

Reducing False Guilty Pleas and Wrongful Convictions through Exoneree Compensation

Murat C. Mungan *Florida State University*

Jonathan Klick *University of Pennsylvania*

Abstract

A great concern with plea bargains is that they may induce innocent individuals to plead guilty to crimes they have not committed. In this article, we identify schemes that reduce the number of innocent pleas without affecting guilty individuals' plea-bargaining incentives. Large compensations for exonerees reduce expected costs associated with wrongful determinations of guilt in trial and thereby reduce the number of innocent pleas. Any distortion in guilty individuals' incentives to take plea bargains caused by these compensations can be offset by a small increase in the discounts offered for pleading guilty. Although there are many statutory-reform proposals for increasing exoneree compensation, no one has yet noted this desirable separating effect of compensations. We argue that such reforms are likely to achieve this result without causing losses in deterrence.

1. Introduction

One of the objectives of the criminal justice system, famously captured in Blackstone's ratio (Blackstone 1765, p. 358), is to minimize wrongful convictions.¹ Although there is no consensus on the rate at which the criminal justice system convicts the innocent,² the advent of postconviction DNA testing provides con-

We are indebted to Giray Ökten and Sam Wiseman for their very valuable contributions. We thank Andrew Daughety, John MacDonald, Alan Miller, Katrina Miller, Jennifer Reinganum, Chris Sanchirico, and Hans-Bernd Schäfer and the participants at the 2015 American Law and Economics Association annual conference for their suggestions and Andrew Butler for research assistance.

¹ See, for example, *In re Winship*, 397 U.S. 358, 364 (1970): "It is critical that the moral force of the criminal law not be diluted by a standard of proof [or a procedure for conviction] that leaves people in doubt whether innocent men are being condemned. It is also important in our free society that every individual going about his ordinary affairs have confidence that his government cannot adjudge him guilty of a criminal offense without convincing a proper fact finder of his guilt with utmost certainty."

² For instance, Gross (2008) summarizes the findings of previous studies and states that "there are two estimates of the false conviction rate for death sentences from 1973 through 1989, and they

[*Journal of Law and Economics*, vol. 59 (February 2016)]

© 2016 by The University of Chicago. All rights reserved. 0022-2186/2016/5901-0006\$10.00

clusive proof that our system does indeed convict the innocent—by both trial and plea—in nontrivial numbers (Gross 2008).³

One obvious, although partial, way to remedy this problem is by offering victims of wrongful conviction postexoneration compensation. Despite this, at the time this article is being written, 20 states do not have exoneration statutes.⁴ Moreover, existing statutes are not uniform,⁵ and many are deemed problematic. “Common shortcomings in existing legislation” include the limitation of compensation through private compensation bills only, the noncompensation of those who are deemed to have contributed to their convictions, and the noncompensation of individuals with felonies unrelated to the alleged offenses for which they were wrongfully convicted.⁶

This situation is troubling for public policy organizations (such as the Innocence Project) and academic scholars (see, for example, Armbrust 2004; Mandery et al. 2013, p. 562n.51), who have defended exoneree compensation on fairness grounds. The behavioral and efficiency gains from exoneree compensation have also been explored to some extent. For instance, the positive potential general (Fon and Schäfer 2007) and specific (Mandery et al. 2013) deterrence effects of exoneree compensation have been pointed out in the existing law and economics literature. Unlike previous work, here we focus on how exoneree compensation can be used to design mechanisms that achieve better separation of innocent and guilty individuals in the criminal justice system through its effect on a seemingly unrelated mechanism—namely, plea bargaining. As we demonstrate, our mechanism reduces the occurrence of wrongful convictions (or, equivalently, type I errors) by incentivizing innocent defendants to refuse plea bargains.

Plea bargaining is widely believed to contribute to the generation of type I errors (see, for example, Bowers 2008, p. 1119n.1). The Innocence Project has identified 29 innocent individuals who pleaded guilty to various crimes and was able to document nicely the reasons why they did so (Innocence Project 2009). This project and academic studies reveal that innocent defendants may plead guilty because, inter alia, they fear even greater punishment after conviction at trial, they are confused about their own conduct or the nature of the charges against them (perhaps because of bad legal advice), the evidence against them appears overwhelming, or they are under pressure from their attorney (perhaps motivated by

range from 2.3% to 5%.” Prior to this, Justice Antonin Scalia had famously greatly underestimated the false conviction rate of the American criminal justice system as .027 percent in *Kansas v. Marsh*, 126 S. Ct. 2516 (2006).

³ See also Innocence Project (2009) for a list of individuals who were wrongfully convicted through pleas.

⁴ These states are Alaska, Arizona, Arkansas, Delaware, Georgia, Hawaii, Idaho, Indiana, Kansas, Kentucky, Michigan, Nevada, New Mexico, North Dakota, Oregon, Pennsylvania, Rhode Island, South Carolina, South Dakota, and Wyoming. The remaining 30 states, the District of Columbia, and the federal government have compensation statutes. See Innocence Project, *Compensating the Wrongly Convicted* (<http://www.innocenceproject.org/compensating-wrongly-convicted/>).

⁵ See, for example, Norris (2012, p. 352), who reviews existing state exoneree compensation statutes and finds that “the assistance offered varies tremendously from state to state.”

⁶ Innocence Project, *Compensating the Wrongly Convicted* (<http://www.innocenceproject.org/compensating-wrongly-convicted/>).

the attorney's desire to avoid the personal costs of trial). A number of scholars, therefore, have sought ways to encourage the innocent to go to trial and, possibly, be acquitted (see, for example, Gazal-Ayal 2006; Bar-Gill and Gazal-Ayal 2006). As we demonstrate, exoneree-compensation schemes can also be used to reduce the number of false convictions resulting from plea bargaining.

The mechanisms that we study rely on the probability of exoneration being greater for innocent defendants than guilty ones. Because of this probability differential, increasing the compensation to defendants who went to trial increases the expected payoff associated with the trial option by a larger amount for innocent than for guilty defendants. Hence, the number of innocent defendants taking plea bargains can be reduced by increasing the amount of compensation available to defendants who go to trial without causing much of a distortion in the incentives of guilty individuals.

Distortions in incentives may exist and cause a small increase in the number of guilty individuals refusing pleas only if false exoneration is probable. Such an increase can be overturned by a small increase in the discount offered through plea bargains. Hence, by simultaneously increasing the exoneree compensation and the discounts offered through plea bargains, one can keep guilty defendants' plea-bargaining incentives constant. The same increases, however, incentivize innocent individuals to refuse plea bargains relatively more often, because the increase in the value of going to trial is greater than the increase in the value of taking plea bargains for innocent defendants.

Our proposal to increase posttrial exoneree compensation should not be understood as being incompatible with the existence of post-plea-bargaining compensation for exonerees. Our main focus here is the posttrial compensation of exonerees and its increased separation effect, which induces a lower rate of type I errors. This does not mean that there are not other, fairness-related, gains that can be achieved by compensating those who have pleaded guilty to crimes they did not commit. Although we abstract from the issue of postplea exoneration in Section 2, in Section 3 we suggest that considerations related to postplea exoneration are unlikely to disrupt the functioning of the mechanism that we propose because, *inter alia*, the probability of postplea exoneration is very low, and, in addition, the mechanism requires only that compensation be greater after trial.

Our proposed mechanism adds to the two strands of the law and economics literature on plea bargaining and exoneree compensation. It complements previously proposed mechanisms to minimize wrongful convictions resulting from plea bargaining (Gazal-Ayal 2006; