Exoneree Compensation and Endogenous Plea Bargaining: Theory and Experiment*

Chulyoung Kim and Sang-Hyun Kim

School of Economics, Yonsei University, Seoul 03722, Korea

October 2018

Abstract

We provide a model of endogenous plea bargaining in which a prosecutor has discretion over her choice of plea bargains in response to a level of exoneree compensation mandated by the state. It is shown that an increase of the compensation may invite a sentence-maximizing prosecutor to offer a higher or lower plea bargain discount, depending on parameter values. We brought this model to the lab, finding that (i) when the exoneration process featured high accuracy, a higher level of exoneree compensation induced no significant change in the average plea bargain discounts but still reduced the number of innocent pleas without affecting the number of guilty individuals pleading guilty, and (ii) when the exoneration process was plagued with low accuracy, a higher level of exoneree compensation increased the average plea bargain discounts but had no significant influence on the number of innocent and guilty individuals pleading guilty. These findings suggest that exoneree compensation could be an effective policy tool in reducing innocent pleas and wrongful convictions when combined with accurate exoneration processes, and that a statute for exoneree compensation could be effective even when one cannot expect coordination between the prosecution office and the state legislative.

JEL Classifications: C92, K14, K41, K42

Keywords: exoneree compensation, plea bargain, wrongful conviction, exoneration process

---

* We are grateful to the participants at the seminar organized by Korean Econometric Society for comments. We thank Ino Cho, Miho Hong, Jisu Lee, Myunghwan Lee, and Sangyoon Nam for research assistance. This work was supported by the Yonsei University Future-leading Research Initiative of 2017 (#2017-22-0123). E-mail addresses: chulyoung.kim@gmail.com (Chulyoung Kim) and kimsan46@gmail.com (Sang-Hyung Kim).

1. Introduction

It is unfortunate reality that our criminal courts do err sometimes, convicting innocent individuals and releasing guilty ones.[1] As society puts more weight on the former type of mistakes (i.e., conviction of innocent individuals) in general, many public policy organizations and academic scholars have taken wrongful convictions as a serious threat to justice.[2] Wrongful convictions are particularly troubling if one considers the aftermath of the wrongly convicted even when they are successfully exonerated afterwards. For instance, it is quite difficult for exonerees to find jobs and to have access to necessary services such as health care, education and housing, and they often live with help from family members or charities (Innocent Project).[3] Moreover, they often suffer from post-traumatic stress and mental disorders, which could have long-lasting impacts on their lives (Campbell and Denov, 2004, Grounds, 2004, and Westervelt and Cook, 2008, 2010).

A partial solution to mitigate the problem could be provision of post-exoneration compensations to the exonerees.[4] Such monetary support from states could help exonerees make a smooth transition back to their previous lives with the feelings of security (Campbell and Denov, 2004).[5] However, an important issue in implementing such a public policy is how

---

[1] Gross et al. (2005) report more than 300 exonerations of the wrongly convicted since 1989. Gross (2008) also reports estimates of the false conviction rate for death sentences from 1973 to 1989, ranging from 2.3% to 5%.

[2] Minimizing the number of wrongful convictions is one of the important goals of our criminal justice system, as captured by Blackstone's ratio (Blackstone, 1765).

[3] Innocent Project is a nonprofit organization based out of the Cardozo Law School, which works to exonerate wrongly convicted individuals through DNA testing. One of their clients, named Lewis Jim Fogle, who was exonerated in 2015 after serving 34 years in jail, told "It's harder to make it out here than what I thought, especially if you have no income, no job to keep you occupied." The average time spent wrongfully incarcerated is substantial: Gross et al. (2005) report 10 years and Mandery et al. (2013) report 12.5 years.

[4] Many legal scholars and advocacy groups have discussed this policy measure. See, e.g., Armbrust (2004), Bernhard (1999, 2009), Mandery et al. (2013), and Norris (2012). Also see the publications from the Innocent Project at https://www.innocenceproject.org/.

[5] Despite strong voices for exoneree compensation, many states do not adopt this policy measure. As of 2017 in the U.S., 32 states, the District of Columbia and the federal government have laws that provide exonerees with compensations. Moreover, the statues are not uniform across the states. In terms of compensation amount, for instance, an exoneree in Texas is provided with USD 80,000 per year of incarceration and an annuity set at the same amount, whereas Wisconsin gives only USD 5,000 per year of incarceration with the maximum of USD 25,000. Some states put a time constraint for a request for compensation: for instance, in Mississippi, an exonerated individual can apply for compensation within 3 years "after either the grant of a pardon or the grant of judicial relief and satisfaction of other conditions described in Section 11-44-3(1) (Miss. Code Ann. §11-44-9 (2012))." Other states place substantial restrictions on eligibility: for instance, Nebraska does not provide compensation to those who pleaded guilty or falsely confessed (Neb. Rev. Stat. §29-4603(4) (Supp. 2012)). In

individuals may respond to it. For instance, if the number of wrongful convictions increases in response to a higher exoneree compensation, it could weaken the case for the provision of compensations despite the strong moral ground for it. In this paper, we study the behavioral responses of the key players in litigation processes to the introduction of exoneree compensation. In particular, we study the effect of a higher level of compensation on wrongful convictions through plea bargaining in which a sentence-maximizing prosecutor offers a discounted sentence to defendants. We believe that this is an important topic because plea bargaining processes account for about 95 percent of all convictions secured in U.S. (U.S. Department of Justice).[6]

In Section 2, we provide a theoretical model in which a sentence-maximizing prosecutor interacts with a risk-averse defendant who has private information regarding whether he is guilty or innocent. In our model, a guilty defendant is more likely to be convicted at trial and less likely to be exonerated (conditional on conviction) than an innocent defendant. In such a situation, we find two effects that influence the prosecutor's choice of plea bargain deals, one increasing the plea bargain discount and the other decreasing it.

The first effect, *competitive compensation effect*, operates to *increase* the discount (hence, lower sentence) chosen by the prosecutor. To understand this effect, let us consider the prosecutor's sentence-maximizing problem. Her optimal plea-discount has to balance marginal gain and loss from one unit increase in discount: the marginal gain (which is positive when the prosecutor expects a low likelihood of conviction at trial) is the increase in the number of defendants who accept the plea deals, and the marginal loss is the decrease in sentence from those who accept the deal. Thus, if more defendants go to trial in response to a higher level of exoneree compensation, it reduces the marginal loss because the number of defendants who accept the offer is now smaller, in which case the prosecutor has an incentive to increase her plea discount.

Montana, exonerated individuals are eligible for compensation only when they were exonerated through DNA testing (Mont. Code Ann. §53-1-214(1) (2011)).

[6] The National Association of Criminal Defense Lawyers (NACDL) recently published a report titled, "The Trial Penalty: The Sixth Amendment Right to Trial on the Verge of Extinction and How to Save It," which examines specific cases, data and statistics to explain the decline in the criminal trial and the steady rise in plea deals. Over the last 50 years, defendants chose trial in less than 3 percent of state and federal criminal cases. The remaining 97 percent of cases were resolved through plea deals. Bowers (2008) argues that plea bargaining generates innocent pleas, and Gazal-Ayal (2006) and Bar-Gill and Gazal-Ayal (2006) study the ways to encourage the innocent to refuse plea bargains and to go to trial.

The second effect, *screening effect*, operates to decrease the discount (hence, higher sentence at plea bargaining). Suppose almost all innocent defendants are willing to face trial when the level of exoneree compensation increases. Then, knowing that she is highly likely dealing with guilty defendants, the prosecutor has an incentive to lower discounts because guilty defendants are willing to accept such unfavorable plea deals. We discuss these effects and provide examples in Section 2.

In Sections 3 and 4, we bring this model to the lab and let subjects play this game. In particular, our aim is to investigate whether subjects' responses to exoneree compensation are different between an accurate exoneration regime and an inaccurate one. Our experimental data reveal interesting and distinct patterns between the two regimes. First, prosecutor subjects increased their plea bargain discounts in both regimes but the increase is statistically significant only when the exoneration process is inaccurate. Second, the subjects who played the role of an innocent defendant declined plea bargains and went to trial more often when exoneree compensations were in place, but their behavioral changes are statistically significant only when the exoneration process is accurate. In contrast, the guilty subjects were not responsive to exoneree compensation on average.

Combining these findings, we obtain our main experimental result: when exoneree compensation is introduced in an accurate exoneration regime, it did not significantly change prosecutor behavior but reduced the number of innocent pleas while keeping constant the number of guilty individuals pleading guilty. In contrast, when exoneree compensation is introduced in an inaccurate exoneration regime, it did significantly change prosecutor behavior but did not significantly change defendant behavior. This result suggests that the compensation statutes can be effective in reducing false guilty pleas and wrongful convictions when our exoneration processes feature high accuracy. Moreover, as we find such desirable effects of the statues in a framework with sentence-maximizing prosecutors, our result suggests that the compensation statutes could be effective even when we cannot expect coordination between the prosecution office (in charge of plea bargaining) and the state legislative (in charge of compensation statues).

