



**Leibniz FH**  
SCHOOL OF BUSINESS

# **Prosecution and Conviction under Hindsight Bias in Adversary Legal Systems**

Autor:

Robin Christmann

Autor:

**Robin Christmann**  
Leibniz-Fachhochschule  
christmann@leibniz-fh.de

ISSN 2511-7491

Redaktion: Robin Christmann

Druck: Quensen Druck+Verlag GmbH

# Prosecution and Conviction under Hindsight Bias in Adversary Legal Systems

ROBIN CHRISTMANN

## Zusammenfassung / Abstract:

The plea bargaining mechanism in criminal procedure serves as a favorable screening device, separating between the guilty and the innocent. Previous literature ignored the impact of asymmetric information on prosecutor performance inside the adversarial court, which degrades his bargaining position. This paper presents a sequential prosecution game with endogenous courts, and shows that the successful conviction in court crucially depends on prosecutor's beliefs and incentives. If the prosecutor is sufficiently convinced of the defendant's guilt ex-ante, he can commit to trial, and the favorable semiseparating equilibrium is obtained. Applying the first formal model of a hindsight biased prosecutor, we find that the negative impact of uncertainty on prosecutor performance is partly mitigated by hindsight bias, and the self-selection of guilty defendants can even improve. Several caveats, like excessive charges, the nature of the case or the quality of investigations by the police force are discussed.

**Fachrichtung / Field of Study:** Economics

**Klassifikation / Classification:** D83, D91, K14, K41 (JEL)

**Schlagworte / Keywords:** criminal procedure, plea bargaining, screening, courts, limited rationality

## 1. INTRODUCTION

Prosecutors are meant to solve crime and bring criminals to trial. A common element of criminal procedure, however, is the plea bargain between the prosecutor and the defendant in which the latter pleads guilty to a reduced charge, and court procedures are avoided. Despite a controversial public and legal debate about such deals, for a long time economically-oriented scholars have argued that plea bargaining is socially desirable. Such pretrial agreements would be negotiated in the shadow of the court's jurisdiction, but save resources and eliminate the risk inherent to any trial.

Another desirable feature of plea bargaining is the revelation of hidden information. Even though the prosecutor does not know the actual guilt of a suspect, guilty and innocent defendants may show different reactions to a given plea offer. Thus, the plea bargaining mechanism can induce at least a partial separation between the guilty and the innocent. In a game-theoretic approach, BAKER and MEZZETTI (2001) examine the strategic interaction between a rational prosecutor and a rational defendant with exogenous courts, and find such a semiseparating equilibrium: some guilty defendants reveal themselves and accept the plea bargain, while the remaining guilty and all innocent defendants reject it and move to trial. In the world of this model, it is the credible threat of the prosecutor to go to trial whenever a plea offer is rejected that drives the favorable self-selection process.

Although the theoretical results appear reassuring, it remains unclear whether the actual bargaining situation is completely understood. Prosecutors are plagued by uncertainty, as they seek to avoid convicting innocent individuals. How can a prosecutor credibly threaten to charge a defendant when uncertainty will persist inside courtroom and the decision of the adversarial court totally hinges on the prosecutor's own line of argument? What happens if guilty defendants further manipulate the prosecutor's confidence in his case? The established models with exogenous courts ignore the potentially negative impact of uncertainty on prosecutor performance inside courts, which also degrades his bargaining power. Yet, plea bargaining still appears to be common practice among prosecutors (see, e.g., GAROUPA 2012, EASTERBROOK 2013).

In order to resolve this puzzle, it is the aim of this paper to study the self-selection mechanism of plea bargaining when trials are endogenous. To distinguish our results from the previous literature, we use the BAKER and MEZZETTI (2001) model as framework, and introduce an additional litigation stage. We show that some negative effects of uncertainty on prosecutor performance are possibly mitigated if the prosecutor is no longer assumed to be a rational (Bayesian) decision-maker. In this regard, we provide the first formal analysis of a prosecutor under hindsight bias<sup>1</sup>, and derive implications for the efficiency of plea bargaining.

The paper is organized as follows: chapter 2 reviews the related literature. In chapter 3, the basic framework of the prosecution game and the formal concept of hindsight bias are introduced. Chapter 4 analyzes the different stages of the game, and chapter 5 identifies its equilibria. Welfare implications are presented in chapter 6. Chapter 7 discusses the main contribution to the literature, and chapter 8 concludes.

## 2. RELATED LITERATURE

The economic analysis of public law enforcement was essentially sparked by the landmark articles of BECKER (1968) and LANDES (1971) who applied the concepts of maximization and scarcity to the criminal justice system. Since then, scholars have focused (i) on deterrence and punishment (see, among others, MALIK 1990, EHRLICH 1996, ENTORF 2011) and (ii) on public choice applications, such as prosecutor elections (see, e.g., GORDON and HUBER 2002, HILTON and KHANNA 2007, VAN AAKEN et al. 2008, MCCANNON 2013). Nevertheless, the main interest of researchers has yet been dedicated to the decision-making of the prosecutorial body and the controversial plea bargaining mechanism, which is also the topic

---

<sup>1</sup> Hindsight biased defines the phenomenon that decision-makers fail to correctly remember how uninformed they were ex-ante when they are confronted with outcome knowledge. As a consequence, hindsight biased agents tend to believe more strongly in what they observe than a rational decision-maker would do (see, e.g., FISCHHOFF 2003, p.304).

of this paper. Compared to civil law or antitrust law, however, the law & economics literature in this field may still be relatively underdeveloped (see, e.g., GAROUPA 2012).

The literature on the plea bargaining mechanism, as the criminal law equivalent to pretrial settlements in commercial law (see, e.g., PRIEST and KLEIN 1984), basically follows two major strands.

The first strand of literature regards plea bargaining as a pricing device for crimes (see, among others, LANDES 1971, ADELSTEIN 1979, COVEY 2009, EASTERBROOK 2013). From that perspective, both the prosecutorial office and the suspect are subject to a scarcity constraint, and bargaining offers a *Pareto*-improvement: Prosecutors want to conclude cases and achieve guilty pleas while defendants yield simply for receiving a shorter sentence. The punishment as a price for a criminal act thus reflects a trade of acquittal probabilities for prison time.<sup>2</sup> Based on this pricing concept, several requirements for successful plea bargains could be derived, such as equal beliefs on trial outcome or risk-neutrality, which are also reminiscent of the traditional settlement literature (see, e.g., SHAVELL 1982, SCOTT and STUNTZ 1992, LEWISH 1999). However, several works have recently questioned the asserted efficiency of this pricing mechanism. BIBAS (2004) argues that the pricing mechanism is also affected in a not systematical manner by impact factors that are not related to the criminal case, such as wealth, class, sex or age, which then creates troubling disparities in public enforcement. Similarly, DELACOTE and ANCELOT (2009) demonstrate that lawyer fee schemes influence bargaining outcome, and finds that hourly wages, which can only be afforded by richer defendants, induce lower sentences and better bargaining offers. Furthermore, the imperfection of trials also carries over to plea bargaining whenever innocent individuals perceive a wrongful admission of guilt as the ‘better price’ (see COVEY 2009). Whether the pricing mechanism of plea bargaining is also applicable and desirable

---

<sup>2</sup> Among legal scholars (not practitioners!), the rationale of plea bargaining, treating the defendant’s right to trial as a commodity, has been heavily criticized (see, among others, ALSCHULER 1983).

for inquisitorial legal regimes remains questionable (see, e.g., BOARI and FIORENTINI 2000, ADELSTEIN and MICELI 2001, GAROUPA and STEPHEN 2008, ALTENHAIN 2010).

The second strand in the literature interprets plea bargaining as a screening device. This game-theoretic perspective was first proposed by GROSSMANN and KATZ (1983) who identified asymmetric information about the defendant's true guilt as the major obstacle to efficient enforcement. The prosecutor's plea offer could then induce an efficiency enhancing self-selection process where the guilty defendants accept the bargain and the innocent defendants reject it. The framework has been extended by REINGANUM (1988), who introduced the observation of a signal about the defendant's type prior to the trial, and KOBAYASHI and LOTT (1996) who considered diverging behavior between guilty and innocent defendants in court. In contrast to the basic model, the latter modification allows for the separating solution even when the defendant types have different attitudes to risk. BAKER and MEZZETTI (2001) then pointed out that the basic screening model relied on the incredible threat of the prosecutor to move to court, as in equilibrium only innocent defendants would reject the plea offer. Given exogenous verification in court, the authors demonstrated that plea bargaining still induces a desirable and subgame perfect self-selection effect where some guilty defendants accept the bargain and all remaining defendants reject it. A similar semiseparating solution is also obtained by BJERK (2007) for endogenous jury decisions, and by KIM (2008) even for the case that the prosecutor cannot credibly commit to charge the defendant. Moreover, BAR-GILL and BEN-SHAHAR (2009) indicate that even though the prosecutor can never credibly commit to charge all defendants, given his budget constraint, the collective refusal of the defendants resembles a public good game, and fails.

Both strands in the plea bargaining literature are typically based on two crucial assumptions, that is the efficient prosecutor hypothesis and rational behavior.

