False confessions: An experimental laboratory study of the innocence problem

Jason Ralston

Jason A. Aimone

Charles M. North

Baylor University

&

Lucas Rentschler

Utah State University

Abstract: The innocence problem, which occurs when an innocent person is falsely accused or convicted of a crime, is impossible to study with empirical data, because "true" innocence and guilt are unobservable in the "real world." In this study, we replicate the criminal justice system in the laboratory with real salient crimes and subjects acting in the roles of defendants, prosecutors, and jurors in order to study the innocence problem. In a controlled environment, we identify individuals who are falsely or accurately accused of a crime and track them through the plea-bargaining system. This allows us to explore how being falsely accused of a crime affects plea bargaining decisions. We find evidence for a substantial innocence problem, reflected by a high willingness of truly innocent defendants to accept plea bargains. However, they do so at a lower rate than the truly guilty suggesting preferences for truth telling, an irrelevant factor in most economic theory. We also find evidence that individual preferences over uncertainty influence plea decisions, as predicted by economic theory. Overall, we find that loss aversion has a significant positive influence on plea decisions and that the reduced propensity of the truly innocent to accept plea bargains is driven by an interaction between their preferences to avoid lying and their preferences over uncertainty.

JEL codes: C91, K42

Keywords: Experiment, Law, Plea Bargains, Risk Preferences, Loss Aversion

Acknowledgements: We gratefully acknowledge funding from the Charles Koch Foundation. For helpful comments and suggestions, we thank Greg DeAngelo, Bryan McCannon, and seminar and workshop participants from the the 3rd Behavioral and Experimental Public Choice Workshop, the 2019 Public Choice Meetings, the 2019 Regional Mont Pelerin Society Meetings, and from Baylor University.

Introduction

Since the seminal paper of Becker (1968), economists have explored how crime can be deterred through the channels of changing detection/conviction probabilities and alternatively through changing the severity of punishments. Within the first channel of deterrence, the practice of plea bargaining is one way to reduce the direct costs of the criminal justice system, since conviction rates can be increased while also avoiding lengthy and expensive trials. However, there is an indirect and potentially costly side effect of plea bargaining; false confessions from truly innocent people rationally choosing to avoid the risks associated with a noisy trial process. This innocence problem has the potential to follow falsely convicted individuals out of the criminal justice system into the remainder of their careers in the form of reduced job opportunities and experiencing the downside of biases against ex-cons.

Why would an innocent person ever plead guilty to a crime they did not commit? Standard economic theory, where agents are represented as expected value maximizers, predicts that comparing the expected payoff from pleading guilty to that of going to trial should reveal the correct course of action. If the agent were truly rational, their knowledge of their own true guilt/innocence would be irrelevant and would not affect their decision to plead guilty or not. The predictions of behavior become more complex after considering the now common observations that people in general tend to be risk averse (e.g., Eckel and Grossman, 2008; Holt and Laury, 2001; Grossman and Katz, 1983), loss averse (Tversky and Kahneman, 1991; Tom et al., 2007), and ambiguity averse (Fox and Tversky, 1995; Borghans et al., 2009; Crockett et al., 2019). At present, there is little to no evidence on the propensity of people to confess to crimes they did not commit, hindering the potential to measure the extent of the innocence problem. This laboratory study offers the first quantitative insights into the problem.

In assessing the state of our criminal justice system, it is important to know how widespread the Innocence Problem is. In general, the guilty plea is dispositive from a legal standpoint – a person who enters an uncoerced plea of guilty is deemed guilty, regardless of whether the person actually committed the criminal act, provided the judge is satisfied that there exists a factual basis for the plea (Erickson, 1973). Unfortunately, when working with data from actual prosecutions, court records, and appeals, no one can know for certain whether a defendant was guilty or innocent of the crime except for the accused and the true perpetrator. Even with advanced DNA analysis and sophisticated evidence analysis systems, evidence degrades over time; it can be fabricated, planted, and misinterpreted; and sometimes it just does not exist. This presents a substantial challenge to policymakers who seek to design legal institutions and to those seeking to identify the costs of crime. However, this problem can be at least partially overcome via experimental methods. By working in a laboratory setting, we can observe perfectly whether a person accused of a crime actually committed it, thereby allowing us to know whether a person entering a guilty plea is indeed guilty or innocent.

Using this experimental approach, we explore several facets of the innocence problem. First, we ask whether and why innocent people accused of a crime would accept a plea bargain and explore whether economic characteristics of defendants can help us predict who is more predisposed to false pleas. We find evidence that both fits economic intuition and some that highlights the important role that psychological effects have on plea-bargaining behavior. Specifically, we find support for the notion that truly innocent defendants will take plea offers some of the time, but that truly innocent defendants plead guilty significantly less often than truly guilty defendants. As economic theory would predict, preferences over risk, loss, and ambiguity also influence plea-bargaining, though these influences are

strongly dependent on true innocence status. Since in the laboratory we can hold other factors constant (e.g., we control charge, evidence levels, plea offers, trial threats, etc.), the only identifiable difference between the truly innocent and the truly guilty in the sample is whether a plea bargain carries with it a lie or not.[1] Thus, we suggest that the differences we observe in the propensity to accept pleas and the interaction between uncertainty preferences and propensity to accept pleas are connected to aversion to lying (see Abeler et al. 2019 for a recent survey/meta-analysis of the literature).

An experimental approach to studying law and economics is not new. In fact, most of the major actors in the court setting have been studied extensively. Several studies have been conducted concerning judges' decision making and how it is affected by anchoring, framing, hindsight, and racial bias ((Rachlinski, Johnson, Wistrich, & Guthrie, 2009), (Rachlinski, Wistrich, & Guthrie, 2011)). Juries and their collective decision making have been studied experimentally, revealing that jurors place more weight on opinions similar to their own (Bornstein & Greene, 2011) and underweight their peers' judgments regardless of agreement with their own (Minson & Mueller, 2012). Experimental work has also revealed that well-trained attorneys in the prosecution and defense roles both follow reasonably closely to the predictions of an expected utility model when making plea decisions (Bushway, Redlich, & Norris, 2014). However, the same paper also found that prosecutors often systematically deviated from the model in certain circumstances, offering pleas exceeding the expected value of the trial. Further, the institution facilitating plea-bargaining shapes a role in settlement behavior (Marselli, McCannon, & Vannini, 2014). A more comprehensive survey of experimental law literature can be found in (Charness & DeAngelo, 2017). We expand the experimental criminal justice literature by exploring the innocence problem using a novel experimental design.

2 Experimental Design

In order to study defendant decision making, we construct an entire criminal justice system in the laboratory, starting with a potential crime (defendant/victim) interaction, followed by a defendant/prosecutor plea bargain interaction and potentially a jury trial. Trials are resolved by a jury who evaluates evidence of the aforementioned crime. Each of the actors in this criminal justice system environment is a real experimental participant, paid for making decisions in their respective roles. We focus in detail in this paper on the first and second interactions mentioned above between the defendant and victim and between the defendant and prosecutor. However, the decisions of the prosecutor and decisions of the jury are relevant and critical for the decisions that interest us as well. In particular, the following are relevant for this defendant study:

1) Prosecutor Decision Making Profiles: results obtained (using the strategy method) from human participants in the role of prosecutors allow us to determine whether or not a defendant is offered a plea bargain, and what that offer consists of, for all possible accusations and all possible levels of evidence of guilt.