There are only a few papers that investigate the incentive effects of exoneree compensation. Using a theoretical model, Fon and Schäfer (2007) argue that compensating the wrongfully convicted would induce a mass of individuals to refrain from committing a crime, thereby increasing the level of deterrence. In contrast, Doménech and Puchades (2015) show that exoneree compensation can either deter or encourage crime, and that the social cost is minimized when some exonerees are left with no compensation. Mandery et al. (2013) show that a higher level of compensation leads exonerees to commit a crime less often. In particular, they show that those who are compensated above the threshold USD 500,000 commit offenses at a significantly lower rate than those who are not compensated or compensated less than the threshold.

While these studies focus on the behavior of the potential criminals, Mungan and Klick (2016, MK henceforth) study the effect of exoneree compensation on wrongful convictions through plea bargaining, as we do in the current paper. In particular, using a mechanism design approach, MK show that the introduction of exoneree compensation, coupled with an increase in plea bargain discounts, could be desirable for society because it can reduce the number of innocent pleas without changing the number of guilty individuals pleading guilty. Their finding rests on the difference in exoneration probability between innocent and guilty individuals. When the level of compensation increases, both innocent and guilty individuals have higher incentive to refuse plea bargains and go to trial because expected trial payoffs are now higher due to compensation. Moreover, innocent individuals experience a higher increase in their expected trial payoffs than guilty individuals because exoneration is more likely for innocent individuals. Therefore, when plea bargain discounts are increased as well in this situation, the number of guilty individuals pleading guilty can be kept constant while some innocent individuals still choose to go to trial. Thus, the increase in exoneree compensation, coupled with an increase in plea bargain discounts, could be welfare improving by reducing false guilty pleas and wrongful convictions in criminal courts. In contrast to their mechanism design approach, we endogenize plea bargain decisions by having sentence-maximizing prosecutors to make plea bargain offers to defendants who then decide whether to accept the offers.

After we present our theoretical models and experimental findings in Sections 2 through 4, we conclude in Section 5. All figures and tables can be found in the end. The Appendix contains the experimental instruction and the mathematical proofs.

2. Theoretical Analysis

2.1 Model

This section theoretically examines a model of endogenous plea bargaining in which a sentence-maximizing prosecutor interacts with risk averse defendants who have private information regarding whether they are guilty or innocent. Our model closely follows MK's model with only a couple of exceptions. First, instead of the power function used in their model, we assume the following expected utility of each defendant:

$$U_\eta(X) = \mathrm{E}(X) - \eta V(X)$$

where X is the lottery that the individual chooses, η is the intensity of risk aversion which differs across individual, and $V(X)$ is the variance of X. This simple "mean-variance" expected utility function keeps the essence of MK's model, and at the same time makes the prosecutor's maximization problem tractable and the model more comprehensible.[7] We assume that η is drawn from a uniform distribution, denoted by $H(\eta)$, with a support $[0, \bar{\eta}]$.

Second, and more importantly, the prosecutor independently makes a discount offer to a defendant in our model whereas the discount is determined by the social planner together with the exoneree compensation in MK's model.

Specifically, defendants initially endowed with wealth $w > 0$ get sanction $s > 0$ if sentenced guilty, and no sanction at all otherwise. The prosecutor makes a take-it-or-leave-it offer of a discount δ for the sanction in the way to maximize the expected sentence. Thus, the utility of an individual who pleads guilty is $w - (1-\delta)s$ which is just the expected value of "accepting the deal"—that is, without the disutility of variance—because there is no uncertainty in plea bargaining, i.e., $V(X) = 0$. In this case, the prosecutor's payoff is $(1-\delta)s$.

If the defendant rejects the plea bargain, one of three events happens to him: he may be acquitted, convicted and later exonerated, or convicted and never exonerated. The probabilities of these outcomes taking place depend on whether he is innocent or guilty. As in MK's model, α_1 denotes the probability that an innocent individual is convicted, and $(1-\alpha_2)$ denotes that of a guilty one. Here, α_1 and α_2, respectively, denote the probability of wrongful conviction (type I error) and the probability of false acquittal (type II error) at trial. Exoneration (conditional on conviction) happens with probability ρ_1 if the individual is innocent and $(1-\rho_2)$ if he is guilty. If the defendant is exonerated, the state provides an exoneree compensation

[7] It is well known that this mean-variance expected utility function is equivalent to the constant absolute risk aversion (CARA) utility function if X follows a normal distribution.

of $\psi \leq s$. Therefore, an innocent defendant's expected utility when refusing plea bargains is given by:

$$\begin{aligned}\mathrm{E}(X_{trial}|\text{innocent}) - \eta V(X_{trial}|\text{innocent}) \\ = (1-\alpha_1)w + \alpha_1\rho_1(w+\psi-s) + \alpha_1(1-\rho_1)(w-s) - \eta\sigma_I(\psi)\end{aligned}$$

where $\sigma_I(\psi)$ denotes $V(X_{trial}|\text{innocent})$ as a function of the compensation. On the other hand, a guilty defendant's expected utility is:

$$\begin{aligned}\mathrm{E}(X_{trial}|\text{guilty}) - \eta V(X_{trial}|\text{guilty}) \\ = \alpha_2 w + (1-\alpha_2)(1-\rho_2)(w+\psi-s) + (1-\alpha_2)\rho_2(w-s) - \eta\sigma_G(\psi)\end{aligned}$$

where $\sigma_G(\psi)$, similarly, denotes $V(X_{trial}|\text{guilty})$ as a function of the compensation. Assuming that the probability of a defendant being innocent is q, the prosecutor's expected payoff from trial is $(q\alpha_1 + (1-q)(1-\alpha_2))s$.

Next, we define the social cost, which is the error cost from the litigation process, as a function of ψ, which would depend on the probabilities of wrongful conviction and false acquittal and also on the weights put on the two outcomes. Thus, the social cost that we consider is as follows:

$$\begin{aligned}SC \equiv q[P(\text{trial}|\text{innocent})\alpha_1 + (1-P(\text{trial}|\text{innocent}))(1-\delta)] \\ + \tau(1-q)[P(\text{trial}|\text{guilty})\alpha_2 + (1-P(\text{trial}|\text{guilty}))\delta]\end{aligned}$$

where τ is the relative importance of false acquittals. This expression reads: with probability q the defendant is innocent, with probability $P(\text{trial}|\text{innocent})$ he goes to trial, and with probability α_1 he is sentenced guilty. With probability $1-P(\text{trial}|\text{innocent})$ the defendant accepts the plea bargain, thus $(1-\delta)$ of the sanction is imposed. The rest of the expression reads similarly.

For the parameter values, we assume the following:

$$1-\alpha_1 > \alpha_2$$
$$\rho_1 > 1-\rho_2$$

The first assumption implies that the guilty are more likely to be sentenced guilty than the innocent. The second means that the probability of exoneration is higher for the innocent than for the guilty.

2.2 Analysis

If there exists an innocent defendant who is indifferent between accepting the plea bargain and going to trial, it is determined by the following equation:

$$(1-\alpha_1)w + \alpha_1\rho_1(w+\psi-s) + \alpha_1(1-\rho_1)(w-s) - \eta\sigma_I(\psi) = w-(1-\delta)s$$

Since increasing η reduces the trial payoff, a more (resp., less) risk averse defendant will strictly prefer to accept the plea bargain (resp., go to trial). Thus, finding the threshold type η^I from the above equation, we can define the proportion of innocent defendants going to trial as $H(\eta^I)$ and that accepting the plea bargain as $1-H(\eta^I)$. We can find the threshold type as:

$$\eta^I = \begin{cases} 0, & \tilde{\eta}^I < 0 \\ \tilde{\eta}^I, & \tilde{\eta}^I \in [0,\bar{\eta}] \\ \bar{\eta}, & \tilde{\eta}^I > \bar{\eta} \end{cases}$$

where $\tilde{\eta}^I = [(1-\delta-\alpha_1)s + \alpha_1\rho_1\psi]/\sigma_I(\psi)$. The threshold type of the guilty, η^G, is determined similarly:

$$\eta^G = \begin{cases} 0, & \tilde{\eta}^G < 0 \\ \tilde{\eta}^G, & \tilde{\eta}^G \in [0,\bar{\eta}] \\ \bar{\eta}, & \tilde{\eta}^G > \bar{\eta} \end{cases}$$

where $\tilde{\eta}^G = [(\alpha_2-\delta)s + (1-\alpha_2)(1-\rho_2)\psi]/\sigma_G(\psi)$. Using this threshold type, we can define the proportion of guilty defendants going to trial as $H(\eta^G)$ and that accepting the plea bargain as $1-H(\eta^G)$.

The prosecutor's objective is to maximize the expected sentence, which is formally written as:

$$\max_{\delta} q\left[H(\eta^I)\alpha_1 + \left(1 - H(\eta^I)\right)(1-\delta)\right] + (1-q)\left[H(\eta^G)(1-\alpha_2) + \left(1 - H(\eta^G)\right)(1-\delta)\right]$$

Recall that H is a uniform distribution. Since η^I and η^G are linear in δ, this objective function is concave as long as at least one of η^I and η^G is in $(0, \bar{\eta})$. Also note that $\delta = 1$ cannot be optimal because if so, everybody will accept the offer, i.e., $H(\eta^I) = H(\eta^G) = 0$, and the expected payoff will become zero. Thus, there exists an optimal discount $\delta^* \in [0,1)$.

Will the positive effects of exoneree compensation prevail in our model as in MK's? The answer will crucially depend on how the prosecutor responds to an increase in the exoneree compensation. We identify two conflicting effects of exoneration: *competitive compensation effect* and *screening effect*. First, the prosecutor may increase δ in response to an increase in ψ to balance the marginal and the infra-marginal effects. To be concrete, let us suppose for a moment that most of individuals are innocent ($q \approx 1$). Then, the first derivative of the prosecutor's objective function would approximately be:

$$\underbrace{\frac{h(\eta^I)(1-\delta-\alpha_1)s}{\sigma_I(\psi)}}_{\text{marginal}} - \underbrace{\left(1 - H(\eta^I)\right)}_{\text{infra-marginal}}$$

where $h(\eta) = 1/\bar{\eta}$ is the density function of η. The first term captures the benefit of increasing δ, namely that more defendants get some, albeit discounted, sanction with certainty. Here, $\alpha_1 s$ is the opportunity cost. The second term is the cost of increasing δ, which is a decrease in the size of sanction imposed on those who accept the plea bargain. The optimal discount level balances these two effects. The prosecutor has a similar incentive for the guilty, so the second half of the objective function can be understood in an analogous way.

If an increase in ψ induces more defendants to go to trial, the infra-marginal effect becomes smaller (i.e., $H(\eta^I)$ in the above expression becomes larger), which would operate to increase δ. In plain words, as the size of the exoneree compensation, thus the defendant's trial payoff, increases, the prosecutor may want to "competitively" increase the discount to keep defendants from going to trial. To formalize this intuition, we introduce the following assumption.

A1. $\sigma_I'(\psi), \sigma_G'(\psi) < 0$

This assumption means that the variances of trial payoffs decrease as the size of exoneree compensation increases. In this case, the "going to trial" option becomes unequivocally more attractive. To formally state the intuition, define $\delta^*(\psi)$ as the optimal discount as a function of ψ.

Proposition 1. Suppose that given the parameter values the prosecutor's problem has an interior solution and that $\delta^*(\psi)$ is differentiable. Then, under A1, the optimal discount δ^* increases in response to the increase of ψ, i.e., $\frac{\partial \delta^*(\psi)}{\partial \psi} > 0$.

Proof. See the Appendix.

The proposition focuses on the situation where $\delta^*(\psi)$ is strictly positive and differentiable. In fact, $\delta^*(\psi)$ is differentiable everywhere but at the point where η^I and/or η^G hit $\bar{\eta}$. In other words, we are assuming that the most risk-averse individuals do not change their decisions in response to the change of ψ.

A1 is a sufficient condition for the infra-marginal effect to get smaller (i.e., $H(\eta^I)$ and/or $H(\eta^G)$ get larger). Because now fewer individuals accept the plea bargain, the cost of increasing the discount (i.e., a decrease in the size of sanction imposed on those who accept the plea bargain) is smaller. Thus, the prosecutor will offer a better plea bargain so as to balance the marginal and the infra-marginal effects. It can be shown that if ψ is sufficiently smaller than s, then A1 is satisfied. If ψ is very close to s, however, the variances may actually increase in ψ depending on the parameter values, and the prosecutor may respond differently.

The second effect of exoneration is termed *screening effect*. If the compensation is sufficiently large, most, if not all, innocent defendants will go to trial, and will not accept the plea bargain unless the discount is very large. In such a case, the prosecutor may want to focus on the guilty defendants who are attracted less to the trial option. Because the willingness of the guilty to accept plea bargains is higher than that of the innocent, the prosecutor may want to make a worse offer (i.e., smaller discount) to the defendant. Therefore, as the exoneree compensation increases, the optimal discount may decrease at some point.

A2. $\eta^I, \eta^G \in (0, \bar{\eta})$ for $\psi = 0$ and $\delta = \delta^*(0)$.

A2 means that when the amount of compensation is negligible, some innocent (and very risk-averse) individuals accept the bargain. This assumption rules out some uninteresting cases where all innocent individuals go to trial no matter what. Under this assumption, and if the probability of exoneration ρ_1 is sufficiently large, then there exists the amount of exoneree compensation which makes the most risk-averse individual indifferent between accepting the bargain and going to trial. Let $\hat{\psi}$ be the compensation such that given $\hat{\psi}$ and $\delta^*(\hat{\psi})$, the most risk-averse innocent individual is indifferent between the two options.

A3. ρ_1 is large enough for $\hat{\psi}(< s)$ to exist.

With these assumptions, we can show that the screening effect of exoneration exists. That is, an increase in the exoneree compensation attracts even the most-risk averse innocent defendants to trial, so the prosecutor lowers the discount in response, focusing on the guilty defendants.

Proposition 2. Suppose that $\delta^*(\psi) > 0$ for some ψ. Then, under A2 and A3, there exist ψ' and ψ'' such that $\psi' > \psi''$ and $\delta^*(\psi') < \delta^*(\psi'')$.

Proof. See the Appendix.

To appreciate the discussion so far in more concrete manner, let us consider the following example. For Figure 1, we assume $\alpha_1 = \alpha_2 = 0.5$, $s = 1$, $q = 1/4$, and $H(\eta)$ is a uniform distribution on $[0,2]$.

[Figure 1]

As Proposition 1 states, the equilibrium discount level first increases as the compensation increases. When the compensation level reaches some point, it drops suddenly, as Proposition 2 states, because all innocent defendants would reject the plea bargain unless the discount is very high, in which case the prosecutor would rather focus on the guilty ones who are willing to take a lower discount. Note that the point at which all innocent defendants start to reject the offer depends on the level of ρ_1 and ρ_2. Specifically, the higher ρ_1 is, the smaller level of compensation at which all innocent ones go to trial is. Also notice that the equilibrium level of discount is higher when ρ_1 and ρ_2 are smaller. That is mainly because a larger number of guilty defendants is attracted to the trial when the exoneration process is inaccurate, in which case the cost (i.e., the infra-marginal effect) of increasing the discount is smaller.[8]

Thus far, we have analyzed the incentive of the prosecutor to increase or decrease the discount offer in plea bargaining. Now, let us examine whether the prosecutor's behavior is in accord with the social planner's objective. Note that the domain of maximization is compact and that the objective function is a linear combination of continuous functions, which guarantees the existence of the optimum. Define the socially optimal levels and the minmax solutions of δ and ψ, respectively, as follows:

$$(\delta^S, \psi^S) \equiv \underset{(\delta,\psi)}{\operatorname{argmin}} SC(\delta, \psi)$$

$$\psi^M \equiv \underset{\psi}{\operatorname{argmin}} SC\,(\delta^M(\psi), \psi)$$

where $\delta^M(\psi) \equiv \underset{\delta}{\operatorname{argmax}}\ SC(\delta, \psi)$

In words, (δ^S, ψ^S) minimizes the social cost, δ^M maximizes it given ψ, and ψ^M minimizes it taking the maximization into account. Similarly, the equilibrium levels of δ and ψ are defined as follows:

$$\psi^* \equiv \underset{\psi}{\operatorname{argmin}} SC\,(\delta^*(\psi), \psi)$$

Then, we can show that (δ^*, ψ^*) has the following relationships with (δ^S, ψ^S) and (δ^M, ψ^M):

Proposition 3. Suppose that (δ^S, ψ^S), (δ^M, ψ^M), and (δ^*, ψ^*) are continuous in q in the

---

8 Indeed, our experimental data show that guilty subjects engaged in risk-taking behavior by choosing uncertain trial outcomes over certain plea bargains more often when exoneration processes were inaccurate (i.e., when ρ_1 and ρ_2 are small). In such a situation, prosecutor subjects offered a greater plea bargain discount as the theory predicts. We further discuss this point in the next section.

neighborhoods of zero and one. Then, (δ^*, ψ^*) converge to (δ^S, ψ^S) as q decreases to zero, and to (δ^M, ψ^M) as q increases to one.