Firstly, the efficient prosecutor model assumes that the prosecutor's preferences coincide with those of society (see, e.g., EASTERBROOK 1983, GROSSMAN and KATZ 1983, REINGANUM 1988, p.717, and largely also BAKER and MEZZETTI 2001, p.165). Consequently, it is socially beneficial that the charge is at the unlimited discretion of the assumed

benevolent prosecutor. A more realistic approach has to consider that the prosecutor is not a social planner but an economic agent. Thus, he reacts to incentives and pursues private goals, such as conviction maximization (see RAMSEYER and RASMUSSEN 2000), career opportunities (BOYLAN and LONG 2005), political success (see BERDEJO and YUCHTMAN 2013), or leisure (see BIBAS 2004). TULLOCK (1975) further highlighted that a private decision, irrespective of the preferences, usually does not consider all relevant costs to society. Given this perspective, the plea bargaining mechanism operates in a principal-agent framework, and it becomes doubtful whether the prosecutor's bargaining strategy is always in society's best interest. As an example, BJERK (2005) found that prosecutors use their discretion about the charged crime to circumvent sentencing guidelines when seeking successful plea bargains. As a positive effect, MICELI (1996) concluded that selfish prosecutors may prevent maximizing legislators from increasing fines as high as possible and cutting apprehension effort, and thus help to maintain deterrence with punishments at reasonable levels.

Secondly, the efficiency of plea bargaining relies on the correct interpretation of punishment prices and observed signals for screening by a rational economic agent (so-called 'REMM-hypothesis'<sup>3</sup>). The growing insights of behavioral economics, however, revealed several constraints to human decision-making, and developed alternative concepts of limited rationality to capture these effects.<sup>4</sup> Concerning the plea bargaining mechanism, several authors suggest distinct biases to distort the behavior of the prosecutor, such as overconfidence, denial, discounting of future costs and the sunk cost fallacy. According to BIBAS (2004), the framing of the plea bargaining situation leads to diverging behavior: the gain-framed prosecutor will be less risk-taking and less aggressive than the loss-framed defendant. Furthermore, the author speculates about a relevant anchoring effect of the initial plea offer, as the initial offer typically serves as a reference for the subsequent

---

<sup>3</sup> Rational-Evaluating-Maximizing-Man.

<sup>4</sup> For example, behavioral economists studied the use of different heuristics and the resulting biases. A more realistic model of human decision-making was proposed by prospect theory (see, e.g., KAHNEMAN and TVERSKY 1984).

negotiations. BURKE (2007) acknowledges that selective information can amplify an a priori opinion, thus inflating the prosecutor's beliefs about the strength of the case. Interestingly, researchers show a remarkable consensus that hindsight bias<sup>5</sup> poses a substantial problem to the correct interpretation of evidence by the prosecutor (see, among others, BIBAS 2004, BURKE 2007, GAROUPA 2012): (i) The prosecutor may subconsciously adjust his ex-ante belief and doubts about the defendant's guilt to fit the observed evidence ('memory distortion'). (ii) When losing/winning a trial, the prosecutor may falsely believe ex-post that he anticipated this outcome all along, and is surprised that others did not ('knew-it-all-along effect'). (iii) When evaluating the defendant's guilt, the prosecutor has the advantage of knowing the outcome and all consequences of the defendant's action with certainty, while the defendant did not when committing the criminal act ('outcome effect').

Previous research has so far established crucial determinants for the efficiency of the plea bargaining mechanism. However, the behavior of the prosecutor *after* having filed the charge is hardly studied, and some further general assumptions remain questionable. It is thus the scope of this paper to study prosecutor behavior under uncertainty inside the court, and to provide a first formal analysis of the effects of limited rationality, exemplified by the hindsight bias, on the efficiency of plea bargaining.

### 3. THE MODEL

In the following, we extend the framework model exemplified by BAKER and MEZZETTI (2001). We endogenize court procedures at the end of the game and thereby introduce

---

<sup>5</sup> Hindsight bias is to be distinguished from other behavioral biases, even though some symptoms appear similar. *Overconfidence* implies that an agent believes himself more capable than the average individual. *Confirmation bias* describes the unconscious selection of information in a way that is always in line with the agent's ex-ante beliefs. *Hindsight bias* distorts the memory of the agent's ex-ante beliefs in a way to be in line with the observed outcome.

asymmetric information to the courtroom. Moreover, we provide a formal concept to integrate hindsight bias and biased learning into our model.

### 3.1 Basic Setup

Consider the legal process as a game with two players, the prosecutor (P) and the defendant (D). The defendant can be either guilty of a crime,  $G$ , or innocent,  $I$ . The type of the defendant is exogenously specified by nature (N) at the beginning of the game, and only the defendant knows his true type. The ex-ante probability of a guilty defendant is defined as  $\phi > 0.5$ , and is common knowledge. This assumption captures the fact that criminal cases are transferred to the prosecutor's office only if the police force has acquired a certain level of evidence (*probable cause*) and at least believes the suspect to be more likely guilty than innocent. Consequently, both players know that a larger fraction  $\phi$  of all defendants is indeed guilty. Information asymmetry exists as the prosecutor does not know for certain the true type of the defendant. Thus, the prosecutor has to form (rational) beliefs about the defendant's type throughout the game. All other information is common knowledge.

The defendant is accused of committing the crime  $X$ , with  $X$  resembling the harm to society. Given a conviction, the defendant receives a utility of  $-X$ , and zero otherwise. The convicted defendant's disutility can be interpreted as prison time or a monetary penalty. The prosecutor receives a utility of  $X$  if the guilty defendant is convicted and a utility of  $-bX$ , with  $b > 0$ , if an innocent defendant is sentenced. The prosecutor always receives a utility of zero if the defendant is set free. Consequently, the prosecutor (and society) is interested in punishing criminals and setting free innocent individuals. The variable  $b$  captures the relative severity of wrongful convictions, which may vary between prosecutors (and societies). Furthermore, the prosecutor receives a utility of  $-cX$ , with  $1 > c > 0$ , if he loses a case in court. This disutility can be interpreted as the damage to the prosecutor's reputation whenever he loses a case he decided to bring to court. The reputational loss is contingent on the severity of the crime, as the prosecutor's reputation

is more affected by the outcome of a severe case, such as murder or rape, than by minor offences.<sup>6</sup> Both players are assumed to be risk-neutral, and maximize their expected utility.

In order to capture major institutional features of adversary legal systems, we assert that the prosecutor has complete bargaining power. Thus, the prosecutor can make a take-it-or-leave-it plea offer to the defendant. Furthermore, as the court has no inquisitorial authority and cannot generate evidence on its own, we stylize litigation in adversary courts in the tradition of TULLOCK (1975) as rent-seeking games. In other words, litigation resembles a “*trial by battle*” (TULLOCK 1975, p.746), and the probability of winning in court is defined by the relative efforts of the litigants. This approach allows us to clarify the nature of the adversarial legal doctrine in our model, even though most adversarial legal systems show some inquisitorial elements, such as the (limited) discretion of the judge to discard evidence, reject motions or advise during an interrogation.

The noncooperative prosecution game consists of four stages as displayed in Fig.1: The plea offer by the prosecutor (Stage I), the reaction to the plea offer by the defendant (Stage II), the prosecutor’s decision to charge (Stage III) and the court (Stage IV). The outcomes  $U_p$ ,  $U_g$ , and  $U_i$  represent the utility of the prosecutor, the guilty defendant and the innocent defendant, respectively, at the end of the game.

At the beginning of the game, nature (N) chooses the defendant’s type, which is either guilty or innocent. At stage 1, the prosecutor then offers a plea bargain  $q$  to the defendant, not knowing his true type. The defendant then can either accept the bargain,

---

<sup>6</sup> In the *BAKER/MEZZETTI model*, the authors propose a constant reputational cost for the prosecutor when losing in trial. We believe this to be unrealistic, as the severity of the crime greatly affects the public interest in a given case and thus puts pressure on the prosecution department. Also, a model with non-contingent reputation costs produces a case separation, and minor cases will never be prosecuted (see BAKER/MEZZETTI, 2001, p.157). This would completely rule out deterrence for minor crimes. Equilibria in a model with contingent reputation costs, however, depend on the effectiveness of the police force, and not on the severity of the crime.

which ends the game at stage 2, or reject it. If the plea offer is rejected, the game continues. The following investigations of the prosecutor then produce an exogenous evidence signal  $s$ , which may reveal the innocence of the defendant (signal  $\underline{s}$ ) with positive probability. At stage 3, the prosecutor observes the rejection of the plea offer and the evidence signal, and decides whether to bring the case to court. If the case is dropped, the game ends. If the prosecutor charges, both players enter the litigation process (stage 4). Then the prosecutor as first-mover can exert effort to convince the judge of the defendant's guilt. The defendant then responds to the accusations and exerts effort as second-mover in the litigation subgame to demonstrate his innocence. The relative efforts of the litigants then specify the probability of success in court, and determine the respective outcomes  $U_p$ ,  $U_g$ , and  $U_i$ .

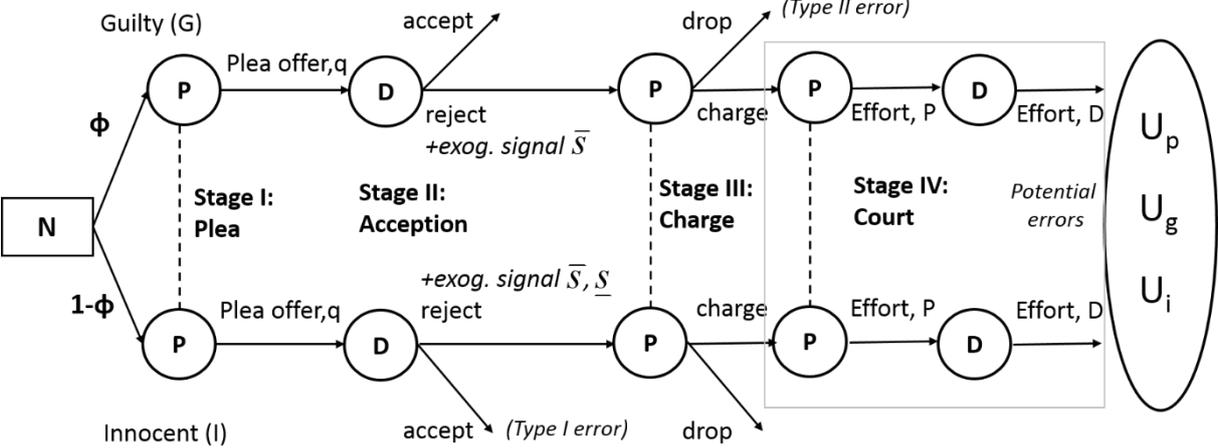


Figure 1. Sequential Prosecution Game with Endogenous Courts.

3.2 Hindsight Bias

Hindsight and foresight differ for decision-makers due to the available information (see FISCHHOFF 2003, p.304). In hindsight, outcome knowledge reduces uncertainty and distinguishes the ex-post evaluation of decisions from the ex-ante perspective. An economic agent will thus revise and update his estimate about the state of the world whenever observing new information. Hindsight biased decision-makers, however, fail to remember how uninformed they were ex-ante, and their judgment is overly affected by outcome knowledge. In other words, the remembrance of their initial estimate will always be closer

to the realized state of the world than their true ex-ante beliefs (*memory distortion*). Hindsight biased decision-makers ‘knew-it-all-along’.

We follow the approach proposed by CAMERER et al. (1989) and BIAS and WEBER (2009) to formally incorporate hindsight bias into our prosecution model. Let  $s$  be a dichotomous evidence signal to the prosecutor indicating the defendant either to be guilty,  $\bar{s}$ , or innocent,  $\underline{s}$ . For the prosecutor, the ex-ante belief about facing a guilty defendant equals the common prior,  $\phi$ . Under hindsight bias, the prosecutor fails to correctly remember this initial estimate as his recollection is tilted towards the observed realization of the signal  $s$ . This can be modelled by defining the distorted remembrance of the common prior,  $\phi_{HB}(\omega)$ , as the weighted average of the true ex-ante probability of guilt and the ex-post probability of supporting evidence, which is one for  $\bar{s}$ , and zero otherwise.

$$\phi_{HB}(\omega) = \omega \cdot s + (1 - \omega) \cdot \phi \quad (1)$$

The distorted remembrance is contingent on the parameter  $\omega \in [0;1]$  which captures the magnitude of the hindsight bias. For  $\omega = 0$ , the decision-maker is unbiased.

The hindsight bias thus distorts the capability of the decision-maker to learn from observations correctly. Given new information, economic agents update their ex-ante beliefs according to Bayes’ Rule. Hindsight biased decision-makers, however, will have to rely on their distorted remembrance of the ex-ante estimate, and are thus subject to *biased Bayesian learning*. In other words, as the remembrance is tilted towards the actual observation, hindsight bias leads to overinference from new information. Moreover, the hindsight biased decision-maker will usually err when estimating the true ex-ante probability from a random sample.<sup>7</sup>

Given the specified prosecution game, only the prosecutor is potentially subject to hindsight bias: the prosecutor does not know whether the defendant is actually guilty, and he has to learn about the defendant’s type from the observed behavior and from the

---

<sup>7</sup> The biased decision-maker may not err if and only if both overinference distortions cancel each other out.

observed evidence of the case. For  $\omega > 0$ , the prosecutor is biased in hindsight and remembers the ex-ante probability of a guilty defendant incorrectly,  $\phi \neq \phi_{HB}(\omega > 0)$ . We assume limited rationality to apply in a way that the prosecutor is not aware of being biased in foresight, thus, e.g., he would be surprised about his incorrect inference from case evidence if the true guilt of the defendants was revealed.

#### 4. PLEA BARGAINING WITH ENDOGENOUS CONVICTION

In this game, the prosecutor's strategy consists of a plea bargain offer  $q$ , the decision to go to trial when observing a rejection of the plea offer and the evidence signal  $s$ , and the effort in court. The defendant's strategy, depending on his type, consists of his reaction to a plea offer, and his effort in court. At each information set, a player thus maximizes his expected payoffs given the strategy choices and beliefs of the other player.<sup>8</sup>

##### 4.1 Stage IV: The Court

The adversary court system is modelled as a TULLOCK (1975) rent-seeking game. Both litigants, the prosecutor and the defendant, may exert costly effort to increase the probability of winning the case. We designate the continuous litigant effort for the prosecutor and defendant as  $P$  and  $D$ , and specify  $P, D \geq 1$ . Thus, the probability of winning the case for the prosecutor can be described by  $P/(P+D)$ , and the probability of success for the defendant yields  $D/(P+D)$ . That is, we assume the probability of prevailing in an adversarial court for a litigant to be determined only by the litigant's effort relative to total effort.<sup>9</sup> For simplicity, marginal effort costs are constant, equal, and set to one. This also

---

<sup>8</sup> Note that due to the circularity between beliefs and equilibrium strategies, the perfect Bayesian equilibrium cannot be perfectly determined via backward induction (see, e.g., FUDENBERG/TIROLE 1999, 326).

<sup>9</sup> This assumption stresses the adversarial nature of the legal system. It also implies that, in the case of equal effort of the litigant, the resulting probability of success would be 50 percent. This specification could easily be altered to produce a higher probability of winning for the prosecutor, capturing effects of further factual

implies, due to  $P, D \geq 1$ , that each party de facto faces a fixed litigation cost when entering the court.

The analysis requires a case separation: (i) the standard case where both parties actively seek to win in court ( $P, D > 1$ ), (ii) the defendant does not actively defend himself ( $P > 1, D = 1$ ), (iii) the prosecutor does not pursue his charge in court ( $P = 1, D > 1$ ), and (iv) both litigants remain inactive and the court is fully arbitrary ( $P, D = 1$ ). In the following, we focus on the most relevant case, case (i).<sup>10</sup>

The defendant, as the second-mover in court, reacts to the effort of the prosecutor and chooses his optimal level of effort  $D$  to defend his case. The defendant thus maximizes his utility function with respect to  $D$ , which yields his reaction function,  $D^R$

$$D^R = \arg \max_D \left[ -X \left( 1 - \frac{D}{P+D} \right) - D \right] = -P + \sqrt{X \cdot P} \quad (2)$$

The defendant receives an expected disutility, dependent on the crime  $X$ , when losing the case with probability  $\left( 1 - \frac{D}{P+D} \right)$ , and incurs effort costs  $D$ . Clearly, his optimal reaction strictly increases with the severity of the crime, and will eventually decrease for high effort levels of the prosecutor. Note that  $D^R > 1$  always holds in the standard case, case (i).

The prosecutor, as the first-mover in court, anticipates the optimal reaction of the defendant and chooses his optimal level of effort  $P$  to prevail in court. As the prosecutor does not know the defendant's true type, he has to form (rational) beliefs  $\mu$  about the defendant's guilt. We specify the prosecutor's belief as  $\mu(G | q, s, \omega)$ , which defines the believed probability that the defendant is truly guilty ( $G$ ), given the observed rejection of the plea offer  $q$  by the defendant, the observed evidence signal  $s$  and the potential

---

evidence, testimonies or superior prosecutor resources. However, our results in equilibrium would be qualitatively unaffected.

<sup>10</sup> An analysis of the remaining cases can be obtained from the author upon request.

hindsight bias  $\omega$ . Given the defendant's behavior  $D^R$  and his beliefs  $\mu$ , the prosecutor's utility function  $U_p(P)$  is defined by

$$U_p(P) = \mu(G|q, s, \omega) \cdot \left[ X \frac{P}{P + D^R} - cX \left( 1 - \frac{P}{P + D^R} \right) \right] + (1 - \mu(G|q, s, \omega)) \left[ (-b)X \frac{P}{P + D^R} - cX \left( 1 - \frac{P}{P + D^R} \right) \right] - P \quad (3)$$

Maximizing with respect to  $P$  then yields the optimal effort of the prosecutor,  $P^*$ , with

$$P^* = \arg \max_P [U_p(P)] = \frac{1}{4} X (\mu(G|q, s, \omega) \cdot (1 + b) - b + c)^2 \quad (4)$$

The prosecutor's effort in court under uncertainty thus increases with the severity of the crime  $X$ , his belief about the defendant's guilt  $\mu$ , and the expected cost to his reputation  $c$  when losing the trial. His effort under uncertainty decreases the higher the disutility from convicting an innocent defendant.

The equilibrium strategies  $[P^*; D^R]$  of the litigants in the court subgame then produce the probability of success for the prosecutor  $\pi_p(\mu)$ , contingent on his beliefs  $\mu$ , as

$$\pi_p(\mu) = \frac{1}{2} \cdot (\mu(G|q, s, \omega) \cdot (1 + b) - b + c) \quad (5)$$

The prosecutor's probability to win the case in court under uncertainty thus increases with his beliefs and the expected reputational costs of losing the trial, and decreases with the disutility of a wrongful conviction. Clearly, information asymmetry produces a strategic disadvantage for the prosecutor and plagues both his effort and his chances in court. The stronger the prosecutor believes the defendant to be guilty, the more confidently he can pursue his charge and win the case. Due to uncertainty in courtroom, the prosecutor wins any given trial with a positive probability which may imply either a correct or a wrongful conviction. In contrast to BAKER/MEZZETTI (2001), the specified court itself shows no positive verifiability to distinguish between guilty and innocent defendants. Thus, our

modelled adversarial court totally relies on the behavior and capability of the prosecutor to be welfare-improving.

#### 4.2 Stage III: The Charge

The prosecutor observes two signals about the defendant's type at stage III. First, he learns that the plea offer  $q$  was rejected. Second, he receives an exogenous evidence signal  $s$  during the following investigations. Given the two signals, the prosecutor is subject to a potential hindsight bias when updating his beliefs and making his decision to charge the defendant.

The generated evidence to the case produces a dichotomous signal  $s \in [\bar{s}; \underline{s}]$  about the defendant's true type with  $\bar{s}$  suggesting that the defendant is potentially guilty and  $\underline{s}$  indicating an innocent defendant. For simplicity, we follow BAKER/MEZZETTI (2001) and assume that the signal  $\underline{s}$  reveals with certainty that the defendant is actually innocent. In other words, the prosecutor interprets this signal as clear proof that the defendant cannot have committed the crime, such as a watertight alibi. While a truly guilty defendant can never provide a rock-solid proof of his innocence, the investigating prosecutor reveals such evidence for the truly innocent defendant with positive probability,  $\sigma$ . We specify  $prob(\underline{s}|G) = 0$  and  $\sigma = prob(\underline{s}|i) > 0$ . We regard  $\sigma$  as the quality of the generated evidence signal, which can be affected by prosecutorial resources and power, procedural rules, previous investigations by the police and the nature of the crime.

Based on his updated beliefs, the prosecutor now decides to charge the defendant with probability  $\theta(q, s)$  for a given rejection of the plea offer  $q$  and the evidence signal  $s$ . Clearly, the prosecutor will never charge on the signal  $\underline{s}$ , which indicates the defendant's innocence, as this implies a certain disutility from effort costs and either from convicting the innocent or losing the trial. The following condition always holds for  $\theta(q, \underline{s}) > 0$ :

$$\begin{aligned}
 U_p(q, \underline{s}) = & -\theta(q, \underline{s}) \cdot [\pi_p(\mu(G|q, \underline{s}, \omega)) \cdot bX + (1 - \pi_p(\mu(G|q, \underline{s}, \omega))) \cdot cX] \\
 & - \theta(q, \underline{s}) \cdot P^* < 0
 \end{aligned} \tag{6}$$

The prosecutor will charge on the signal  $\bar{s}$ , if and only if

$$\begin{aligned}
U_p(q, \bar{s}) = & \mu(G | q, \bar{s}, \omega) \cdot \theta(q, \bar{s}) \cdot [\pi_p(\mu(G | q, \bar{s}, \omega)) \cdot X - (1 - \pi_p(\mu(G | q, \bar{s}, \omega))) \cdot cX] \\
& - (1 - (\mu(G | q, \bar{s}, \omega))) \cdot \theta(q, \bar{s}) [\pi_p(\mu(G | q, \bar{s}, \omega)) \cdot bX + (1 - \pi_p(\mu(G | q, \bar{s}, \omega))) \cdot cX] \quad (7) \\
& - \theta(q, \bar{s}) \cdot P^* > 0
\end{aligned}$$

Inserting conditions (4) and (5), and some further simplifications now yield

$$\mu(G | q, \bar{s}, \omega) > \frac{\sqrt{4c - c + b}}{1 + b} \quad (8)$$

Thus, the prosecutor will only charge the defendant if his beliefs about the defendant's guilt are sufficiently strong. Condition (8) imposes a lower threshold for the prosecutor's beliefs to actually move to court. The higher the disutility from convicting an innocent defendant or the higher the expected reputational cost from losing a case, the more convinced the prosecutor has to be that the suspect is indeed guilty.

Integrating the (biased) Bayesian updating of the prosecutor's beliefs yields further insights about the determinants for pressing a charge. The updated belief is given as

$$\mu(G | q, \bar{s}, \omega) = \frac{\phi_{HB}(\omega)}{\phi_{HB}(\omega) + (1 - \phi_{HB}(\omega))(1 - \sigma)} \quad (9)$$

This defines the probability of actually facing a guilty defendant out of all cases where the suspect rejected the plea offer and the evidence signal  $\bar{s}$  was generated, given that the prosecutor remembers the ex-ante probability of a guilty suspect as  $\phi_{HB}(\omega)$ . The higher the quality of the provided evidence signal,  $\sigma$ , the more unlikely it is to mistakenly face an innocent defendant. Inserting equations (9) and (1) into condition (8) further specifies the above discussed threshold for charging the defendant, and yields

$$\phi > \frac{1}{1-\omega} \left[ \frac{(1-\sigma + \omega\sigma)(\sqrt{4c-c+b}) - \omega(1+b)}{1+b-\sigma(\sqrt{4c-c+b})} \right] = \phi_{MIN} \quad (10)^{11}$$

Based on his incentive structure  $(b, c)$ , the quality of the evidence signal  $\sigma$ , and his limited rationality through hindsight bias  $\omega$ , a distinct ex-ante probability of a guilty defendant,  $\phi_{MIN}$ , is required for the prosecutor to be confident enough to bring the case to court under uncertainty. He will always charge if the prior is sufficiently high,  $\phi > \phi_{MIN}$ , and never otherwise. We find that  $\frac{\partial \phi_{MIN}}{\partial \sigma} < 0$ , which implies that a better evidence signal reduces uncertainty for the prosecutor and thus relaxes the threshold  $\phi_{MIN}$ . Interestingly, we also find  $\frac{\partial \phi_{MIN}}{\partial \omega} < 0$ . Thus, the threshold is also lowered and the prosecutor becomes more confident to go to court if hindsight bias increases. As hindsight bias induces overinference from the observed evidence signal, the prosecutor deems it more likely that the defendant is guilty than a rational decision-maker would. This potentially produces an inefficiency whenever  $\phi_{MIN}(\omega > 0) < \phi < \phi_{MIN}(\omega = 0)$ , as the hindsight biased prosecutor then charges cases that a rational prosecutor would never pursue. To put it differently, the hindsight biased prosecutor is overly confident to confront guilty defendants while he de facto excessively charges innocent individuals.

### 4.3 Stage II: Acception

The defendant can accept or reject an offered plea bargain  $q$  of the prosecutor. Let  $G(q)$  be the probability that the guilty defendant rejects the offer, and  $I(q)$  the probability that the innocent defendant rejects it. A defendant who rejects the plea offer receives a disutility through the expected sentence and costs through litigation effort only if the prosecutor actually decides to put him to trial. Accepting the plea bargain however implies a certain

---

<sup>11</sup> In contrast to the BAKER/MEZZETTI (2001) model, a case distinction for the receiver's behavior based on the common prior is more in line with the standard literature on signaling games (see, e.g., KREPS/WILSON 1982, TIROLE/FUDENBERG 1999).

punishment  $q$  for the defendant. The guilty defendant chooses a strategy  $G(q)$  to maximize his utility, specified by

$$U_g(\cdot) = -G(q) \cdot \theta(q, \bar{s}) \cdot [\pi_p(\mu(G|q, \bar{s}, \omega)) \cdot X + D^*] - (1 - G(q)) \cdot q \quad (11)$$

In the same manner, the innocent defendant chooses his strategy  $I(q)$  to maximize his expected utility, given as

$$U_i(\cdot) = -I(q) \cdot (1 - \sigma) \cdot \theta(q, \bar{s}) \cdot [\pi_p(\mu(G|q, \bar{s}, \omega)) \cdot X + D^*] - (1 - I(q)) \cdot q \quad (12)$$

As a distinct feature, only the innocent defendant benefits from improved evidence to the case,  $\sigma$ , and thus is less likely to face a charge upon rejecting the bargain. Given these considerations, guilty and innocent defendants may choose different strategies for the plea bargain. Consequently, the defendant's decision to accept or reject the plea offer is potentially informative to the prosecutor.

#### 4.4 Stage I: Plea Offer

A perfect Bayesian equilibrium (PBE) for the prosecution game consists of the strategies  $\{I^*(q), G^*(q), D^*, q^*, \theta^*, P^*\}$  and the beliefs  $\mu^*(G|q, s, \omega)$  such that, at any stage of the game, strategies are optimal given the beliefs, and the beliefs are obtained from the equilibrium strategies and observed actions using Bayes' rule (FUDENBERG/TIROLE 1999).

