conviction rate reports*. Daily Press.

Dyke, A. (2007). Electoral cycles in the administration of criminal justice. *Public Choice, 133*, 417–437.

Fender, J. (2011). DA Chambers offers bonuses for prosecutors who hit conviction targets. Denver Post.

Ferguson-Gilbert, C. (2001). It is not whether you win or lose, it is how you play the game: Is the win-loss scorekeeping mentality doing justice for prosecutors? *California Western Law Review, 38*, 283.

Fischbacher, U. (2007). *z-tree: Zurich toolbox for ready-made economic experiments*. Experimental Economics.

Gordon, S. C., & Huber, G. A. (2002). The effect of electoral competitiveness on incumbent behavior. *American Journal of Political Science, 46*(2), 334–351.

Joy, P. A., & McMunigal, K. C. (2011). Contingent rewards for prosecutors? *Criminal Justice, 26*(3), 55–58.

Leonetti, C. (2012). When the emperor has no clothes III: Personnel policies and conflicts of interest in prosecutors' offices. *Cornell Journal of Law and Public Policy, 22*(1), 53–92.

McCannon, B. C., Asaad, C. T., & Wilson, M. (2016). Financial competence, overconfidence, and trusting investments: Results from an experiment. *Journal of Economics and Finance, 40*, 590–606.

Mueller, D. (2003). *Public choice III*. Cambridge University Press.

Ostrom, B. J., & Hanson, R. A. (1999). *Efficiency, timeliness, and quality: A new perspective from nine state criminal courts*. National Center for State Courts.

Ralston, J., Aimone, J., Rentschler L., & North, C. (2019). False confessions: An experimental laboratory study of the innocence problem. Working Paper.

Rasmusen, E., Raghav, M., & Ramseyer, M. (2009). Convictions versus conviction rates: The prosecutor's choice. *American Law and Economics Review, 11*(1), 47–78.

Reinganum, J. F. (1988). Plea bargaining and prosecutorial discretion. *American Economic Review, 78*(4), 713–728.

Rodricks, D. (2022). Marilyn Mosby claim as an effective prosecutor a hard case to make as Baltimore violence continues. Baltimore Sun.

Schulhofer, S. J. (1984). Is plea bargaining inevitable? *Harvard Law Review, 97*, 1037–1107.

Shamir, J., & Shamir, N. (2012). The role of prosecutor's incentives in creating congestion in criminal courts. *Review of Law and Economics, 8*(3), 579–618.

Silverglate, H. (2011). The revolving door at the department of justice. Forbes.

Sipes, D. A., Oram, M. E., Thornton, M. A., Valluzzi, D. J., & van Duizend, R. (1988). *On trial: The length of civil and criminal trials*. Williamsburg, VA: National Center for State Courts.

Smith, V. (1982). Microeconomic systems as an experimental science. *American Economic Review, 72*(5), 923–955.

Stroud A. M. (2015). Lead prosecutor apologizes for role in sending man to death row. Shreveport Times.

The Cost of Justice. (1998). Los Angeles Times.

Wright, R., & Miller, M. (2002). The screening/bargaining tradeoff. *Stanford Law Review, 55*(1), 29–118.

Zacharias, F. C. (1991). Structuring the ethics of prosecutorial trial practice: Can prosecutors do justice? *Vanderbilt Law Review, 44*(1), 45–114.

Zeiler, K., & Puccetti, E. (2018). Crime, punishment, and legal error: A review of the experimental literature. Boston University School of Law, Law and Economics Research Paper.