[1] Note, 171 of our 202 participants are willing to steal at some point in the experiment. Since we observe many theft opportunities and only bring charges up in one randomly chosen situation, we can observe people who are willing to be thieves both in situations where they are truly guilty of a theft and in situations where they were truly innocent in the situation they are accused of a crime, even if they had committed a crime in a different situation.

2) Jury Verdict Decisions: results obtained (using the strategy method) from human participants in the role of jurors allow us to determine jury trial outcomes for all possible crime levels and all possible evidence profiles.

Prosecutor and jury participants participated in their own self-contained experimental sessions with their own detailed experimental designs and treatment conditions.[2] Aside from describing the specific places where these previous experiments inform defendant decisions, we refer the interested reader to the prosecutor study (Aimone et al. 2019) and the juror study (Aimone, North, & Rentschler, 2019) for complete descriptions of the corresponding experimental designs. From these studies, we take strategy method decisions from prosecutors that outline what specific human prosecutors would do in each possible situation in which the defendants in the current study could find themselves. Likewise, we take strategy method decisions from jurors that outline what trial verdict specific human jurors would reach in each possible trial situation defendants in the current study could find themselves facing.

2.1 Defendant Study

We utilize a strategy method design in our study. All participants complete each of the three main decision situations we are interested in in this paper: a theft game, a plea-bargain game, and a series of elicitations to reveal individual participants' preferences over uncertainty. We first provide a detailed description of each of these tasks then discuss our procedure for varying the order of the tasks at a session level. Screenshots of instructions and decision-making screens are in appendix A.

2.2 Theft Game

In order for a subject to enter our simulated criminal justice system as a defendant, they must have the opportunity to commit some offense. This opportunity is presented in the theft game and involves a pair of subjects. In each pair, one subject is the potential offender, while the other plays a passive role. Each subject was simultaneously in two pairs, appearing as the potential offender player in one and the passive player in the other. As such, every participant in our sample played in the role of a potential offender. They were not aware that they were also in the passive role (with a different partner) until the end of the experiment.

A theft game proceeds as follows. To begin, the computer randomly divides $10 between the two subjects, where divisions are rounded to the nearest $0.10 and no subject is initially allocated less than $3.00. For example, the computer might have initially allocated the potential offender $6.50 and their partner the remaining $3.50. Another (of the many different possibilities) example is that the computer might have initially allocated the potential offender $3.50 and their partner the remaining $6.50. The computer then informs the potential offender of this division and presents them an opportunity.

The opportunity provides the potential offender four options: leave the initial division unchanged, or take a small, medium, or a large amount of the money allocated to their partner. A small amount is an amount between $0.10 and $1.00. A medium amount is between $1.10 and $2.00. A large amount is

[2] Note that using strategy method decision making procedures, and self-contained seed sessions to generate population data, participants in this experiment and the other experiments referred to here could be paid saliently based upon their decisions in their respective experiments. All studies received IRB review and followed standard economics non-deception protocols.

between $2.10 and $3.00. The specific amount is always in $0.10 intervals. The exact value in each range is determined by a uniform random draw made by the computer. After the potential offender makes their decision, the theft, if chosen, is implemented and used in the final allocation.

	Prosecutorial Evidence	
	If Truly Innocent Defendant	If Truly Guilty Defendant
Strong Evidence of Guilt	5 of every 100 truly innocent	15 of every 100 truly guilty
Medium Evidence of Guilt	10 of every 100 truly innocent	25 of every 100 truly guilty
Weak Evidence of Guilt	15 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

	Defense Evidence	
	If Truly Innocent Defendant	If Truly Guilty Defendant
Strong Evidence of Innocence	15 of every 100 truly innocent	5 of every 100 truly guilty
Medium Evidence of Innocence	25 of every 100 truly innocent	10 of every 100 truly guilty
Weak Evidence of Innocence	40 of every 100 truly innocent	15 of every 100 truly guilty
No Evidence of Innocence	20 of every 100 truly innocent	70 of every 100 truly guilty

Figure 1: Detailed evidence generation process (not shown to participants)

When deciding whether or not to alter the final division, subjects can also weigh how their actions could affect the rest of the experiment. Specifically, subjects are informed that their decisions probabilistically influence the evidence of guilt and/or innocence generated about them. The evidence generating process we use is described in Figure 1, above. A modified simplified version of Figure 1, shown in Figure 2, is presented and explained in the subject's instructions. Note participants are told the probability of getting some level of evidence of innocence or guilt. However, participants are not told the exact probabilities of getting each different level of evidence.

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

	Defense Evidence of *Innocence*	
	If Truly Innocent Defendant	If Truly Guilty Defendant
Some Evidence of Innocence	80 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

Figure 2: Simplified evidence generation process (shown to participants)

The less detailed presentation is used intentionally to preserve ambiguity over the relationship between true innocence/guilt and strength of evidence of innocence/guilt that is present in a real trial. For example, no participant in a real trial knows exactly how much more likely it is for a truly innocent person to have a strong alibi (e.g., CCTV footage suggesting innocence) compared to a medium alibi (e.g., a potentially biased friend's testimony). However, most people may be likely to hold the reasonable belief that innocent people are more likely to obtain a strong alibi than a guilty person and that a weak alibi is always easier to obtain than a strong alibi. Indeed, to make subjects aware of the relationship between true innocence/guilt and evidence of innocence/guilt the instructions explicitly stated 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 truly innocent defendant is always more likely to get a particular level of evidence of innocence compared to a truly guilty defendant." Similar language was used when explaining the evidence of guilt.

The theft game described above is played 10 times in succession. Participants are told that "the computer will choose one of your reports for payment". We programed to the computer to choose one of the 10 rounds chosen quasi-randomly for payment: the round in which the subject first stole was assigned a 0.5 chance of being selected for payment, and the other 9 rounds split the remaining 0.5 chance of selection evenly.[3] Behavior in the round selected for payment determines evidence of guilt and evidence of innocence generated. After the theft game concludes, subjects immediately proceed to the plea-bargaining game. Subjects are not told which of the 10 repetitions of the theft game is chosen for payment by the randomization procedure until the plea-bargaining game begins.

2.3 Plea-Bargaining Game

After reviewing the instructions of the plea-bargain game, participants are told which theft game was chosen by the computer for the plea-bargaining game. They are reminded of: i) the preset division from that round; ii) whether they are truly guilty of taking money; and, if they are guilty iii) how much money they took from their partner. We ask defendants to make decisions in many possible legal situations and tell them that if their prosecutor offers them a plea bargain, we will look at their relevant plea bargain response. Though the computer generates only one set of evidence and one crime severity for a subject, the subject also considers many other sets of evidence and crime severities as a result of the strategy method we employ. That is, subjects make decisions for half of all the different possible evidence and crime level conditions they could have found themselves in. The half of the situations they make decisions over is guaranteed to include all the possible plea offers the prosecutor could have made for their actual level of guilt. The process we use to construct the different scenarios the subjects consider is described next.

First, evidence of innocence is grouped into two bins, low and high with the low bin containing nonexistent and weak evidence of innocence and the high bin containing the medium and strong evidence of innocence. From the random theft round chosen for payment, we know the true evidence

[3] As we wanted some truly guilty individuals in our final data set, and the rate of stealing was relatively low in our seed study, we set the algorithm to choose the first theft with a higher probability. This precaution wound up not being necessary.

of innocence associated with a particular subject and which bin they are truly in. We then choose two levels of evidence of innocence to present to the subject throughout the plea-bargaining game: the true level of evidence of innocence, and another randomly from the remaining bin that the subject was not truly in. For example, if a subject truly has a strong evidence of innocence, then they would, with certainty, consider their true strong evidence of innocence as well as a level of evidence from the low bin, either nonexistent or weak evidence of innocence, randomly chosen. This process keeps the temporal burden low for participants in the lab while always guaranteeing that a participant makes a decision over their true evidence of innocence. This process also gives us rich within-subject data regarding how subjects make decisions in both favorable and non-favorable evidence situations. The subjects make decisions over these two levels of evidence of innocence in conjunction with all of the possible levels of evidence of guilt and severities of crime. On average, participants made 47.4 such decisions. Therefore, every new plea-bargaining screen the subject encountered contained a new set of circumstances for them to consider.

2.3.1 Plea-Bargaining Decision

As described above, subjects make many decisions concerning how to respond to potential plea bargains. On every plea bargain decision screen, subjects face a number of plea bargain decision opportunities. The number of decisions on a screen can be as many as five or as few as one and are situation dependent.

The subjects always have several pieces of information available when making their decisions which describe the defendant's situation. In each situation they are always reminded of their true innocence or guilt. If they have any evidence of innocence, then they are told the strength of their evidence (weak, medium, or strong). Likewise, they are informed in every situation of the evidence of guilt the prosecutor has at their disposal (weak, medium, or strong) and the specifics of the plea deal the prosecutor is offering for the situation. The specifics of the plea deal include the crime the subject would be pleading guilty to (small, medium, or large crime), what monetary punishment the plea deal entails, and the trial threat the prosecutor will pursue if the defendant does not plead guilty. Specifically, the trial threat consists of the level of crime a defendant will be tried for and (if found guilty) the accompanying monetary punishment which is always more than the monetary punishment associated with accepting the offered plea deal. This trial threat is credible – that is, if the subject declines the prosecutor's relevant plea deal, they knew the case will be resolved by that specific trial with certainty.[4]

After observing all relevant information, defendants decide how to proceed. When deciding, subjects have two or three options to select from. When subjects have some evidence of innocence, their three options are to: i)accept the plea deal and plead guilty; ii)decline the deal and proceed to trial and present their evidence of innocence; and iii)decline the deal and proceed to trial *without* presenting their evidence of innocence. The last option is included to maintain similarity to the real criminal justice

[4] Due to the large number of threatened trials within each plea bargain decision environment, we did not have a procedure for eliciting salient beliefs about the probability of being found guilty at trial should a participant reject a plea bargain. As such some participants may view the trial from a rejected plea as a risky environment (if they have a strong probability prior or belief in mind) or may view such a trial as more ambiguous (if they do not have well established probabilities associated with trial outcome likelihoods). We do gather loose non-salient beliefs on a smaller number of trial situations (1 per decision screen, framed as a trial associated with the crime a guilty plea was associated with), but do not analyze those beliefs here. (See Appendix A for more details).

system, which guarantees the right not to testify against oneself. In the case that a subject has no evidence of innocence, the defendant is not provided the option to decline the plea deal and proceed to trial and present their evidence of innocence (since they have no evidence of innocence.)

As with the standard non-laboratory criminal justice system, prosecutors are not forced to offer a defendant a plea bargain. As such, defendants are also asked to make decisions for the situation that their prosecutor does not offer a plea bargain and instead takes their case straight to trial. In this situation, defendants choose to either present their evidence of innocence at trial (if they have evidence) or not. Therefore, for any possible prosecutor decision, the defendant always makes any relevant decisions they have available in response to the prosecutor's decision.

2.3.2 Resolution of Plea-Bargaining Game

When the plea-bargaining game starts, the computer randomly chooses a representative real human prosecutor from a pool of prosecutors who previously completed the prosecutor experiment. This representative prosecutor's reported strategy profile for each possible situation that could arise is used for all participants in the plea-bargaining game for a particular defendant experiment session (e.g. that prosecutor acts as the local district attorney). Since prosecutors previously decided over all the pairings of evidence of guilt and crime severity, the representative prosecutor for the defendants certainly made a decision for the true level of guilt associated with each defendant and true crime severity the defendant stood accused of in this experiment. The prosecutors knew in advance their decisions could be used to affect real human defendants in future sessions.

As previously stated, at the end of the theft game, defendants have a true evidence of guilt, true evidence of innocence, and true severity of crime generated for them and associated with them. The computer then finds the decision of the representative prosecutor that corresponds to the true evidence of guilt and true severity of crime associated with each defendant. This combination of true evidence of guilt, true evidence of innocence, true severity of crime, and selected prosecutor decision represents exactly one of the settings described to the defendant in their strategy method plea-bargain game. Each defendant is rendered a final criminal justice decision from that specific setting.

If a defendant goes to trial either via a rejected plea deal or the representative prosecutor opted to take a case straight to trial, decisions from previous juror experiments are used to render a verdict. The jurors in these experiments were presented with all the combinations of evidence of guilt, evidence of innocence, and severity of crime and were asked to render a verdict of guilty or not guilty for each combination. Jurors made judgments independently of one another. They were told that in the event a defendant goes to trial, three jurors from the pool of all jurors who had completed the juror experiment would be pulled at random to generate a guilty/not guilty ruling. If all three jurors agreed that the defendant was guilty (or not guilty), then the defendant would be ruled guilty (or not guilty) and fined some amount (or not fined). If their judgments did not agree, the trial would be labeled a mistrial, and the process repeated until a set of 3 jurors were found that all agreed. Decisions from these jurors are also used to determine what the punishment from the trial punishment interval would be. More information on the juror experiment can be found in Aimone et al. (2019).

At the conclusion of the plea-bargaining game, the results are displayed to the subject. The screen reminds the subject of the theft decision that was under consideration, what their decision was in that situation, and whether they were truly innocent or truly guilty. The screen also reports whether the

defendant plead guilty or whether the subject was found guilty or not guilty at trial and how much they have been fined (if they were found guilty.) If there was no accusation by the prosecutor, that was reported to the participant.