Proof. See the Appendix.

The above proposition implies that the equilibrium in general diverges from the social optimum, which is not too surprising given that the objective of the prosecutor is not in line with the social planner. If, however, the proportion of innocent defendants is sufficiently small, the equilibrium outcome will be close to the socially desirable outcome. A more practical question is whether the social cost would decrease in exoneree compensation. The following example shows that it may not be the case. Based on our analysis of the model, the social cost can be written as follows:

$$SC \equiv q\left[H(\eta^I)\alpha_1 + \left(1 - H(\eta^I)\right)(1-\delta)\right] + \tau(1-q)\left[H(\eta^G)\alpha_2 + \left(1 - H(\eta^G)\right)\delta\right]$$

[Figure 2]

For Figure 2, we use most parameter values same as for Figure 1, and assume $\tau = 1$, that is, a false acquittal is as important as a wrongful conviction. When the exoneration process is accurate ($\rho_1 = \rho_2 = 0.9$), and the compensation is not too large, the social cost decreases as the compensation increases. However, if ψ is already large enough, an increase of it only increases the social cost, because the equilibrium discount level is too low for the innocent to accept, and an increase of ψ makes only the guilty better off. When the exoneration process is inaccurate ($\rho_1 = \rho_2 = 0.7$), the social cost increases in most range of ψ. This implies that although the equilibrium δ increases in response to an increase in ψ as shown in Figure 1, it is not enough to keep guilty defendants from going to trial.

3. Experiment

We conducted our experiment at the laboratory managed by the Center for Research in Experimental and Theoretical Economics (CREATE) at Yonsei University, South Korea. Our experiments were computerized by the O-tree software (Chen, Schonger and Wickens, 2016). We recruited 110 undergraduate and graduate students from our subject pool and each subject participated in one treatment (between-subject design).

Each experimental session consisted of 12 rounds and proceeded as follows. Every subject was labeled as either P1 or P2 (subjects were not informed of their labels). In each round, a P1 subject was randomly matched to a P2 subject, and their roles were assigned as shown in Table 1, where P stands for prosecutor, G for guilty defendant, I for innocent defendant, and ψ for exoneree compensation.[9] The defendant subject knew whether he is innocent or guilty, but the prosecutor subject was only informed that the defendant is equally likely to be innocent or guilty (i.e., $q = 1/2$). For instance, in Round 8, P1 subjects played the role of guilty defendant, P2 subjects played the role of prosecutor, and the amount of exoneree compensation was 0 experimental coins.

[Table 1]

The first four rounds were practice rounds, designed for providing subjects with an opportunity to learn, and therefore we do not use them in our data analysis. We implemented such a fixed sequence of executed rounds to collect balanced data by having all subjects experience all possible moves and to avoid double counting issues.

In the beginning of the experiment, subjects were given 600 experimental coins in their virtual accounts, which were used for the entire experiment. It was common knowledge that in each round: 50 experimental coins are deducted from the defendant's account if the defendant is convicted (i.e., $s = 50$); an exoneree compensation is paid to the defendant if the defendant is convicted but exonerated afterwards, where the amount of exoneree compensation is common knowledge and either 0 or 40 experimental coins (i.e., $\psi = 0$ or 40).

[Figure 3]

[9] Subjects were informed that in each round two subjects are randomly matched, their roles (prosecutor or defendant) are randomly assigned with equal probabilities, and a defendant is equally likely to be guilty or innocent. According to Table 1, each subject played a prosecutor for 6 times, a guilty defendant for 3 times, and an innocent defendant for 3 times. We checked our data for any behavioral difference between P1 and P2 subjects and found no significant difference.

Each round consisted of three stages. Figure 3, distributed to subjects to enhance their understanding of our experiment, shows the sequence of stages in each round. In the first stage (plea bargaining stage), without knowing whether the defendant was guilty or innocent, the prosecutor subject decided a plea bargaining discount, δ, which was a number between 0 and 50. If the defendant accepted this offer, the round ended with the prosecutor obtaining $50-\delta$ experimental coins and the defendant losing $50-\delta$ experimental coins. If the defendant rejected the offer, the round continued to the next stage.

In the second stage (trial stage), 10 experimental coins were deducted from both subjects' accounts as litigation costs, and the server computer randomly decided whether the defendant was convicted or acquitted. The innocent defendant faced a higher chance of getting acquitted: the innocent defendant was acquitted with probability 60% while the acquittal probability for the guilty defendant was 40% (i.e., $\alpha_1=\alpha_2=0.4$). When the defendant was acquitted, the round ended without any further change in subjects' accounts. In contrast, when the defendant was convicted, the prosecutor obtained 50 experimental coins, the defendant lost 50 experimental coins, and the round continued to the next stage.

The third stage (exoneration stage) ensues only when the defendant was convicted in the previous stage. In this stage, the server computer randomly decided whether the convicted defendant was to be exonerated or not. In particular, the innocent, convicted defendant was exonerated with probability K while the exoneration probability for the guilty, convicted defendant was $1-K$. When the defendant was exonerated, the defendant obtained the exoneree compensation that was either 0 or 40 experimental coins, and the round ended. Otherwise, the round ended without any further change in subjects' accounts.

We introduced two treatments at this exoneration stage in our experiment. In the first treatment, Accurate Exoneration Process (AEP), we had $K=0.9$ (i.e., $\rho_1=\rho_2=0.9$). That is, in AEP, the innocent, convicted defendant was exonerated with probability 0.9 while the exoneration probability for the guilty, convicted defendant was 0.1. In the second treatment, Inaccurate Exoneration Process (IEP), we had $K=0.6$ (i.e., $\rho_1=\rho_2=0.6$), which represents a relatively inaccurate exoneration stage compared to AEP. In total, 56 subjects participated in the AEP treatment and 54 subjects in the IEP treatment.

After the experiment ended, each experimental coin in a subject's virtual account was converted to KRW 15, and payed to the subject in an envelope. Each session took about 50 minutes and the average payment was about KRW 8,500 (approximately USD 8).

4. Results

In this section we present our main experimental findings. First, we begin with the average plea bargain discounts offered by prosecutors. Figure 4 shows the average plea bargain discount offered by prosecutor subjects. When exoneree compensation is introduced in the AEP treatment, we find a slight increase in the average plea bargain discount: prosecutor subjects increased their discount offers from 20.4 to 21.8. This increase is, however, not statistically significant (t-test p-value = 0.2647). In contrast, we do find a significant increase in the average plea bargain discount in the IEP treatment: the discount increased from 19.2 to 23.3, which is statistically significant (t-test p-value = 0.0074). Thus, we obtain our first experimental finding:

Result 1 (Prosecutor Behavior). When exoneree compensation is introduced, the average plea bargain discount increases in both treatments. However, the increase is statistically significant only in IEP.

[Figure 4]

Prosecutor subjects increased their discount offers when the exoneration process featured low accuracy, exhibiting competitive compensation behavior in response to a higher exoneree compensation. In contrast, the average plea bargain discount also increases in the case of highly accurate exoneration processes, but its effect is marginal and not statistically significant.

[Table 2]

Table 2 provides the results from a formal regression analysis. In the first two columns, we regress the plea bargain discount on the dummy variable taking a value of 1 if the level of exoneree compensation is 40 and zero otherwise, without and with control variables, respectively, where control variables include age, gender, major, and religion. As expected, the coefficient on the compensation dummy is not statistically significant in AEP. The next two columns show the regression results for the IEP treatment. They show that a higher level of exoneree compensation increases plea bargain discounts by more than 4 experimental coins on average, and this effect is highly significant. The coefficient on compensation is quite stable regardless of the inclusion of controls. Thus, using a formal analysis, we can again confirm that a higher level of exoneree compensation does not influence plea bargain discounts in AEP whereas its effect is highly significant in IEP. This might be because our subjects faced a trade-off between two countervailing effects, competitive compensation and screening, in AEP treatment, while the competitive compensation effect clearly dominated screening effect in IEP. Another possible reason why prosecutors decided to offer a greater discount in IEP than in AEP is that more defendants went to trial in IEP, which, as explained in the previous section, reduces the marginal cost (i.e., the infra-marginal effect) of increasing the discount. Below, we investigate defendants' behavior to see whether this is indeed the case.