## 5. EQUILIBRIA

Distinct from the established law and economics literature on plea bargaining, we consider limited rationality and introduce endogenous verification under uncertainty to the prosecution game. In this model, the prosecutor may be overly affected by the acquired information and '*knew-it-all-along*'. Furthermore, he is strained by the potential threat to convict an innocent defendant when presenting his accusations before the court and cannot rely on exogenous verification. Both modifications imply a more realistic approach to the prosecution process. We show that the major finding of the plea bargaining literature, the existence of a semi-separating perfect Bayesian equilibrium, still holds, but further

conclusions can be made. As before, we focus on the most relevant case where the litigants exert positive efforts in court ( $P, D > 1$ ).<sup>12</sup>

**Proposition 1.** *There is no separating equilibrium (SE) where the defendant always reveals his type and the prosecutor makes a positive plea offer  $q > 0$ , given the evidence technology is not meaningless,  $\sigma > 0$ .*

Proof. The first candidate for a SE is  $I(q) = 0$  and  $G(q) = 1$ . For  $I(q) = 0$  to be optimal,  $q$  must be  $q < (1 - \sigma) \cdot \theta(q, \bar{s}) \cdot [\pi_p(\mu(G | q, \bar{s}, \omega) = 1) \cdot X + D^*] = q_1$ . For  $G(q) = 1$  to be optimal,  $q$  must be  $q > \theta(q, \bar{s}) \cdot [\pi_p(\mu(G | q, \bar{s}, \omega) = 1) \cdot X + D^*] = q_2$ . As  $q_2 > q_1$ , no such  $q$  exists. The second candidate for a SE is  $I(q) = 1$  and  $G(q) = 0$ . The prosecutor infers  $\mu(G | q, \bar{s}, \omega) = 0$ , and chooses  $\theta(q, \bar{s}) = 0$ . Then,  $G(q) = 0$  is no longer optimal for the guilty defendant. ■

This result is in line with the previous argumentation of BAKER/MEZZETTI (2001), who demonstrated that any separating equilibrium cannot be subgame perfect. First, the guilty defendants expect, on average, a higher sentence than the innocent defendants upon rejection of the plea offer, and thus will still accept plea offers that the innocent will rationally reject. Consequently, a state where all innocent individuals accept the plea offer, and all guilty defendants reject it, cannot be achieved. Second, if all guilty defendants self-select and accept the plea offer, even a prosecutor with limited rationality has no interest in suing the remaining innocent defendants.<sup>13</sup> Clearly, the threat of litigation is no longer credible, and the guilty defendants will then prefer to reject the plea offer as well. Subgame perfectness implies that the plea bargaining mechanism can never achieve a complete separation between guilty and innocent defendants.

**Proposition 2.** *i) Given a case with  $\phi < \phi_{MIN}(\omega)$ , a pooling equilibrium (PE) exists with  $I(q) = 1$ ,  $G(q) = 1$  and  $\theta(q, s) = 0$  for  $q \geq 0$ . ii) Given a case with  $\phi < \phi_{MIN}(\omega)$ ,*

---

<sup>12</sup> An analysis of the remaining cases can be obtained from the author upon request.

<sup>13</sup> Note that the complete separation effectively rules out the problem of overinference from observed signals.

a separating equilibrium (SE) exists for  $q = 0$  and  $\theta(q, s) = 0$  with  $I(1), G(0)$  or  $I(0), G(1)$ , respectively.

Proof. i.) Due to  $\phi < \phi_{MIN}(\omega)$ ,  $\theta(q, s) = 0$  is optimal for the prosecutor. The defendants optimally react with  $I(q) = 1$  and  $G(q) = 1$  for all  $q \geq 0$ . ii.) For  $q = 0$  and  $\phi < \phi_{MIN}(\omega)$ ,  $\theta(q, \bar{s}) = 0$  applies, and the strategies  $I(1), G(0)$  or  $I(0), G(1)$  are (also) optimal for the defendant. ■

This finding shows that plea bargaining cannot serve as a screening device if the prosecutor evaluates the case itself as not sufficiently favorable to move to court. Given the prosecutor never charges a defendant, both types of defendants have no incentive to accept a positive plea offer to evade litigation, which constitutes the pooling equilibrium. Distinct from previous literature, we show that overinference affects the equilibrium conditions. As the threshold  $\phi_{MIN}(\omega)$  decreases with the degree of hindsight bias  $\omega$ , the range for the pooling equilibrium narrows under hindsight bias. To put it differently, the biased prosecutor is more confident about his case, which makes a potential charge more credible and thus incentivizes some defendants to self-select and accept the plea offer.

**Proposition 3.** *Given a case with  $\phi > \phi_{MIN}(\omega)$ , no pooling equilibrium (PE) exists.*

Proof. Due to  $\phi > \phi_{MIN}(\omega)$ , the prosecutor chooses  $\theta(q, \bar{s}) = 1$  and  $\theta(q, \underline{s}) = 0$ . For the prosecutor, very low plea offers cannot be optimal. The innocent defendant accepts any plea offer with  $q < (1 - \sigma) \cdot \theta(q, \bar{s}) \cdot [\pi_p(\mu(G | q, \bar{s}, \omega) = 1) \cdot X + D^*] = q_1$  and the guilty defendant accepts any plea offer with  $q < \theta(q, \bar{s}) \cdot [\pi_p(\mu(G | q, \bar{s}, \omega) = 1) \cdot X + D^*] = q_2$ . As it cannot be optimal for the defendants to accept very high plea offers, a potential pooling strategy must reject some plea offers with a positive probability. A pooling strategy then implies that all defendants accept plea offers with  $q < q_2$ . However,  $q_2 > q_1$  holds and this contradicts the PE. ■

This result points out that for cases with a sufficient ex-ante probability of a guilty suspect,  $\phi > \phi_{MIN}(\omega)$ , no pooling equilibrium exists and thus the behavior of the defendants is potentially informative to the prosecutor. This stresses the fact that the effective groundwork of the police is essential for the plea bargaining mechanism. Furthermore, due to  $\frac{\partial \phi_{MIN}}{\partial \omega} < 0$ , the (wrongfully) increased confidence of the prosecutor, based on overinference, forces the defendants to use diverging strategies, and thus turns the game more informative.

**Proposition 4.** *Given a case with  $\phi > \phi_{MIN}(\omega)$ , then the following strategies and beliefs constitute a semiseparating perfect Bayesian equilibrium:*

$$G^*(q) = \begin{cases} 1, & \text{if } q > q^* \\ \gamma, & \text{else} \end{cases} \quad \text{with } \gamma = \frac{(1 - (\omega + (1 - \omega) \cdot \phi))(1 - \sigma)(\sqrt{4c} - c + b)}{(\omega + (1 - \omega) \cdot \phi)(1 - \sqrt{4c} + c)}$$

$$I^*(q) = 1$$

$$D^R = -P + \sqrt{X \cdot P}$$

$$\theta^*(q, \underline{s}) = 0$$

$$\theta^*(q, \bar{s}) = \begin{cases} 1, & \text{if } q > q^* \\ \frac{q}{X \cdot \pi_p(\mu(G | q, \bar{s}, \omega) + D^*)}, & \text{else} \end{cases}$$

$$q^* = \pi_p(\mu(G | q, \bar{s}, \omega) \cdot X + D^*)$$

$$P^* = \frac{1}{4} X (\mu(G | q, s, \omega) \cdot (1 + b) - b + c)^2$$

$$\mu^*(G | q, s, \omega) = 0 \text{ for all } q$$

$$\mu^*(G | q, \bar{s}, \omega) = \begin{cases} \frac{(\omega + (1 - \omega) \cdot \phi)}{((\omega + (1 - \omega) \cdot \phi) + (1 - (\omega + (1 - \omega) \cdot \phi))(1 - \sigma))}, & \text{if } q > q^* \\ \frac{\gamma \cdot (\omega + (1 - \omega) \cdot \phi)}{(\gamma \cdot (\omega + (1 - \omega) \cdot \phi) + (1 - (\omega + (1 - \omega) \cdot \phi))(1 - \sigma))}, & \text{else} \end{cases}$$

Proof. Given a plea offer  $q$  that satisfies  $q = q_2 > q_1$ , all innocent Defendants will reject  $q$ ,  $I^*(q) = 1$ , and the guilty defendants reject  $q$  with positive probability  $\gamma$ . Firstly, the prosecutor becomes indifferent between charging the defendant or dropping the case, when  $q$  is rejected and  $\bar{s}$  is observed if and only if condition (8)

is binding:  $\mu(G|q, \bar{s}, \omega) = \frac{\sqrt{4c-c+b}}{1+b}$ . Inserting the Bayes's update (9), modified

with the probability  $\gamma$  that the guilty defendant actually rejects the plea offer, this

leads to  $\frac{\gamma \cdot \phi_{HB}(\omega)}{\gamma \cdot \phi_{HB}(\omega) + (1 - \phi_{HB}(\omega))(1 - \sigma)} = \frac{\sqrt{4c-c+b}}{1+b}$ . Solving for  $\gamma$  yields

$$\gamma^* = \frac{(1 - \phi_{HB})(1 - \sigma)(\sqrt{4c-c+b})}{\phi_{HB}(1 - \sqrt{4c+c})} = \frac{(1 - (\omega + (1 - \omega) \cdot \phi))(1 - \sigma)(\sqrt{4c-c+b})}{(\omega + (1 - \omega) \cdot \phi)(1 - \sqrt{4c+c})},$$
 which is

the mixed strategy for the randomizing guilty defendant. Secondly, the guilty defendant randomizes between accepting or rejecting the plea offer  $q$  if  $q = q_2 = \theta(q, \bar{s}) \cdot [\pi_p(\mu(G|q, \bar{s}, \omega)) \cdot X + D^*]$ . The mixed strategy for the randomizing

prosecutor thus is  $\theta(q, \bar{s}) = \frac{q}{X \cdot \pi_p(\mu(G|q, \bar{s}, \omega)) + D^*}$ . Thirdly, the prosecutor offers

a plea bargain that maximizes his utility. A natural candidate for the optimal plea offer is  $q^*$  with  $q^* = q_2 = \pi_p(\mu(G|q, \bar{s}, \omega)) \cdot X + D^*$ . The plea offer  $q^*$  yields a higher utility for the prosecutor than any higher plea offer  $q > q^*$ , if  $\phi \cdot (1 - \gamma) \cdot (\pi_p(\mu(G|q, \bar{s}, \omega)) \cdot X + D^*)$

$$\begin{aligned} &\geq \phi \cdot [\pi_p(\mu(G|q, \bar{s}, \omega)) \cdot X - (1 - \pi_p(\mu(G|q, \bar{s}, \omega))) \cdot cX - P^*] \\ &\quad - (1 - \phi) \cdot (1 - \sigma) \cdot [bX \cdot \pi_p(\mu(G|q, \bar{s}, \omega)) + (1 - \pi_p(\mu(G|q, \bar{s}, \omega))) \cdot cX + P^*] \end{aligned}$$

Some transformations lead to  $D^* \geq -(1 - \pi_p(\mu(G|\bar{q}, s, \omega))) \cdot cX - P^*$  which always holds. The plea offer  $q^*$  also yields a higher utility for the prosecutor than any lower plea offer  $q < q^*$ , as  $\phi \cdot (1 - \gamma) \cdot q^* > \phi \cdot (1 - \gamma) \cdot q$  is always fulfilled. ■

This finding demonstrates that the previously established semiseparating perfect Bayesian equilibrium in a plea bargaining game (see, e.g., BAKER/MEZZETTI 2001) also holds under endogenous litigation with asymmetric information and limited rationality of the prosecutor. If the prosecutor is confident that bringing the case to court is generally favorable,  $\phi > \phi_{MIN}(\omega)$ , the known self-selection mechanism of plea bargaining is induced: some guilty defendants will accept the plea offer while all innocent defendants and the remaining guilty defendants reject it, and the prosecutor always moves to court.