2.4 Elicitation Tasks

In addition to the theft and plea-bargaining games, subjects complete three different elicitations for their preferences over uncertainty: a risk attitudes elicitation, a loss-aversion elicitation, and an ambiguity aversion task. The risk elicitation task is similar to that found in Eckel & Grossman (2008), the loss aversion task is similar to that found in Fehr & Goette (2007) and Gachter et al. (2010), and the ambiguity aversion task is similar to that found in Borghans et al. (2009). See Figure 3 below for details. We apply definitions from those papers to categorize risk, loss, and ambiguity preferences of our subjects. One of these three tasks is randomly chosen to be paid.

A: $2.80 or $2.80
B: $2.40 or $3.60
C: $2.00 or $4.40
D: $1.60 or $5.20
E: $1.20 or $6.00
F: $0.20 or $7.00

A) Risk Aversion Task: Choose One of the 50/50 Gambles

1) Lose 2 Dimes or Gain 6 Dimes
3) Lose 3 Dimes or Gain 6 Dimes
5) Lose 4 Dimes or Gain 6 Dimes
7) Lose 5 Dimes or Gain 6 Dimes
9) Lose 6 Dimes or Gain 6 Dimes
11) Lose 7 Dimes or Gain 6 Dimes
Even # Decisions) 6 Independent Repetitions of Preceding Odd # Decision

B) Loss Aversion Task: 12 Accept or Reject Decisions. Each Gamble Referred to is a 50/50 Gamble. One chosen to be played out.

Box 1: 10 Balls; 5 Red & 5 Blue
Box 2: 10 Balls; 4-6 Red & 4-6 Blue
Box 3: 10 Balls; 2-8 Red & 2-8 Blue
Box 4: 10 Balls; 0-10 Red & 0-10 Blue
For Each Box: Indicate your color choice. If right earn $2.50.
For Each Box: Indicate min price to sell your bet from [$0, $2.5]

C) Ambiguity Task[5]

Figure 3: Risk Aversion, Loss Aversion, and Ambiguity Aversion Elicitation Tasks

[5] Note. We pay participants for all four urns in this task. While this is not traditional for this task, the recent work by (Bade, 2015) indicates that paying for one of multiple ambiguity choices leads to the potential misrepresentation of ambiguity preferences and ambiguity averse individuals

2.5 Session Order details:

In a given session, each subject participates in the theft game, the plea-bargaining game, the uncertainty preference elicitation procedures, and two repetitions of an additional economic game. This additional

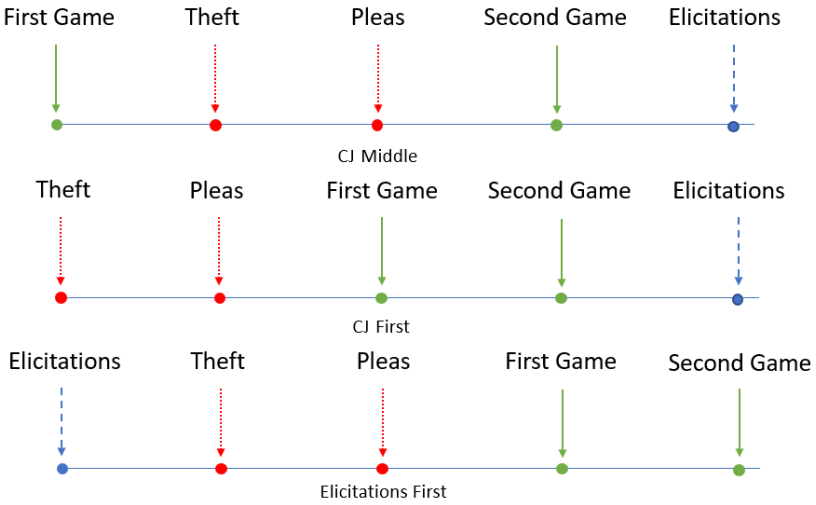

Figure 4: Orderings of games and tasks

game differs across sessions, and is either a gift exchange game, a trust game, or public goods game. These additional economic games are not the subject of this paper and are not discussed further here. Finally, sessions end with an unincentivized survey. To control for order effects, the order of all incentivized tasks are varied between subjects. Figure 4 illustrates the three orders in which tasks are presented to subjects.

2.6 Notes/Discussion about design

There are a few points to note about this design. First, given our focus on the innocence problem, we took the approach of making this design highly framed which is a departure from most studies in experimental economics. Second, one may have concerns that in plea bargaining situations outside of the lab, the worries about external shame of public exposure of confession/conviction would be much greater than in the laboratory setting. This is undeniably true, but a laboratory study serves an important role as a test bed to explore what may be happening outside the laboratory. While the stake size is not as large in the laboratory, innocent participants who take a plea in the lab are still misrepresenting themselves. Even very small stakes lies in the laboratory have been observed to have costs that are very salient to laboratory participants who will pay a price to not be perceived as a liar (see Abeler et al. (2019) for a review.) There is no reason to suspect that lying about taking money from another participant would be any less costly for participants than lying about a die roll, as in those other lying studies. We do believe that plea bargains coming from prosecutors in the "real world" will be more costly than the laboratory setting but likewise the trial threats are also going to be analogously more costly, both in terms of pecuniary (fines) and non-pecuniary (reputation) costs. With such higher stakes on both ends, it would lead one to believe that lying costs, risk, ambiguity, and loss aversion concerns to be even more salient in the real world than they are in the laboratory. As such if we fail to

find effects in the laboratory, that would not preclude there from being effects of risk, loss, and ambiguity preferences on plea bargaining decisions in the outside world. However, any effects we do find would likely be intensified in the outside world and indicate a need to explore that environment in much more detail. As such, we interpret our study as generally being a lower bound.

3. Hypotheses and Results

3.1 Hypotheses

We begin by stating our three main hypotheses regarding our experimental environment. These hypotheses are informed both by the criminal justice system and economic intuition.

Hypothesis 1: *Only truly guilty defendants plead guilty.*

The above hypothesis reflects of the position of the criminal justice system (Erickson, 1973). An economist, however, would predict that a defendant's willingness to accept a plea deal is not informed by a defendant's true innocence or guilt itself, but rather the comparison between the utility from pleading guilty and the expected utility of going to trial.

Hypothesis 2: *True guilt or innocence status of an agent does not influence plea bargaining behavior*

If all other factors are equal between the innocent and the guilty, most standard economic models would lean towards supporting hypothesis 2. However, much economics work suggests that individuals are averse to lying or misrepresenting themselves (see Abeler (2019) for a meta-analysis), leading one to reject hypothesis 2. Instead, aversion to lying leads to the prediction that truly innocent individuals would be less likely to plead guilty as part of a plea bargain. Economic models of preferences over uncertainty would lead to further expectations about differences between individuals. These additional preferences lead to the variations of hypothesis 3.