[Figures 5 and 6]

Figures 5 and 6 show the average accept rates of the innocent and guilty defendants, respectively, in the plea bargaining stage. Figure 5 shows that innocent defendants' responses to a higher exoneree compensation are quite different across the treatments. In AEP, the average accept rate was 17.9% when there was no compensation for exonerated defendants, but it dropped to 3.6% when compensation was introduced. This decrease is statistically significant (t-test p-value = 0.0149). Thus, with exoneree compensation in place, almost all innocent defendants in the AEP treatment rejected the plea bargain and chose to proceed to trial. In contrast, in IEP, we found only small decrease in the average accept rate of the innocent defendants: the average accept rate went down from 18.5% to 14.8%, which is not a statistically significant change (t-test p-value = 0.6096).

Result 2 (Innocent Behavior). When exoneree compensation is introduced, the average accept rate of the innocent defendants decreases in both treatments. However, this decrease is statistically significant only in AEP.

As Figure 6 shows, the guilty defendants were not responsive to a higher exoneree compensation. Their average accept rates were stable around 70% regardless of the magnitude of compensation in AEP (t-test p-value = 0.5383) and the corresponding rates were around 55% in IEP (t-test p-value = 0.8485).

Result 3 (Guilty Behavior). When exoneree compensation is introduced, the change in the average accept rate of guilty defendants is not statistically significant in both treatments.

Understandably, many innocent individuals chose to proceed to trial only when the exoneration process was accurate. Results 2 and 3 together suggest that screening effect—more innocent defendants go to trial than guilty ones do in response to an increase in exoneree compensation— might be salient in AEP, but not in IEP.

[Table 3]

Table 3 provides results from a formal regression analysis. In the first column, restricting our sample to the AEP treatment, we regressed the defendant's decision to accept on the following three variables: *Compensation* is a binary variable which takes a value of 1 if the level of compensation is 40, and zero otherwise. *Discount* is the level of plea bargain discount offered by the prosecutor subject, and *Innocent* is a dummy variable indicating whether the defendant is innocent.

The estimated coefficients on these independent variables are highly significant at 1% level. As expected, a higher level of compensation leads defendants to accept plea bargain deals less often and proceed to trial; a higher level of plea bargain discount increases the defendant's accept rate; and innocent defendants are less likely to plead guilty. The second column shows that our estimated coefficients are quite robust to the inclusion of controls. The two columns in the middle contain regression results for the IEP treatment. When controls are not included in the regression, *Discount* and *Innocent* are highly significant but the effect of *Compensation* is only marginally significant (p-value = 0.073). With controls, only *Discount* and *Innocent* are significant. Thus, our formal regression analysis suggests that increasing the level of exoneree compensation could be an effect policy tool when our exoneration process features high accuracy, but might not produce a noticeable change otherwise.

In the last two columns we report the regression results for the entire sample. *Compensation*, *Discount*, and *Innocent* are highly significant, and the magnitudes and signs of the coefficients are in line with those from subsample regressions. *IEP* is a dummy variable for the IEP treatment, and its coefficient shows that defendant subjects rejected plea bargains more often when exoneration processes are plagued with low accuracy. Although this finding may seem somewhat puzzling, running separate regressions for guilty and innocent individuals provides us with interesting behavioral patterns.

[Table 4]

Table 4 shows regression results for guilty and innocent individuals, respectively, and it shows that it is guilty individuals who drive the above finding on the IEP dummy variable: the coefficients on *IEP* are negative and highly statistically significant for the guilty sample whereas they are insignificant for the innocent sample. These results suggest that guilty individuals exhibit risk-taking behavior in choosing the uncertain trial outcome over the certain plea bargain deal more often when the accuracy decreases.

Thus far we have presented our experimental findings separately for prosecutor and defendant behavior in Results 1 to 3. Combining them, we obtain the following result.

Result 4 (Effectiveness of Exoneree Compensation). When exoneree compensation is introduced in AEP, it did not significantly change prosecutor behavior but reduced the number of innocent pleas while keeping constant the number of guilty individuals pleading guilty. In contrast, when exoneree compensation is introduced in IEP, it did significantly change prosecutor behavior but did not significantly change defendant behavior.

Result 4 suggests that the post-release compensation could be effective in reducing false guilty pleas and wrongful convictions if the exoneration process features high accuracy: a higher level of exoneree compensation induces many innocent defendants, who might have agreed to plea bargain, to reject the plea bargain and proceed to trial without changing the number of guilty defendants pleading guilty. Thus, increasing exoneree compensation, society can reduce the number of wrongful convictions by sending more innocent defendants to trial while keeping the number of false acquittals constant, which improves welfare. This finding suggests that the accuracy of exoneration processes is an important element to be considered in the draft of compensation statues. In light of MK's result, our experimental findings provide a stronger argument for exoneree compensation: these statues could be welfare improving even without the coordination between the prosecution office in charge of plea bargains and the state legislative in charge of drafting statues.

[Table 5]

Table 5 shows breakdowns of our experimental data in terms of final outcomes for defendant subjects with data for innocent subjects in (a) and data for guilty in (b). In the AEP treatment, when compensations are paid to exonerees, the number of innocent pleas went down from 10 to 2 and the number of acquitted individuals increased from 28 to 35. In addition, almost all innocent subjects were exonerated after convicted. Thus, the total number of wrongful convictions, which consists of the number of innocent pleas and convictions without exoneration, went down from 12 to 3. In contrast, for the IEP treatment, the total number of wrongful convictions was constant at 14.

In Table 5(b), final outcomes for guilty subjects in AEP are quite similar regardless of compensation statues: the total number of false acquittals, which includes the number of acquitted and convicted-but-exonerated individuals, are 7 and 8, respectively, for $\psi = 0$ and $\psi = 40$. In the IEP treatment, the total number of false acquittals increases from 12 to 17, but the difference does not seem to arise from behavioral changes because the number of guilty pleas is quite similar regardless of the magnitude of compensation.

However, the transition from AEP to IEP significantly increases the error rate: the number of false acquittals increases from 7 to 12 with $\psi = 0$, and 8 to 17 with $\psi = 40$. This dramatic increase in error rates is driven by the risk-taking behavior of guilty subjects. That is, more guilty subjects refused pleas and went to trial when the accuracy declined from AEP to IEP: the number of guilty pleas went down from 41 to 29 with $\psi = 0$, and 38 to 30 with $\psi = 40$. Such an increase in the number of guilty subjects facing trial leads to a higher number of acquittals at trial and a larger population of exonerees who are erroneously paid compensation by state funds. Thus, unless our exonerations processes are sufficiently accurate, mandating and introducing post-release compensations could be quite costly, helping those who should not be helped, without any desirable behavioral change on part of defendants.

Using data in Table 5, we can also calculate the social cost for each treatment. Recall that our social cost expression is given as follows:

$$SC \equiv q\left[H(\eta^I)\alpha_1 + \left(1 - H(\eta^I)\right)(1-\delta)\right] + \tau(1-q)\left[H(\eta^G)\alpha_2 + \left(1 - H(\eta^G)\right)\delta\right]$$

We calculate each term in the social cost expression as follows:

- $q = 1/2$
- $H(\eta^I)$: the proportion of innocent individuals going to trial
- $H(\eta^G)$: the proportion of guilty individuals going to trial
- $\alpha_1 = \alpha_2 = 0.4$
- δ: the average plea bargain discount

Table 6 shows the social cost from each treatment. Subtracting the social cost under $\psi = 40$ from that under $\psi = 0$ in AEP and IEP, respectively, we obtain the following linear equations in τ:

$$\text{AEP: } 0.00929 * \tau - 0.01421$$
$$\text{IEP: } 0.02263 * \tau - 0.01007$$

These equations show that the social cost decreases if τ is sufficiently small. In particular, we can find three areas for τ such that

- $\tau < 0.445$: the social cost decreases in both treatments
- $\tau \in [0.445, 1.531)$: the social cost decreases in AEP but increases in IEP
- $\tau > 1.531$: the social cost increases in both treatments

These findings suggest that a higher level of compensation improves welfare if society is sufficiently averse to wrongful convictions relative to false acquittals. The main factor behind this finding is that the error rate for guilty (innocent, respectively) individuals increases (decreases, respectively) in response to a higher level of compensation in both treatments.

5. Concluding Remarks and Discussion

We presented a theory of endogenous plea bargaining in which a sentence-maximizing prosecutor offers a discounted sentence to a risk averse defendant who has private information regarding whether he is guilty or innocent. Using this model, we found two opposing effects of exoneree compensation on the prosecutor's discount choice, *competitive compensation effect* and *screening effect*, and discussed the mechanism behind these effects. We then brought this model to the lab, letting subjects play a simplified version of our model, and found interesting behavioral responses to exoneree compensation.