Guilty defendants generally try to mimic the behavior of the innocent defendants and reject the plea offer. However, any increase in rejected plea offers also raises the share of guilty defendants which the prosecutor may face in court. Thus, the ability to imitate the innocent defendants is limited because increased rejections make the charge more favorable to the prosecutor and this further increases prosecutor performance. Our model also reveals a positive effect of hindsight bias, as we find  $\frac{\partial \gamma^*}{\partial \omega} < 0$ . This implies that an increased hindsight bias makes the prosecutor more confident in his charge, and turns it more difficult for the guilty defendant to mimic the innocent individuals. Thus, quite surprisingly, hindsight bias amplifies the self-selection mechanism of plea bargaining.

## 6. WELFARE IMPLICATIONS

### 6.1 Comparative Statics

In the following, we focus on the interesting case of the semiseparating equilibrium, and calculate court errors<sup>14</sup> and social welfare for the equilibrium path. We determine the type I court error (wrongful conviction) with probability

$$e_I = \frac{(1-\phi)(1-\sigma)}{(1-\phi)(1-\sigma) + \gamma^* \cdot \phi} \cdot \pi_p(\mu(G | q^*, \bar{s}, \omega, \gamma = \gamma^*)). \quad (13)$$

The court produces a type II error (wrongful acquittal) with probability

$$e_{II} = \frac{\gamma^* \cdot \phi}{(1-\phi)(1-\sigma) + \gamma^* \cdot \phi} \cdot [1 - \pi_p(\mu(G | q^*, \bar{s}, \omega, \gamma = \gamma^*))] \quad (14)$$

The efficiency of the prosecutorial system is described by a social welfare function  $W$  with

$$W = (1 - \gamma^*) \cdot q^* + \frac{\gamma^* \cdot \phi}{(1-\phi)(1-\sigma) + \gamma^* \cdot \phi} \cdot \pi_p(\mu(G | q^*, \bar{s}, \omega, \gamma = \gamma^*)) \cdot X - (\alpha_1 \cdot e_I + \alpha_2 \cdot e_{II}) \cdot X - (\gamma^* \cdot \phi + (1-\phi)(1-\sigma))(P^* + D^*) \quad (15)$$

---

<sup>14</sup> Given the equilibrium path, type I and type II errors can only occur inside the court: no innocent defendant wrongfully accepts a plea offer, and all guilty defendants who reject the offer are charged by the prosecutor.

Social welfare is enhanced by two factors: the first summand shows the positive effect of accepted plea offers, defined by the probability of an accepted bargain,  $(1-\gamma^*)$ , and the offered sentence,  $q^*$ . The second summand identifies the benefit from correct convictions of guilty suspects, that is the probability that a guilty suspect finds himself in court multiplied with the expected sentence,  $\pi_p \cdot X$ . However, welfare is degraded by two factors: the first subtrahend specifies the costs of court errors, given the weighted probability of a court error and the seriousness of the crime  $X$ . We assume that type I and type II error are weighted with  $\alpha_1$  and  $\alpha_2$  respectively. The second subtrahend illustrated the expenditure of legal resources in court, given the prosecutor files a charge.

Table 1 shows the comparative statics of the semiseparating equilibrium for a given crime  $X$ . Each column considers a change in an exogenous variable of the game. The rows identify the effect of the changed exogenous variable on the endogenous variables, the threshold  $\phi_{MIN}$ , the optimal mixed strategy of the guilty defendant  $\gamma^*$ , the produced probability of court errors  $e_I$  and  $e_{II}$ , and the generated social welfare  $W$ .

	$\Phi$ <sup>1)</sup>	$X$	$b$	$c$ <sup>2)</sup>	$\sigma$	$\omega$
$\Phi_{MIN}$	0	0	+	+	-	-
$\gamma$	-	0	+	+	-	-
$e_I$	0/- <sup>3)</sup>	0	-	?	0	+
$e_{II}$	0/+ <sup>3)</sup>	0	+	?	0	-
$W$	+/? <sup>3)</sup>	?	?	?	+	?

1)  $\Phi \geq 0.5$     2)  $c < 1$     3) for  $(\omega=0/ \omega>0)$

**Table 1. Comparative Statics Results.**

The first two exogenous variables show expectable effects. A higher share of guilty defendants clearly, c.p., increases the prosecutor's effort in court to win a given case. Thus, less plea offers can be rejected by the guilty defendant in order to make the prosecutor indifferent again. This require the defendant to manipulate  $\gamma^*$  in a way that the (perceived) share of guilty defendants in court does not change. Thus, court errors are at least not affected whenever the prosecutor is unbiased. In this case, less costly court procedures are required and  $W$  increases unambiguously. The total effect on welfare is unclear, however,

if the prosecutor is biased. The impact of the seriousness of the crime,  $X$ , is similarly straightforward: Following condition (15),  $X$  amplifies equally all the positive effects of plea bargains and correct decisions, and also the negative impact of court errors and the resources spent. In other words, any positive or negative total effect on social welfare is enlarged by the seriousness of the crime.<sup>15</sup>

The prosecutor's incentives,  $b$  and  $c$ , show more complex implications. As both variables, c.p., increase the expected loss of going to court, the prosecutor has to be more certain ex-ante that he actually faces a guilty suspect, and the threshold  $\phi_{MIN}$  increases. Consequently, it is easier for the guilty defendant to mimic the innocent and reject more pleas. Then, the prosecutor will eventually face a higher share of guilty defendants inside the courtroom. This generates less wrongful convictions and more wrongful acquittals given  $b$  increases, as the prosecutor is more afraid to convict the innocent and equilibrium effort is unaffected. While the negative impact of less successful plea bargains is considerably strong, the total effect on social welfare is also dependent on the assigned weights of court errors,  $\alpha_1$  and  $\alpha_2$ . However, if  $c$  increases, the prosecutor's stakes are raised in court and this boosts his equilibrium effort and probability of winning. Consequently, the net outcome effect on welfare stays ambiguous.

The evidence technology,  $\sigma$ , determines the probability that innocent defendants are not recognized and ruled out from prosecution when assessing the facts of a criminal case. Consequently, an increasing  $\sigma$  improves the strength of the evidence signal to the prosecutor, and lowers the threshold  $\phi_{MIN}$ . Less innocent defendants will be wrongfully charged, and thus the guilty defendants have to accept more plea offers in order to keep the prosecutor indifferent about his charge. As the share of guilty defendants in court in equilibrium does not change, court errors are unaffected. Total welfare increases due to more successful plea bargains.

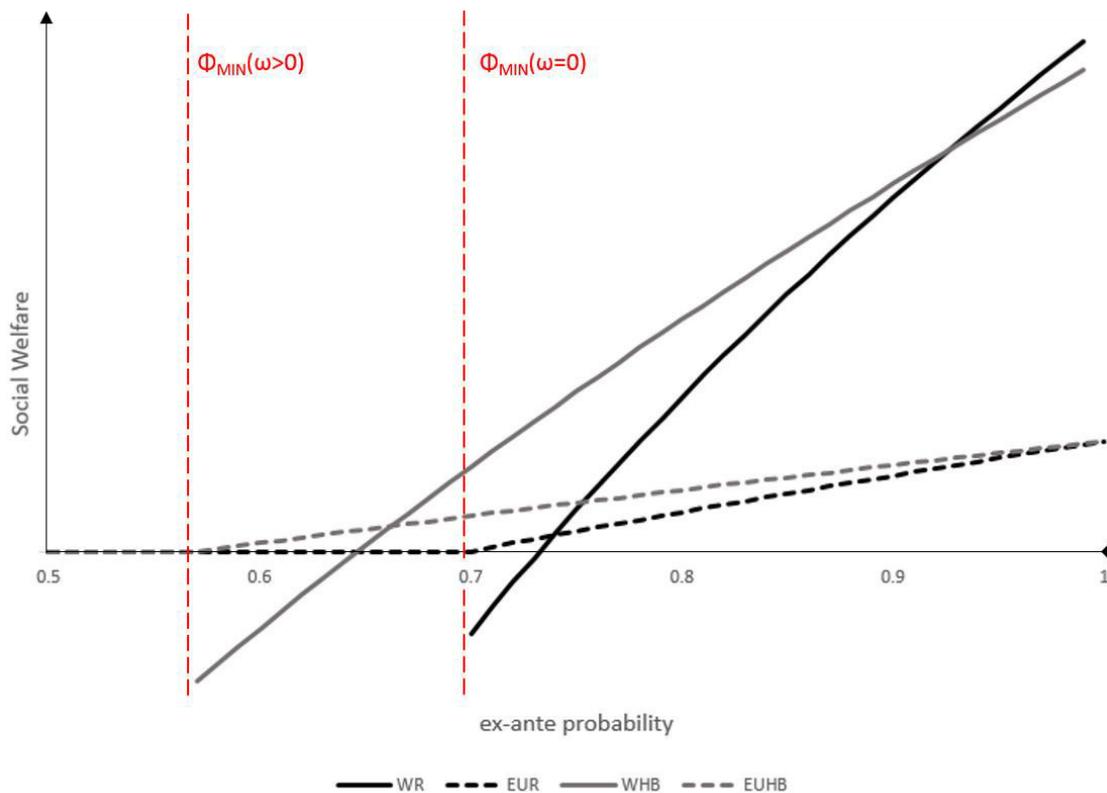
---

<sup>15</sup> Using condition (15) and inserting the equilibrium strategies  $q^*$ ,  $P^*$  and  $D^*$ ,  $X$  could simply be factorized in the social welfare function. Thus,  $X$  only amplifies a positive or negative sign of the remaining factor.