Hypothesis 3.a: *More risk-averse subjects will plead guilty more frequently than relatively more risk-loving subjects, all else held equal.*

Hypothesis 3.b: *More loss-averse subjects will plead guilty more frequently than less loss-averse subjects, all else held equal.*

Hypothesis 3.c: *Ambiguity-averse subjects will plead guilty more frequently than relatively more ambiguity-loving subjects, all else held equal.*

In our study we have data from 202 participants across 9 sessions conducted at the University of California at Irvine and 4 sessions conducted at Baylor University. 72 participated in the Trust Game version of the experiment, 76 in the Gift Exchange version, and 54 in the public goods version. Of those, 120 participants played in the Econ Game/Theft-Plea/Econ Game/Elicitation order. Fifty-eight participants played in the Theft-Plea/ Econ Game/Econ Game/Elicitation order and the remaining twenty-four participants played in the Elicitation/ Theft-Plea/ Econ Game/Econ Game order. Participants spent approximately two hours in the lab and made an average of $27.14 in the Trust Game, $28 in the Public Goods Game versions and $32.36 in the Gift Exchange version of the task. The study was classified as "Exempt" by our University's Internal Review Board.

3.2 Non-Parametric Analysis

To begin, we illustrate how the percentage of plea-bargains accepted differs by the true guilt or innocence status of a participant. These can be found in Table 1 and Figure 5.

	Never Took a Plea	Took at Least One Plea	Percent of Pleas Taken	Obs
Truly Guilty	0.9%	99.1%	40.8%	111
Truly Innocent	12.1%	87.9%	27.6%	91
MW	***<0.001***		***<0.001***	

Table 1: Plea acceptance by true innocence/guilt

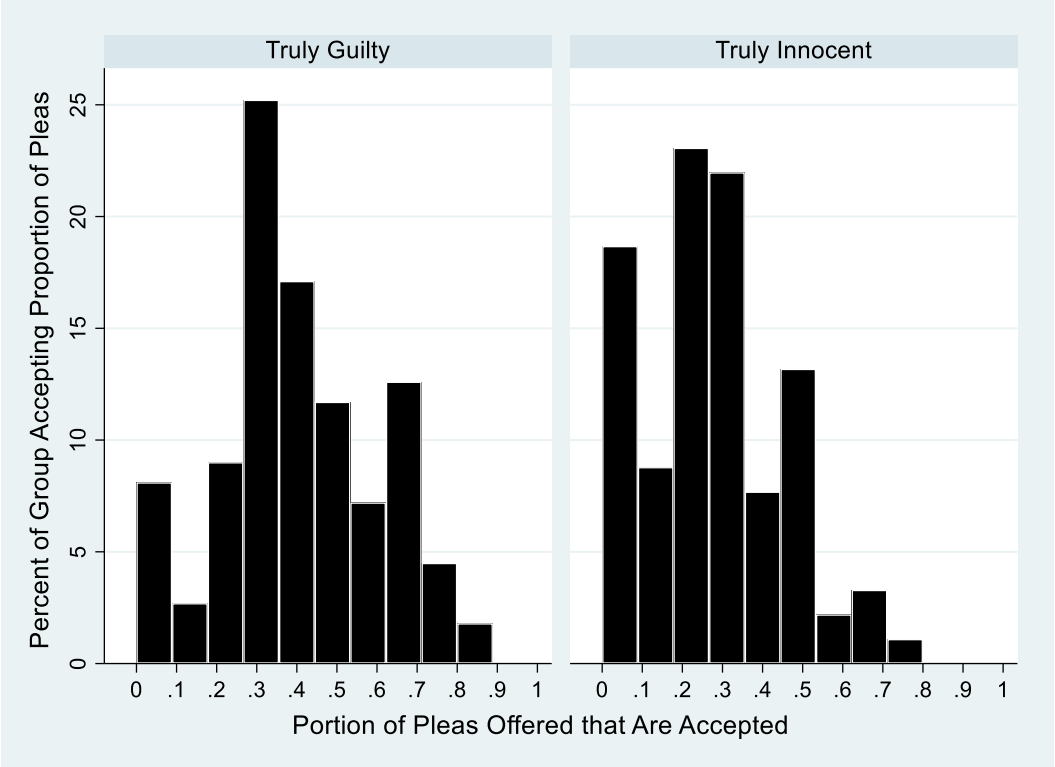

Figure 5: Distributions of plea acceptance over true guilt status

Finding 1: *Truly innocent subjects are willing to accept plea deals.*

Finding 2: *Truly Innocent subjects accepted fewer pleas than truly guilty subjects.*

The first thing to note from Table 1 and Figure 5 is that truly innocent individuals do agree to plead guilty to crimes they did not commit indicating the potential for a substantial innocence problem. Only 12.1% of truly innocent individuals never take a plea bargain. The fraction of those who never take a plea is significantly higher among those who are truly innocent than those are truly guilty (*Mann-*

Whitney U-test, p<0.001). Evident in these graphs as well is the fact that the truly innocent accept plea-bargains at a lower rate (*Mann-Whitney U-test, p<0.001*).[6]

The failure of both hypotheses 1 and 2 may be interpreted as indicating, as we discussed earlier, that some subjects may have preferences not only over money, but also over their own self-image or reputation. When one accepts a plea bargain (i.e., admits guilt for a crime that they did not commit), this is tantamount to a lie, which subjects participants to an internal cost (see (Abeler, Nosenzo, & Raymon, 2019).) The presence of such internal costs is supported by the prior results showing that truly innocent subjects in our experiment exhibit systematically different plea-bargaining behavior than their truly-guilty counterparts. Truly guilty individuals are not lying when they accept a plea bargain unlike those who are truly innocent, nor are they lying by taking their case to court.

Next, we explore the role of subject-level preferences over uncertainty and loss in explaining plea-bargaining behavior. We begin our analysis by examining differences in uncertainty and loss attitudes across truly guilty and truly innocent subjects. Key differences in these measures are reported in Table 2. Columns labeled "MW" indicate p-values of two-tailed Mann-Whitney tests.

	(i)	(ii)	(i) vs (ii)	(iii)	(i) vs (iii)	(iv)	(i) vs (iv)	(iii) vs (iv)
	Baseline	After CJ Exposure	MW	Truly Guilty	MW	Truly Innocent	MW	MW
EG Risk Choice	4.33	3.71	***0.048***	4.00	0.31	3.41	***0.006***	***0.007***
Gachter LA Safe	2.67	2.61	0.935	2.37	0.39	2.87	0.446	***0.031***
Fehr LA Safe	2.04	2.35	0.424	2.13	0.93	2.59	0.143	**0.043**
Total LA Safe	4.71	4.97	0.597	4.50	0.69	5.47	0.147	***0.021***
Amiguity Aversion Level	0.21	0.27	0.651	0.28	0.77	0.27	0.569	0.618
Ambiguity Averse	0.46	0.51	0.627	0.48	0.86	0.55	0.446	0.364
Obsv	24	178		92		86		

Table 2: Differences in subject level characteristics between truly guilty and truly innocent

First note that the criminal justice system itself appears to cause an increase in risk aversion as measured by the EG risk task). Comparing risk choices of participants who participated in the elicitation games prior to any other games in the experiment and those made by participants who participated in the elicitation games after all other games in the experiment, we see similar choices in the loss aversion and ambiguity tasks, but significant ($p<0.05$) increases in risk aversion on the risk aversion task. Comparing truly guilty and truly innocent individuals' decisions, the greatest difference between pre-and post-criminal justice system elicitation task decisions is coming from those who are truly innocent ($p<0.01$, average choice dropping an entire choice level from 4.33 to 3.41 compared to those who took the task at the beginning of the experiment). Comparing the truly guilty and truly innocent to one another, we see that truly innocent subjects in our study choose significantly less risky lotteries than truly guilty subjects ($p<0.01$) and similarly make more loss averse choices as well (for all three ways of measuring loss aversion, $p<0.05$). Given there are differences in uncertainty preferences between our truly innocent and truly guilty groups, in Table 3 we explore plea behavior segmenting the population by uncertainty preferences rather than by innocence status.