Our main experimental finding shows that subjects' behaviors are different depending on the accuracy of exoneration processes. When exoneree compensation is introduced in an accurate exoneration regime, it did not significantly change prosecutor behavior but reduced the number of innocent pleas while keeping constant the number of guilty individuals pleading guilty. In contrast, when exoneree compensation is introduced in an inaccurate exoneration regime, it did significantly change prosecutor behavior but did not significantly change defendant behavior. Thus, our findings suggest that the compensation statutes can be an effective policy tool in reducing false guilty pleas and wrongful convictions when combined with accurate exoneration processes. Moreover, as our prosecutor subjects were free to choose their plea discounts regardless of the magnitude of compensation, our result suggests that the effectiveness of the statues does not require coordination between the prosecution office in charge of plea bargaining and the state legislative in charge of the statutes.

In our experiment, whether the defendant is either guilty or innocent was given by the experimenter. An interesting avenue for future research is to endogenize the crime decision and study the effect of compensation on the level of deterrence, which has not been addressed in the literature except by Fon and Schäfer (2007) and Doménech and Puchades (2015). In light of our experimental findings, the introduction of exoneree compensation could reduce the level of deterrence if exoneration processes are not accurate. In particular, if compensation statues increase the number of false acquittals as shown in Table 5(b), potential criminals could have higher incentive to commit a crime, thereby decreasing the level of deterrence. We leave this topic for future research.

Reference

Abeler, J., A. Falk, L. Goette, and D. Huffman (2011), "Reference Points and Effort Provision", *American Economic Review*, 101, 470-92.

Ambrust, S. (2004), "When Money Isn't Enough: The Case for Holistic Compensation of the

Wrongfully Convicted", *American Criminal Law Review*, 41, 157-182.

Bar-Gill, O. and O. Gazal-Ayal (2006), "Plea Bargains Only for the Guilty", *Journal of Law and Economics*, 49, 353-364.

Bernhard, A. (1999), "When Justice Fails: Indemnification for Unjust Conviction", *University of Chicago Law School Roundtable*, 6, 73-112.

Bernhard, A. (2009), "A Short Overview of the Statutory Remedies for the Wrongly Convicted: What Works, What Doesn't, and Why", *Public Interest Law Journal*, 18, 403-425.

Blackstone, W. (1765), *Commentaries on the Laws of England*, Oxford: Clarendon Press.

Bowers, J. (2008), "Punishing the Innocent", *University of Pennsylvania Law Review*, 156, 1117-1179.

Campbell, K. and M. Denov (2004), "The Burden of Innocence: Coping with Wrongful Imprisonment", *Canadian Journal of Criminology and Criminal Justice*, 46, 139-164.

Chen, D. L., M. Schonger, and C. Wickens (2016), "oTree—An Open-Source Platform for Laboratory, Online, and Field Experiments", *Journal of Behavioral and Experimental Finance*, 9, 88-97.

Doménech, G. and M. Puchades (2015), "Compensating Acquitted Pre-trial Detainees", *International Review of Law and Economics*, 43, 167-177.

Fon, V. and H.-B. Schäfer (2007), "State Liability for Wrongful Conviction: Incentive Effects on Crime Levels", *Journal of Institutional and Theoretical Economics*, 163, 269-284.

Gazal-Ayal, O. (2006), "Partial Ban on Plea Bargains", *Cardozo Law Review*, 27, 2295-2349.

Gross, S., K. Jacoby, D. J. Matheson, N. Montgomery, and S. Patil (2005), "Exonerations in the United States 1989 through 2003", *Journal of Criminal Law and Criminology*, 95, 523-553.

Gross, S. (2008), "Convicting the Innocent", *Annual Review of Law and Social Science*, 4, 173-192.

Grounds, A. (2004), "Psychological Consequences of Wrongful Conviction and Imprisonment", *Canadian Journal of Criminology and Criminal Justice*, 46, 165-182.

Mandery, E. J., A. Schlosberg, V. West, and B. Callaghan (2013), "Compensation Statues and Post-exoneration Offending", *Journal of Criminal Law and Criminology*, 103, 553-584.

Mungan, M. C. and J. Klick (2016), "Reducing False Guilty Pleas and Wrongful Convictions through Exoneree Compensation", *Journal of Law and Economics*, 59, 173-189.

Norris, R. J. (2012), "Assessing Compensation Statues for the Wrongly Convicted", *Criminal Justice Policy Review*, 23, 352-374.

Westervelt, S. D. and K. J. Cook (2008), “Coping with Innocence after Death Row”, *Contexts*, 7, 32-37.

Westervelt, S. D. and K. J. Cook (2010), “Framing Innocents: The Wrongly Convicted as Victims of State Harm”, *Crime, Law, and Social Change*, 53, 259-275.

Table 1: Sequence of Rounds

Round	1	2	3	4	5	6	7	8	9	10	11	12
P1	P	I	G	P	G	P	I	G	P	P	I	P
P2	G	P	P	I	P	I	P	P	G	I	P	G
ψ	0	40	0	40	40	40	0	0	40	0	40	0

Note: P = prosecutor; G = guilty defendant; I = innocent defendant; ψ = exoneree compensation; Gray = practice rounds.

Table 2: Regression Results for Plea Bargain Discounts

	AEP		IEP		Full	
Dep. Var.	Plea Bargain Discounts					
Comp	1.46	1.46	4.14***	4.29***	4.14***	4.29***
	(1.31)	(1.30)	(1.53)	(1.53)	(1.43)	(1.44)
AEP					1.19	0.78
					(1.42)	(1.43)
AEP*Comp					-2.67	-2.83
					(2.01)	(2.01)
Control	no	yes	no	yes	no	yes
N	224	224	216	212	440	436

Note: *** indicates statistical significance at 0.01.

Table 3: Regression Results for Accept Decisions

	AEP		IEP		Full	
Dep. Var.	Accept					
Compensation	-0.13***	-0.12***	-0.09*	-0.08	-0.11***	-0.11***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.03)	(0.04)
Discount	0.02***	0.02***	0.02***	0.02***	0.02***	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Innocent	-0.58***	-0.58***	-0.37***	-0.37***	-0.58***	-0.58***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
IEP					-0.16***	-0.17***
					(0.05)	(0.05)
Innocent*IEP					0.21***	0.22***
					(0.07)	(0.07)
Control	no	yes	no	yes	no	yes
# of Obs.	224	224	216	212	440	436

Note: *** and * indicate statistical significance at 0.01 and 0.1, respectively.

Table 4: Subsample Regression Results for Accept Decisions

	Guilty		Innocent	
Dep. Var.	Accept			
Compensation	-0.11**	-0.11**	-0.12***	-0.11***
	(0.05)	(0.06)	(0.04)	(0.04)
Discount	0.02***	0.02***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)
IEP	-0.16***	-0.16***	0.06	0.03
	(0.05)	(0.06)	(0.04)	(0.04)
Control	no	yes	no	yes
# of Obs.	220	218	220	218

Note: *** and ** indicate statistical significance at 0.01 and 0.05, respectively.

Table 5: The Number of Wrongful Convictions and False Acquittals

(a) Wrongful Convictions for Innocent Individuals

	AEP				IEP			
	$\psi = 0$		$\psi = 40$		$\psi = 0$		$\psi = 40$	
	# of Obs.	Per.	# of Obs.	Per.	# of Obs.	Per.	# of Obs.	Per.
Total	56		56		54		54	
Pleas	10	0.179	2	0.036	10	0.185	8	0.148
Trial	46	0.821	54	0.964	44	0.815	46	0.852
Acquitted	28	0.500	35	0.625	20	0.370	30	0.556
Convicted	18	0.321	19	0.339	24	0.444	16	0.296
Comp.	16	0.286	18	0.321	20	0.370	10	0.185
~Comp.	2	0.036	1	0.018	4	0.074	6	0.111

(b) False Acquittals for Guilty Individuals

	AEP				IEP			
	$\psi = 0$		$\psi = 40$		$\psi = 0$		$\psi = 40$	
	# of Obs.	Per.	# of Obs.	Per.	# of Obs.	Per.	# of Obs.	Per.
Total	56		56		54		54	
Pleas	41	0.732	38	0.679	29	0.537	30	0.556
Trial	15	0.268	18	0.321	25	0.463	24	0.444
Acquitted	6	0.107	7	0.125	8	0.148	11	0.204
Convicted	9	0.161	11	0.196	17	0.315	13	0.241
Comp.	1	0.018	1	0.018	4	0.074	6	0.111
~Comp.	8	0.143	10	0.179	13	0.241	7	0.130

Table 6: Social Cost

		Social Cost
AEP	$\psi = 0$	$0.20293 * \tau + 0.217143$
	$\psi = 40$	$0.21221 * \tau + 0.20293$
IEP	$\psi = 0$	$0.19570 * \tau + 0.22000$
	$\psi = 40$	$0.21833 * \tau + 0.20993$

Note: τ represents the relative weight on letting guilty individuals set free.