The impact of the evidence technology on outcome has to be carefully distinguished from the effects of hindsight bias  $\omega$ . First, a higher bias (mistakenly) increases the confidence of the prosecutor in the strength of the signal, as he over-infers from the observed facts and behavior. This (mistakenly) lowers the threshold  $\phi_{MIN}$ . The higher confidence of the prosecutor in his charge, c.p., also boost his effort in court. Thus, a higher share of guilty defendants is forced to accept the plea bargains, and turn the confident prosecutor indifferent again. The positive effect of hindsight bias clearly is this increased self-selection effect. The limited rationality of the prosecutor induces de facto a lower share of guilty defendants in court, compared to a fully rational prosecutor, and this produces more wrongful convictions and less wrongful acquittals. While the positive effect of a higher self-selection is quite strong, the net effect on social welfare is eventually also dependent on the chosen weights for court errors,  $\alpha_1$  and  $\alpha_2$ , and stays ambiguous.

## 6.2 Hindsight Bias and the Ex-Ante Probability

So far, we revealed an ambiguous impact of hindsight bias on social welfare if *Proposition 4* holds: while hindsight bias increases self-selection through the plea bargaining mechanism, which is clearly desirable, it also amplifies wrongful convictions in court. A second, potentially negative effect of hindsight bias emerges if we explore its nexus to the ex-ante probability. A hindsight biased prosecutor, who is overly confident that he actually faces a guilty suspect, will litigate cases that a rational prosecutor never would. Fig. 2 illustrates the efficiency effects of this argument.



**Figure 2.** *Hindsight Bias and the Decision to Charge.*

For the case of the rational prosecutor,  $\omega = 0$ , the graph EUR depicts his expected utility of moving to court and charging the defendant.<sup>16</sup> The threshold  $\phi_{MIN}(\omega = 0)$  identifies the ex-ante probability at which the rational prosecutor would be indifferent between a charge or dropping the case, as his expected utility equals zero. For  $\phi > \phi_{MIN}(\omega = 0)$ , the potential charge is a credible threat to the defendant, and the semi-separating equilibrium with some self-selecting guilty defendants is achieved (*Proposition 4*). Expected utility EUR and the generated social welfare WR increase in  $\phi$ . Note that WR turns negative when  $\phi$  falls towards the threshold  $\phi_{MIN}(\omega = 0)$ , as the rational prosecutor does not consider the additional rent-seeking costs for society due to the defendant's effort in court. For  $\phi < \phi_{MIN}(\omega = 0)$ , a charge is not credible to the defendants, and the pooling equilibrium unfolds (*Proposition 2*).

<sup>16</sup> The expected utility is calculated before the defendant randomizes.

For the case of the hindsight biased prosecutor,  $\omega > 0$ , different effects can be observed. First, the expected utility of the biased prosecutor becomes inflated due to overinference, as depicted by EUHB. As a consequence, the biased prosecutor will move to court even when the ex-ante probability is comparatively low,  $\phi_{MIN}(\omega > 0) < \phi < \phi_{MIN}(\omega = 0)$ . Second, the improved confidence of the prosecutor amplifies the beneficial self-selection, but also increases wrongful convictions. The social welfare under hindsight bias, WHB, shows that the gain in self-selection even overcompensates potential error costs for some  $\phi$ . Furthermore, the (wrongfully) overly confident prosecutor achieves some efficiency enhancing litigation for even lower ex-ante probabilities (below, but close to the threshold  $\phi_{MIN}(\omega = 0)$ ). As  $\phi$  drops, the costs of court errors increase, however, and eventually dominate the gains from self-selection. The nexus between the ex-ante probability of a guilty defendant and the extent of limited rationality thus are crucial to understand whether a limited hindsight bias can actually help the prosecutorial system to perform under uncertainty or whether it floods the legal system with unsubstantiated lawsuits.

## 7. DISCUSSION

Our model clearly contributes to the functioning of plea bargaining as a screening device. In contrast to the exogenous verification models (REINGANUM 1988, BAKER and MEZZETTI 2001), the prosecutor in our game is still plagued by uncertainty inside the courtroom, as he still does not know the defendant's true type. Even though uncertainty clearly degrades his performance, we show that the semiseparating equilibrium is still feasible as long as the ex-ante probability of guilt exceeds a certain threshold. Distinct from BJERK (2007), we also find a sizeable self-selection of the guilty defendants, and the negotiated sentence is not reduced whenever self-selection increases. Self-selection increases with the ex-ante probability. Interpreting this probability as the quality of the police groundwork, our results also emphasize the role of the police force in public enforcement: it is the quality of their groundwork that enables the separation in the first place, and that shapes the amount of possible self-selection. In a similar manner, one can also conclude that routine cases should

lead to more successful plea bargains, as the ex-ante probability is high, while complicated cases with ambiguous investigations lead to considerably less self-selection. This fits to the observation that prosecutors manage to settle plain cases to free resources for more difficult ones (see, e.g., BURKE 2007).

The hindsight bias has previously been studied in other asymmetric information settings, for example, in investment theory (BIAIS and WEBER 2009). To our knowledge, we offer the first formal integration of hindsight bias into the plea bargaining literature. A surprisingly positive effect of hindsight bias, which is in line with some previous findings by psychologists and behavioral economists (see, e.g., ROESE and VOHS 2012, BECK 2017), is the increased confidence of the prosecutor in the strength of his case. As a consequence, the bias (wrongfully) mitigates the negative effect of uncertainty on his performance. This is particularly interesting for the case of plea bargaining, as this increased confidence of the prosecutor puts pressure on the guilty defendants and amplifies the socially favorable self-selection mechanism. The negative effect is, however, that a hindsight biased prosecutor overestimates the guilt of his adversaries, thus wrongfully convicts some innocent individuals and also litigates cases that produce de facto a net loss to society.

Our model allows for some general implications about the hindsight bias in plea bargaining. First, the hindsight bias is possibly beneficial for society if the ex-ante probability is higher than the threshold for litigation. In other words, the positive effect dominates if the overly confident prosecutor can rely on the qualitative groundwork of the police force, as this effectively limits the amount of potential type I errors. Second, we believe that the nature of the case further determines the net effect on social welfare. For cases that are usually ambiguous, that is, show limited evidence or a low ex-ante probability of a guilty suspect, the negative effect of excessive charges will clearly dominate the favorable self-selection. We thus assume that hindsight bias is particularly worrisome in typically difficult cases, such as murder, embezzlement or harassment. Third, our theoretical framework allows for some testable hypotheses for future research: (i) hindsight bias increases the number of accepted plea bargains, (ii) hindsight bias does, on average, not

affect the negotiated sentence, (iii) hindsight bias increases type I errors in adversarial courts.

## 8. CONCLUSION

In this paper, we present a plea bargaining model with endogenous courts in which (limitedly) rational prosecutors seek to convict the guilty defendants. In contrast to the previous literature, we consider the impact of information asymmetry on prosecutor performance inside the adversary court, and thus on his bargaining position. Applying a sequential litigation game where the charging prosecutor acts as the less informed first-mover, we show that the probability of successful conviction in court is contingent on the prosecutor's beliefs about the defendant's guilt, his fear to convict the innocent, and reputational concerns.

The effectiveness of plea bargaining crucially hinges on the anticipated prosecutor performance inside the court. Under uncertainty, the equilibrium strategies are dependent on the *ex-ante* probability that the suspect did actually commit the crime, which, for example, can be interpreted as the quality of the investigations by the police force. If the investigations are very ambiguous and the *ex-ante* probability is lower than a specific threshold, the prosecutor cannot credibly commit to go to trial and no self-selection is achieved. If the case is sufficiently clear, however, a semiseparating equilibrium is still obtained where some guilty defendants accept the plea offer and the remaining guilty and all innocent defendants go to trial. Self-selection of guilty defendants and, eventually, social welfare increase the higher the *ex-ante* probability of a guilty defendant or the more conclusive the available evidence. On the contrary, the more the prosecutor fears to convict an innocent individual in court, the less self-selection is induced and social welfare decreases. Similarly, reputational concerns turn the prosecutor more cautious in his decision to charge, and thus narrow the room for successful plea bargains.