[6] If we limit our sample to only those that accept at least one plea bargain the difference between truly innocent and truly guilty defendants remains. The former accept 31.4% of plea offers and the later 41.2%, still significantly different ($p<0.001$).

	Highest Risk Aversion (EG choice 1 or 2)	Medium Risk Aversion (EG choice 3 or 4)	Low/No Risk Aversion (EG choice 5 or 6)	Highest Loss Aversion (4, 5, or 6 Safe GLA)	Medium Loss Aversion (2 or 3 Safe GLA)	Low/No Loss Aversion (0 or 1 Safe GLA)	Ambiguity Averse	Ambiguity Neutral	Ambiguity Seeking
Obs	36	104	62	71	78	53	102	64	36
% That Never Took a Plea	**13.89%**	3.85%	4.84%	**11.27%**	3.85%	1.89%	5.88%	9.38%	0.00%

Table 3:Fraction of subjects who never took a plea over uncertainty and loss preferences

Finding 3.a: *More risk averse subjects are less likely to take a plea deal than less risk averse subjects.*

Table 3 displays the percentage of subjects that never accepted a plea across measurements of uncertainty attitudes. Finding 3.a reports how Table 3 shows that the highest category of risk aversion is also that with the higher propensity to reject all plea deals (MW, $p<0.05$). This implies that those who are the most risk averse are also those who would rather go to an uncertain trial rather than take a certain plea deal. This pushes against the standard expected utility theory prediction expressed in Hypothesis 3.a. We explore this result in more detail shortly.

Finding 3.b: *More loss averse subjects are less likely to take a plea deal than less loss averse subjects.*

Finding 3.c: *Ambiguity-averse subjects plea bargain at rates insignificantly different from ambiguity-neutral and ambiguity-loving subjects.*

Table 3 shows that the most loss averse category of subjects also exhibit the higher propensity to reject all plea deals (MW, $p<0.05$). This is in line with Hypothesis 3.b, that those who are most averse to losses will avoid the trial that exposes them to larger losses. Table 3 fails to find evidence in support of our third hypothesis that ambiguity aversion is related to greater propensity to take pleas.

Next, we explore the heterogenous effects preferences about uncertainty have on willingness to accept a plea. Table 4 displays differences in plea-bargaining behavior between truly innocent and truly guilty subjects across the different levels of evidence of innocence.

		Evidence of Innocence			
	All Levels	**None**	**Weak**	**Medium**	**Strong**
Truly Innocent	27.59%	50.80%	28.49%	24.13%	10.96%
Truly Guilty	40.79%	57.87%	33.99%	31.74%	23.88%
Mann-Whitney	$p<0.001$	$p>0.1$	$p>0.1$	$p<0.1$	$p<0.001$

Table 4: Willingness to accept plea offers by level of evidence of innocence

Table 4 shows that when evidence of innocence increases, plea acceptance rates decrease for both the truly innocent and the truly guilty (2-tailed trend test, $p<0.001$). When there is no evidence of innocence both the truly innocent and truly guilty take pleas more often than not, and when there is any evidence of innocence, taking a plea bargain is the rarer action. While the truly guilty are seen accepting pleas more often than the truly innocent overall (MW, $p<0.001$), these differences are most significant when evidence of innocence is highest (marginally for medium evidence, $p<0.1$, and highly for strong evidence, $p<0.001$).

3.3 Regression Analysis

The non-parametric analysis in the previous section provides initial evidence that there is a significant difference between truly innocent and truly guilty defendants' propensity to take plea bargains. Further, it provides initial evidence that risk and loss aversion are related to plea decisions, but that ambiguity aversion likely is not. However, the preceding nonparametric tables, figures, and analyses do not tell the full story. Participants are making many decisions each, and plea bargain conditions differ not only based upon evidence of innocence but also based upon evidence of guilt, the plea punishment amount, and the threatened punishment at a potential trial. To explore such factors in more detail, we estimate a panel probit model of plea acceptance on individual levels of risk aversion (coded as 7- Eckel Grossman gamble choice), loss aversion (the total number of safe choices in the Fehr & Goette (2007) and Gachter et al. (2010) tasks), ambiguity aversion, and their interactions with true innocence/guilt status of the decision maker. The results of this estimation can be seen in Table 5.

Not shown in Table 5 are several dummies included to control for factors that are likely relevant for plea decisions but are otherwise not the focus of this study. First, these regressions include a variable that indicates the amount of punishment that would be borne by the defendant if they accept the plea bargain and separate dummies for each possible level of punishment that the prosecutor could threaten as the trial charge/punishment. Second, in order to capture how challenging a potential trial would be, we include a dummy that is equal to 1 when the prosecution's evidence of guilt at trial would be stronger than the defendant's evidence of innocence (and zero otherwise.) Third, we include a dummy for situations when the defense and prosecution evidence are tied. Fourth, we include a dummy equal to one when the defense evidence is non-existent. Such situations mirror "real world" settings in which defendants rely upon their 5th Amendment right not to testify at trial, making those cases somewhat different from others; Table 4 shows that pleas are almost twice as likely in this situation compared to other levels of evidence of innocence. Fifth, we include a dummy variable equal to 1 if the data was collected at Baylor and zero if it was collected at the University of California at Irvine. Finally, note that p-values in our regression are empirical p-values calculated from 10000 simulations of random behavior in the experiment. The use of such randomization-based inference has recently been shown to be of critical importance in experimental work (see Paz and West (2019), Young (2019), and Athey and Imbens (2017). We adopt procedures similar to that used Paz and West (2019) for the calculation of our empirical p-values.