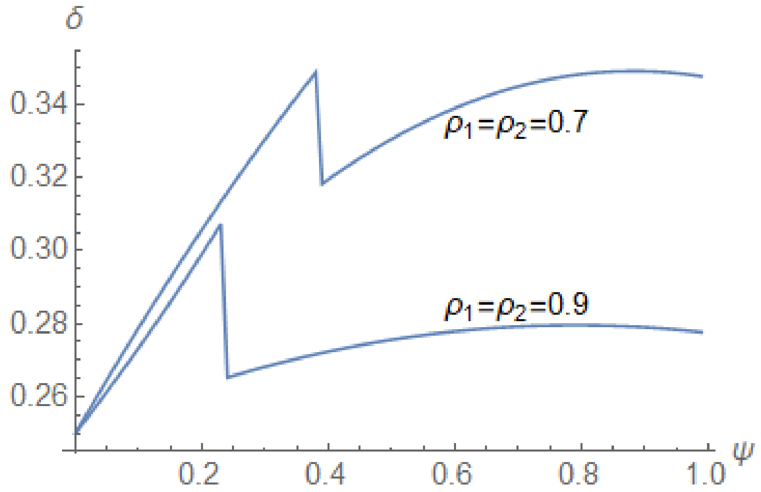


Figure 1. Trajectory of Equilibrium Discount

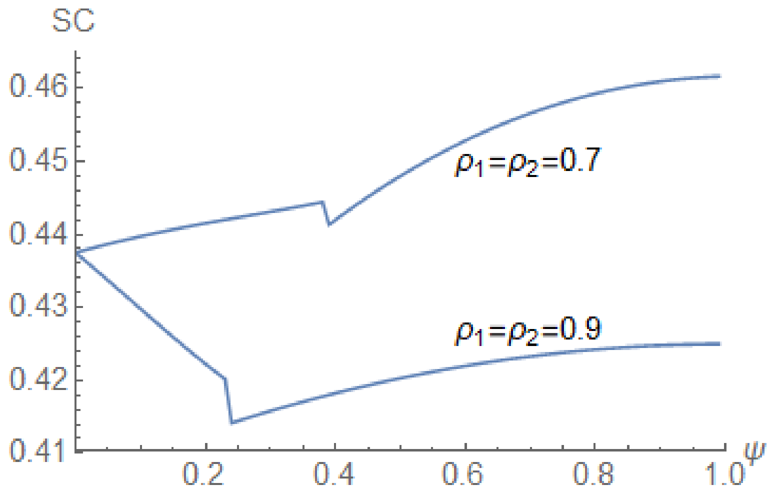


Figure 2. Equilibrium Social Cost

Figure 3. The Sequence of Stages in Each Round

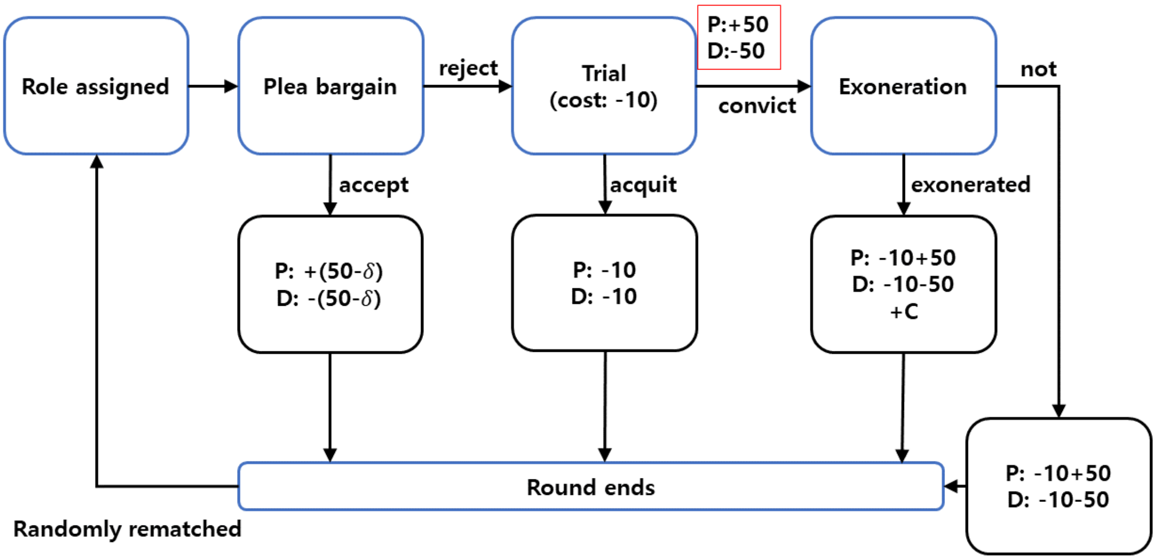


Figure 4. Average Discounts offered by Prosecutors

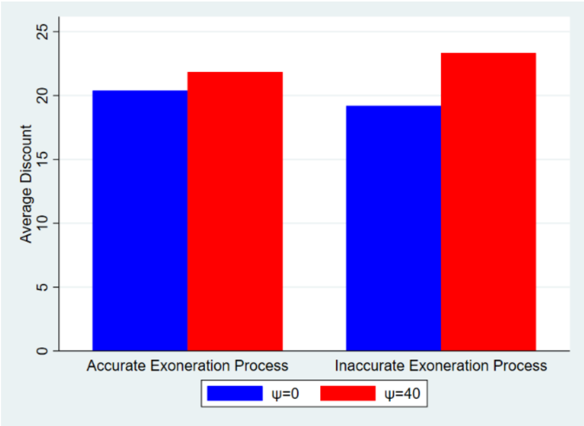


Figure 5: Average Accept Rates of the Innocent

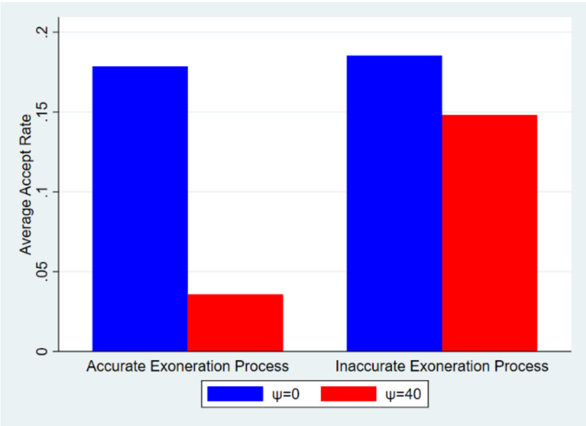
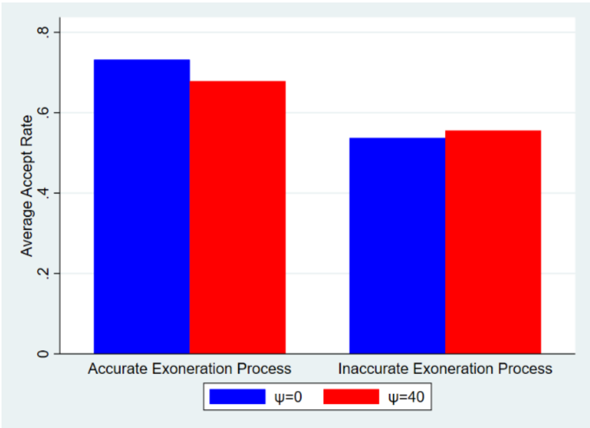


Figure 6: Average Accept Rates of the Guilty



Appendix: Experimental Instruction

<u>Instruction</u>

Thank you for participating in the experiment. Please read the following instruction carefully. All decisions of participants in the experiment are anonymously collected and used only for research. No one knows what your decisions are in the experiment.

The coins in your "account" are given to you in the end of experiment. 600 coins are added to your account in the beginning of experiment. Coins are added to and subtracted from your account in the experiment.

You will be paired with someone in this room. You and your partner do not know each other during and after the experiment.

The experiment proceeds as follows.

1. Basic setting:

- There are four practice rounds and eight real rounds (twelve rounds in total) in this experiment. The process explained below is one round, and the same process is repeated for twelve times.
- In each round, participants in this room are randomly grouped into pairs, and one participant becomes Prosecutor and the other participant becomes Defendant.
- Defendant is equally likely to be innocent or guilty. Defendant knows whether he/she is innocent or guilty, but Prosecutor does not know it. If Defendant is convicted at trial, Defendant is given a sentence of 50.
- Each participant begins the experiment with 600 coins in his/her account. In the end of experiment, each coin in your account is converted to KRW 15 and given to you in cash.
- Prosecutor obtains more coins when Defendant is given a higher sentence, and Defendant obtains more coins when his/her sentence is lower.