The existence of hindsight bias can possibly explain why these theoretical caveats may be potentially less restrictive in actual criminal procedure. Under hindsight bias, decision-makers are overly susceptible to what they observe. As a consequence, economic

agents become (falsely) more confident under uncertainty and act bolder, even though decision errors increase. In the first formal model of a hindsight biased prosecutor, we find, surprisingly, ambiguous effects of this behavioral bias. The positive effect of hindsight bias is that the (falsely) increased confidence of the prosecutor puts pressure on the guilty defendants and thus amplifies the desirable self-selection. Compared to this welfare enhancing effect, the impact of increased court errors remains negligible as long as the ex-ante probability of a guilty defendant is sufficiently high. The negative effect of hindsight bias, however, is that the (falsely) increased confidence also makes the prosecutor press charges that a rational prosecutor never would consider. If the ex-ante probability falls further below the threshold for a rational charge, the additional social costs of such excessive charges then rapidly dominates the social gains from further self-selection, and social welfare deteriorates. We thus conclude that (minor) hindsight bias is potentially beneficial in plain proceedings, but likely produces additional social costs in rather ambiguous criminal cases, such as murder, embezzlement or harassment.

## REFERENCES

- [1] Adelstein, Richard. 1979. Informational Paradox and the Pricing of the Crime. *Journal of Criminal Law and Criminology* 70: 281-298.
- [2] Adelstein, Richard, and Thomas Miceli. 2001. Toward a Comparative Economics of Plea Bargaining. *European Journal of Law and Economics* 11: 47-67.
- [3] Alschuler, Albert. 1983. Implementing the Criminal Defendant's Right to Trial: Alternatives to the Plea Bargaining System. *University of Chicago Law Review* 50: 931-1050.
- [4] Altenhain, Karsten. 2010. Die gesetzliche Regelung der Verständigung im Strafverfahren – eine verweigerter Reform. *Juristenzeitung* 7: 327-337.
- [5] Baker, Scott, and Claudio Mezzetti. 2001. Prosecutorial Resources, Plea Bargaining, and the Decision to Go to Trial. *Journal of Law, Economics & Organization* 17: 149-167.
- [6] Bar-Gill, Oren and Omri Ben-Shahar. 2009. The Prisoners' (Plea Bargain) Dilemma. *Journal of Legal Analysis* 1: 737-773.
- [7] Beck, Hanno. 2017. *Behavioral Economics*. Wiesbaden: Springer.
- [8] Berdejo, Carlos, and Noam Yuchtman. 2013. Crime, Punishment, and Politics: An Analysis of Political Cycles in Criminal Sentencing. *Review of Economics and Statistics* 95: 741-756.

- [9] Biais, Bruno, and Martin Weber. 2009. Hindsight Bias, Risk Perception and Investment Performance. *Management Science* 55: 1018-1029.
- [10] Bibas, Stephanos. 2004. Plea Bargaining Outside the Shadow of Trial. *Harvard Law Review*, 2463-2547.
- [11] Bjerck, David. 2005. Making the Crime Fit the Penalty: The Role of Prosecutorial Discretion under Mandatory Minimum Sentencing. *Journal of Law and Economics* 48: 591-625.
- [12] Bjerck, David. 2007. Guilt shall not escape or innocence suffer? The limits of plea bargaining when defendant guilt is uncertain. *American Law and Economics Review* 9: 305-329.
- [13] Boari, Nicola, and Gianluca Fiorentini. 2000. An Economic Analysis of Plea Bargaining: The Incentives of the Parties in a Mixed Penal System. *International Review of Law and Economics* 21: 213-231.
- [14] Boylan, Richard, and Cheryl Long. 2005. Salaries, Plea Rates, and the Career Objectives of Federal Prosecutors. *Journal of Law and Economics* 48: 627-651.
- [15] Burke, Alafair. 2007. Prosecutorial Passion, Cognitive Bias, and Plea Bargaining. *Marquette Law Review* 91: 183-211.
- [16] Camerer, Colin, Loewenstein, George, and Martin Weber. 1989. The Curse of Knowledge in Economic Settings: An Experimental Analysis. *Journal of Political Economy* 97: 1232-1254.
- [17] Covey, Russel. 2009. Signaling and Plea Bargaining's Innocence Problem. *Wash. & Lee Law Review* 66: 73-130.
- [18] Delacote, Philippe, and Lydie Ancelot. 2009. Prosecutor and Lawyers in Plea Bargaining with Complete Information. *Economics Bulletin* 29: 1929-1936.
- [19] Easterbrook, Frank. 1983. Criminal Procedure as a Market System. *Journal of Legal Studies* 12: 289-332.
- [20] Easterbrook, Frank. 2013. Plea Bargaining is a Shadow Market. *Duquesne Law Review* 51: 551-558.
- [21] Ehrlich, Isaac. 1996. Crime, Punishment, and the Market for Offenses. *Journal of Economic Perspectives* 10: 43-67.
- [22] Entorf, Horst. 2011. Crime, Prosecutors, and the Certainty of Conviction. IZA Discussion Paper No. 5670.
- [23] Fischhoff, Baruch. 2003. Hindsight  $\neq$  Foresight: The Effect of Outcome Knowledge on Judgment under Uncertainty. *BMJ Quality and Safety* 12: 304-312.
- [24] Fudenberg, Drew, and Jean Tirole. 1999. *Game Theory*. Cambridge: MIT Press.

- [25] Garoupa, Nuno, and Frank Stephen. 2008. Why Plea Bargaining fails to Achieve Results in so many Criminal Justice Systems: A new Framework for Assessment. *Maastricht Journal of European and Comparative Law* 15: 323-358.
- [26] Garoupa, Nuno. 2012. 'The Economics of Prosecutors', in: A. Harel & K. Hylton (eds.), *Research Handbook on the Economics of Criminal Law*. Northampton: Edward Elgar Publishing.
- [27] Gordon, Sanford, and Gregory Huber. 2002. Citizen Oversight and the Electoral Incentives of Criminal Prosecutors. *American Journal of Political Science* 46: 334-351.
- [28] Grossmann, Gene, and Michael Katz. 1983. Plea Bargaining and Social Welfare. *American Economic Review* 73: 749-757.
- [29] Hylton, Keith, and Vikramaditya Khanna. 2007. A Public Choice Theory of Criminal Procedure. *Supreme Court Economic Review* 15: 61-118.
- [30] Kahneman, Daniel, and Amos Tversky. 1984. Choices, Values and Frames. *American Psychologist* 39: 342-350.
- [31] Kim, Jeong-Yoo. 2010. Credible Plea Bargaining. *European Journal of Law and Economics* 29: 279-293.
- [32] Kobayashi, Bruce, and John Lott. 1996. In Defense of Criminal Defense Expenditures and Plea Bargaining. *International Review of Law and Economics* 16: 397-415.
- [33] Kreps, David, and Robert Wilson. 1982. Reputation and Imperfect Information. *Journal of Economic Theory* 27: 253-279.
- [34] Landes, William. 1971. An Economic Analysis of the Courts. *Journal of Law and Economics* 14: 61-107.
- [35] Lewisch, Peter. 2000. 'Criminal Procedure', in: A. Marciano & G. Ramello (Eds.) *Encyclopedia of Law and Economics*. New York: Springer.
- [36] Malik, Arun. 1990. Avoidance, Screening and Optimum Enforcement. *Rand Journal of Economics* 21: 341-353.
- [37] McCannon, Bryan. 2013. Prosecutor Elections, Mistakes and Appeals. *Journal of Empirical Legal Studies* 10: 696-714.
- [38] Miceli, Thomas. 1996. Plea Bargaining and Deterrence: An Institutional Approach. *European Journal of Law and Economics* 3: 249-264.
- [39] Priest, George, and Benjamin Klein. 1984. The Selection of Disputes for Litigation. *Journal of Legal Studies* 13: 1-55.

- [40] Ramseyer, Mark, and Eric Rasmusen. 2001. Why is the Japanese Conviction Rate so High? *Journal of Legal Studies* 30: 53-88.
- [41] Reinganum, Jennifer. 1988. Plea Bargaining and Prosecutorial Discretion. *American Economic Review* 78: 713-728.
- [42] Roese, Neal, and Kathleen Vohs. 2012. Hindsight Bias. *Perspectives on Psychological Science* 7: 411-426.
- [43] Scott, Robert, and William Stuntz. 1992. Plea Bargaining as Contract. *Yale Law Review* 101: 1909-1968.
- [44] Shavell, Steven. 1982. The Social Versus the Private Incentive to Bring Suit in a Costly Legal System. *Journal of Legal Studies* 11: 333-339.
- [45] Tullock, Gordon. 1975. On the Efficient Organization of Trials. *Kyklos* 28: 745-762.
- [46] Van Aaken, Anne, Feld, Lars, and Stefan Voigt. 2010. Do Independent Prosecutors Deter Political Corruption? An Empirical Evaluation across Seventy-eight Countries. *American Law and Economics Review* 12: 204-244.

## **Bisherige Research Papers:**

Nr. 1 Februar 2017

Berlemann, Michael; Christmann, Robin.

**The Role of Precedents on Court Delay.**

**Evidence from a Civil Law Country.**

Nr. 2 September 2017

Matthes, Roland.

**A note on the Saito-Kurokawa lift for Hermitian forms.**

Nr. 3 Februar 2018

Christmann, Robin.

**Prosecution and Conviction under Hindsight Bias in**

**Adversary Legal Systems.**



Leibniz-Fachhochschule  
Expo Plaza 11  
30539 Hannover

[www.leibniz-fh.de](http://www.leibniz-fh.de)