	(i)	(ii)
Risk Aversion	-0.012	0.019
	(0.112)	(0.912)
Loss Aversion (# safe choices)	0.004	0.0149**
	(0.765)	(0.987)
Ambiguity Aversion (Level)	-0.024	0.000
	(0.089)	(0.501)
Risk Aversion*Truly Innocent	-	-0.0649***
	-	(0.000)
Loss Aversion*Truly Innocent	-	-0.0318***
	-	(0.001)
Ambiguity Aversion Level*Truly Innocent	-	-0.0633*
	-	(0.042)
Truly Innocent	-0.4574***	-0.064
	(0.000)	(0.219)
Obs	9,834	9,834
I.O.s	202	202

Notes: Randomized p-values are reported in parentheses. Each value is the proportion of coefficients from 10,000 regressions using simulations of randomized plea bargain decisions that are less than the indicated coefficient. Included, but not shown in the table, are dummies for when the prosecutions level of evidence of guilt is greater than the defense evidence of innocennce, tied evidence levels, no evidence of innocence (when defense is forced to plead the 5th at trial), the severity of the plea punishment offered, dummies for trial threat severity, and a dummy for whether the data was collected at Baylor (rather than Univ. of California at Irvine). *** p<0.005 or p>0.995, ** p<0.025 or p>0.975, * p<0.05 or p>0.95

Table 5:Probit estimation of plea acceptance on individual level preferences/characteristics

In the probit regressions in Table 5, we explore how economic preferences over uncertainty factor into plea decisions. Several key results come from the regressions. First, hypothesis 3 (a, b, and c) predicts that, all else equal, risk aversion, loss aversion, and ambiguity aversion will lead defendants to be relatively more likely to accept plea bargains. In our first regression (column (i)), we fail to see support for these hypotheses as there is no significant influence of uncertainty preferences on plea bargain decisions. However, echoing our non-parametric results, we do see in this regression that true innocence status has a significant predictive effect on the propensity to take plea bargains, with the truly innocent accepting pleas at a significantly reduced rate compared to the truly guilty (empirical p<0.001). This puts our experimental data in between the hypotheses that the criminal justice system would make (hypothesis 1 above, that the truly innocent would never take a plea) and the hypotheses that economists might make (hypothesis 2 above, that the expected value of the trial with respect to the plea bargain should determine acceptance regardless of true guilt or innocence).

In column (ii) of Table 5, we add interaction terms for each of the uncertainty preferences; doing so yields non-negative coefficients on all three uncertainty preference measures in the direction expected in hypothesis 3. Further, the main effect of loss aversion has a significant positive influence on plea

acceptance propensity (empirical p = 0.987, two-tailed p<0.05). Turning to the differences between the truly innocent and the truly guilty, we see in Table 5 that the significant negative correlation between true innocence status and plea acceptance probability from column (i) disappears in column (ii) (empirical p>0.1). The general negative correlation from column (i) is fully captured by the three separate interaction effects, which show significant and negative differences between truly innocent and truly guilty defendants in how each measure of uncertainty aversion affects the propensity to accept plea bargains (empirical p<0.001 for risk and loss aversion, marginally significant p<0.10 for ambiguity aversion).

3.4 Discussion of Probit Results

Table 5 suggests that the significantly reduced propensity for innocent individuals to accept plea bargain offers is related to how they respond to the uncertainty involved in the situation. The more averse to uncertainty a truly innocent individual is, the less likely they are to accept a plea bargain and the more likely they are to face a trial. On the surface, this sounds strange. Why would a truly innocent individual who is more averse to uncertainty prefer the uncertain outcome of a trial relative to a comparable truly guilty individual?

These differences between the truly guilty and the truly innocent start to make sense, as we previously discussed, when one considers that truly innocent individuals would be suffering an extra cost of lying when misrepresenting themselves by falsely entering a guilty plea, and thus they would be expected to accept plea bargains less often (as in column (i)). Taking a plea reduces monetary risk but guarantees that the truly innocent individual will incur the added cost of misrepresenting themselves,[7] while a truly guilty person does not suffer this cost. The natural next question is why this cost of misrepresentation seems to vary at the individual level, connected to uncertainty preferences, as opposed to being fully captured in the truly innocent dummy variable. One potential explanation is that confessing to wrongdoing is an uncommon action, and thus individuals do not know for sure how badly they will suffer from lying and misrepresenting themselves. Such unfamiliar misrepresentation/lying thus has an uncertain impact on their own self-image and will have an uncertain impact on how others will see them. Their own internal perceptions and feelings towards this uncertainty appear to influence how they respond in their plea decisions. Thus, to a truly innocent subject, the "certainty" of punishment that comes from accepting a plea deal may not be so certain from their perspective. In essence, they may be comparing two uncertain lotteries.

In their recent meta-analysis of the lying literature, Abeler et al. (2019) pool the individual participant data from 90 lying studies to explore what model best explains lying behavior and aversion to lying in the laboratory. They attribute the behavioral data to being best described by a model where people

[7] To demonstrate the strength and robustness of this misrepresentation effect, we re-ran model (ii) excluding individuals who never stole in any of their 10 opportunities to steal, leaving us with 171 individuals. All of these individuals chose to steal at least once and thus were guilty of something, but only some of them were guilty of the crime they were being accused of committing. Our results largely hold. The general (positive) loss aversion coefficient remains highly significant (empirical p=0.987, two-tailed p<0.05), as do the significant (negative) interactions between loss aversion (empirical p=0.000, two-tailed p<0.001), and ambiguity aversion (empirical p=0.014, two-tailed p<0.05), and true innocence status. The significant negative interaction between risk aversion and true innocence status disappears though (empirical p= 0.279). This may be driven by the significant change in risk preferences that truly innocent people seem to face when presented with the criminal justice system environment that we discussed earlier and illustrated in Table 2.

balance a desire to be honest (self-image) with a desire to appear honest (reputation concerns or other-image). In the typical laboratory experiment, this delicate balance of preferences seems to lead people to lie for profit, but not to the maximum extent they can get away with. Only rarely are populations seen avoiding lying altogether and even more rare are situations where populations lie maximally (Aimone et al., 2019). This behavior over lying is consistent with the behavior of truly innocent individuals in our study. By taking a plea, these individuals are lying, and thus they suffer disutility from both reduced self-image and loss of reputation. These individuals typically are willing to bear these costs in some circumstances but not all, and the willingness to bear these costs, as we see in Table 5, appears to correlate with their preferences over uncertainty. As one may expect, the more averse to uncertainty they are, the less willing they are to expose themselves to the uncertain costs of lying.

Along these lines, there is some evidence in the lying literature that lying aversion is connected to preferences over losses and the probability of a loss (Garbarino et al., 2018; Grolleau et al., 2016). The literature shows that theoretically and empirically, people exhibiting loss aversion tend to lie more to avoid potential losses and the propensity to lie to avoid a loss is increased when the probability of a loss is small as opposed to large. The literature has less to say when it comes to lying behavior's correlation with risk attitudes, though what has been found in previous studies were null results (Kocher et al., 2017). Our study is the first, we believe, to report significant results between risk preferences and lying behavior.

4. Discussion/Conclusion

Empirical studies cannot fully explore the innocence problem due to the inability to perfectly identify those that are truly innocent. Fortunately, laboratory experiments such as ours provide a means to begin exploring the innocence problem with new data from a controlled environment. In the lab, we can observe both true innocence and true guilt status and also hold factors like evidence of innocence, evidence of guilt, types of charges, threats, and plea offers constant and observe how both types of defendants respond under otherwise identical situations. In our study, we bring laboratory data to bear on two initial questions related to the innocence problem.

First, we show the propensity of a standard subject pool to be willing to accept plea bargains when truly innocent or truly guilty. As expected, we find that the truly guilty are willing to accept pleas at fairly high rates. We also find that truly innocent defendants are generally willing to accept pleas at (to us) a surprisingly high rate (though not as high a rate as truly guilty defendants.) In doing so, these truly innocent defendants face the costs of misrepresenting themselves, even to avoid relatively minor losses like those seen in standard laboratory experiments exploring lying aversion. This suggests the extent of the innocence problem stemming from the US criminal justice system may be even higher, when the stakes of trials are even higher.