2. Plea Bargaining Stage:

- A negotiation begins in which Prosecutor lowers sentence in exchange for Defendant's admitting guilt.
- Prosecutor can choose to lower sentence by D (a number between 0 and 50). After observing Prosecutor's offer, Defendant chooses whether to admit guilt or not.
- If Defendant chooses to admit guilt, his/her sentence is determined to be 50-D. In this case, 50-D coins are added to Prosecutor's account, 50-D coins are subtracted from Defendant's account, and the round ends.
- If Defendant chooses to not admit guilt, they move to trial.

3. Trial Stage:

- In Trial Stage, 10 coins are subtracted from Prosecutor's and Defendant's accounts as litigation costs.
- In Trial Stage, Defendant is convicted or acquitted according to a given probability. An innocent Defendant is more likely to be acquitted than a guilty Defendant. More precisely, an innocent Defendant is acquitted with probability 60% whereas a guilty Defendant is acquitted with probability 40%.
- If Defendant is acquitted, no change is made to Prosecutor's and Defendant's accounts and the round ends.
- If Defendant is convicted, Defendant's sentence is determined to be 50. In this case, 50 coins are added to Prosecutor's account and 50 coins are subtracted from Defendant's account. They move to Exoneration Stage afterwards.

4. Exoneration Stage

- Although Defendant is convicted in Trial Stage, the decision can be reversed in Exoneration Stage.
- An innocent Defendant is more likely to be exonerated than a guilty Defendant in Exoneration Stage. More precisely, an innocent Defendant is exonerated with probability 90% whereas a guilty Defendant is exonerated with probability 10%. <u>[This is the instruction for the AEP treatment. For the IEP treatment, these probabilities are 60% and 40%, respectively.]</u>
- If Defendant is not exonerated, no change is made to Prosecutor's and Defendant's accounts, and the round ends.

- If Defendant is exonerated, no change is made to Prosecutor's account but Defendant obtains exoneree compensation for the loss from wrongful conviction, and the round ends. The amount of exoneree compensation is randomly chosen to be either 0 or 40 coins, and the amount is announced in the beginning of each round.

Please do not talk to each other and do not use a cell phone and internet until the experiment ends. You do not have to hurry when others finish early. If you have any question, please raise your hand. Please wait for further instruction from the experimenter.

Appendix: Proofs

Proof of Proposition 1

For ease of presentation, let us first define the followings:

$$F^I(\delta,\psi) \equiv H(\eta^I)\alpha_1 + [1-H(\eta^I)](1-\delta)$$

$$F^G(\delta,\psi) \equiv H(\eta^G)(1-\alpha_2) + [1-H(\eta^G)](1-\delta)$$

Then, $\delta^*(\psi)$ can be written as:

$$\delta^*(\psi) = \operatorname*{argmax}_{\delta} qF^I(\delta,\psi) + (1-q)F^G(\delta,\psi)$$

Provided that an interior solution exists, the first-order condition is given by:

$$qF^I_\delta(\delta^*,\psi) + (1-q)F^G_\delta(\delta^*,\psi) = 0$$

We differentiate the first-order condition with respect to ψ, and arrange the terms to obtain:

$$\frac{\partial\delta^*(\psi)}{\partial\psi} = -\frac{qF^I_{\delta\psi}(\delta,\psi) + (1-q)F^G_{\delta\psi}(\delta,\psi)}{qF^I_{\delta\delta}(\delta,\psi) + (1-q)F^G_{\delta\delta}(\delta,\psi)}$$

Because the denominator is negative due to the concavity of $F^I(\delta,\psi)$ and $F^G(\delta,\psi)$ (i.e., the second-order condition), $\frac{\partial\delta^*}{\partial\psi} > 0$ if $F^I_{\delta\psi}(\delta,\psi)$ and $F^G_{\delta\psi}(\delta,\psi)$ are positive. In other words, the optimal discount increases in exoneree compensation if they are "complementary" in the prosecutor's objective function. If $\eta^I \in (0,\bar{\eta})$, $F^I_{\delta\psi}(\delta,\psi)$ is:

$$F^I_{\delta\psi}(\delta,\psi) = \frac{1}{\bar{\eta}}\left[\frac{\alpha_1\rho_1}{\sigma_I(\psi)} - \frac{\sigma_I'(\psi)}{\sigma_I(\psi)^2}(s(1-\delta-\alpha_1) + \alpha_1\rho_1\psi)\right]$$

which is positive if $\sigma_I'(\psi)$ is negative. Similarly, one can easily show that $F^G_{\delta\psi}(\delta,\psi)$ is positive if $\sigma_G'(\psi)$ is negative.

Proof of Proposition 2

When $\delta^*(\psi)$ is an interior solution, it is characterized by the following first-order condition:

$$qF^I_\delta(\delta^*,\psi) + (1-q)F^G_\delta(\delta^*,\psi) = 0$$

If $\psi > \hat{\psi}$, $F^I_\delta(\delta^*,\psi) = 0$. Otherwise,

$$F^I_\delta(\delta^*,\psi) = \frac{1}{\bar{\eta}}\frac{s}{\sigma_I(\psi)}(1-\delta^*-\alpha_1) - [1-H(\eta^I)]$$

which is approximately $\frac{1}{\eta}\frac{s}{\sigma_I(\psi)}(1-\delta^*-\alpha_1)$ when ψ is slightly smaller than $\hat{\psi}$. Note that $1-\delta^*-\alpha_1$ is positive because the first-order condition cannot be satisfied otherwise. Thus, when ψ is smaller than and sufficiently close to $\hat{\psi}$, $F_\delta^I(\delta^*,\psi)$ is positive.

Define $\psi_0 \equiv \hat{\psi}-\varepsilon$ and $\psi_1 \equiv \hat{\psi}+\varepsilon$ for some positive ε, and let δ_0 and δ_1 be the corresponding optimal discounts. Then, the first-order conditions for ψ_0 and ψ_1 are:

$$q \times (\text{some positive value}) + (1-q)F_\delta^G(\delta_0,\psi_0) = 0$$

$$(1-q)F_\delta^G(\delta_1,\psi_1) = 0$$

These two equations imply that $F_\delta^G(\delta_1,\psi_1) > F_\delta^G(\delta_0,\psi_0)$. Note that given a fixed δ, $F_\delta^G(\delta,\psi)$ is continuous in ψ. Furthermore, $\psi_0 \approx \psi_1$ for small ε. Thus, for small ε, $F_\delta^G(\delta_1,\psi_1) > F_\delta^G(\delta_0,\psi_0) \approx F_\delta^G(\delta_0,\psi_1)$, which implies that $\delta_0 > \delta_1$ since $F_\delta^G(\delta,\psi)$ is decreasing in δ.

Proof of Proposition 3

Recall that the social cost is defined as:

$$SC = q\left[H(\eta^I)\alpha_1 + \left(1-H(\eta^I)\right)(1-\delta)\right] + \tau(1-q)\left[H(\eta^G)\alpha_2 + \left(1-H(\eta^G)\right)\delta\right]$$

And, (δ^S,ψ^S) and (δ^M,ψ^M) are defined as:

$$(\delta^S,\psi^S) = \underset{(\delta,\psi)}{\operatorname{argmin}}\, SC(\delta,\psi)$$

$$\psi^M = \underset{\psi}{\operatorname{argmin}}\, SC\,(\delta^M(\psi),\psi)$$

where $\delta^M(\psi) = \underset{\delta}{\operatorname{argmax}}\ SC(\delta,\psi)$. Also, recall that the equilibrium levels (δ^*,ψ^*) are defined as:

$$\psi^* = \underset{\psi}{\operatorname{argmin}}\, SC\,(\delta^*(\psi),\psi)$$

where $\delta^*(\psi)$ is the solution of the following problem.

$$\max_\delta q\left[H(\eta^I)\alpha_1 + \left(1-H(\eta^I)\right)(1-\delta)\right] + (1-q)\left[H(\eta^G)(1-\alpha_2) + \left(1-H(\eta^G)\right)(1-\delta)\right]$$

Note that the prosecutor's and the social planner's objective functions are identical if $q=1$. However, the prosecutor wants to maximize it, while the social planner wants to minimize it. Thus, when $q=1$, $(\delta^*,\psi^*) = (\delta^M,\psi^M)$. Similarly, when $q=0$, the two objective functions are practically identical, but have the opposite signs. Because the prosecutor wants to maximize it, while the social planner wants to minimize it, they end up working for the same goal. Thus, $(\delta^*, \psi^*) = (\delta^S, \psi^S)$.

Provided that (δ^S, ψ^S), (δ^M, ψ^M), and (δ^*, ψ^*) are continuous in q in the neighborhoods of zero and one, (δ^*, ψ^*) converges to (δ^S, ψ^S) as q decreases to zero, and to (δ^M, ψ^M) as q increases to one.