Second, we explore how economic preferences over uncertainty relate to both false and true guilty pleas. As expected utility theory suggests, we find that aversion to uncertainty has a general influence on plea bargaining decisions. We find that loss aversion appears to have a significant influence on plea acceptance decisions while we do not see a general significant influence of risk or ambiguity aversion; this finding is reasonable, considering that trials and plea bargains are framed and set predominantly in the loss frame. However, we observe that risk, ambiguity, and loss aversion preferences are each

important in explaining the reduced propensity to accept plea bargains seen among the truly innocent compared to the truly guilty. The more averse to uncertainty (either risk, ambiguity, or losses) a truly innocent individual is, the less willing they are to accept a plea bargain. We speculate that this is due to the additional cost of misrepresentation/loss of self-image that comes from the inherent lie in accepting a guilty plea and thus influences the plea decisions of the truly innocent but not the truly guilty.

Future studies may profit from studying this connection between lying aversion, uncertainty preferences, and the innocence problem. Similarly, empirical studies may find value in exploring whether marginalized groups are disproportionately affected by the innocence problem. For instance, indigent or high risk defendants who cannot afford bail or were denied bail may make plea bargain decisions from jail, giving them different risk/loss/ambiguity preferences than comparable populations who have been released on bail. Finally, past research showing that cognitive load adversely affects risk preferences (see Deck and Jahedi, 2015) indicates that plea decision making in these diverse environments would likely reveal different behavior.

References

Abeler, Johannes, Daniele Nosenzo, and Collin Raymond. "Preferences for truth-telling." *Econometrica* 87.4 (2019): 1115-1153.

Athey, Susan, and Guido W. Imbens. "The econometrics of randomized experiments." *Handbook of Economic Field Experiments*. Vol. 1. North-Holland, 2017. 73-140.

Aimone, Jason A., Charles North, and Lucas Rentschler. "Priming the jury by asking for Donations: An empirical and experimental study." *Journal of Economic Behavior & Organization* 160 (2019): 158-167.

Aimone, Jason A., Brittany Ward, and James E. West. *Dishonest Behavior: Sin Big or Go Home*. No. w25746. National Bureau of Economic Research, 2019.

Bade, Sophie. "Randomization devices and the elicitation of ambiguity-averse preferences." *Journal of Economic Theory*159 (2015): 221-235.

Berg, Joyce, John Dickhaut, and Kevin McCabe. "Trust, reciprocity, and social history." *Games and economic behavior* 10.1 (1995): 122-142.

Becker, Gary S. "Crime and punishment: An economic approach." *The economic dimensions of crime*. Palgrave Macmillan, London, 1968. 13-68.

Borghans, Lex, et al. "Gender differences in risk aversion and ambiguity aversion." *Journal of the European Economic Association* 7.2-3 (2009): 649-658.

Bornstein, Brian H., and Edie Greene. "Jury decision making: Implications for and from psychology." *Current Directions in Psychological Science* 20.1 (2011): 63-67.

Bushway, Shawn D., Allison D. Redlich, and Robert J. Norris. "An explicit test of plea bargaining in the “shadow of the trial”." *Criminology* 52.4 (2014): 723-754.

Charness, Gary. "Attribution and reciprocity in an experimental labor market." *Journal of labor Economics* 22.3 (2004): 665-688.

Charness, Gary, and Gregory DeAngelo. "12. Law and economics in the laboratory." *Research Handbook on Behavioral Law and Economics* (2018): 321.

Crockett, Sean, Yehuda Yud Izhakian, and Julian C. Jamison. "Ellsberg's Hidden Paradox." *Julian C., Ellsberg's Hidden Paradox (July 20, 2019)* (2019).

Deck, Cary, and Salar Jahedi. "The effect of cognitive load on economic decision making: A survey and new experiments." *European Economic Review* 78 (2015): 97-119.

Eckel, Catherine C., and Philip J. Grossman. "Men, women and risk aversion: Experimental evidence." *Handbook of experimental economics results* 1 (2008): 1061-1073.

Fehr, Ernst, and Urs Fischbacher. "Why social preferences matter–the impact of non-selfish motives on competition, cooperation and incentives." *The economic journal* 112.478 (2002): C1-C33.

Fehr, Ernst, and Lorenz Goette. "Do workers work more if wages are high? Evidence from a randomized field experiment." *American Economic Review* 97.1 (2007): 298-317.

Fischbacher, Urs, and Simon Gachter. "Social preferences, beliefs, and the dynamics of free riding in public goods experiments." *American economic review* 100.1 (2010): 541-56.

Fox, Craig R., and Amos Tversky. "Ambiguity aversion and comparative ignorance." *The quarterly journal of economics*110.3 (1995): 585-603.

Garbarino, Ellen, Robert Slonim, and Marie Claire Villeval. "Loss aversion and lying behavior." *Journal of Economic Behavior & Organization* 158 (2019): 379-393.

Grolleau, Gilles, Martin G. Kocher, and Angela Sutan. "Cheating and loss aversion: Do people cheat more to avoid a loss?." *Management Science* 62.12 (2016): 3428-3438.

Grossman, Gene M., and Michael L. Katz. "Plea bargaining and social welfare." *The American Economic Review* 73.4 (1983): 749-757.

Guthrie, Chris, Jeffrey J. Rachlinski, and Andrew J. Wistrich. "Blinking on the bench: How judges decide cases." *Cornell L. Rev.* 93 (2007): 1.

Marselli, Riccardo, Bryan C. McCannon, and Marco Vannini. "Bargaining in the shadow of arbitration." *Journal of Economic Behavior & Organization* 117 (2015): 356-368.

Minson, Julia A., and Jennifer S. Mueller. "The cost of collaboration: Why joint decision making exacerbates rejection of outside information." *Psychological Science* 23.3 (2012): 219-224.

Paz, Lourenço S., and James E. West. *Should We Trust Clustered Standard Errors? A Comparison with Randomization-Based Methods*. No. w25926. National Bureau of Economic Research, 2019.

Rachlinski, J., et al. "Does unconscious racial bias affect trial court judges." *Notre Dame Law Review* 84.3 (2009): 1195-1246.

Rachlinski, Jeffrey J., Chris Guthrie, and Andrew J. Wistrich. "Probable cause, probability, and hindsight." *Journal of Empirical Legal Studies* 8 (2011): 72-98.

Tom, Sabrina, Fox, Craig R., Trepel, Christopher, and Russell A. Poldrack. “The neural basis of loss aversion in decision-making under risk.” *Science* 315.5811 (2007): 515-518

Tversky, Amos, and Daniel Kahneman. "Loss aversion in riskless choice: A reference-dependent model." *The Quarterly Journal of Economics* 106.4 (1991): 1039-1061.

Young, Alwyn. "Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results." *The Quarterly Journal of Economics* 134.2 (2018): 557-598.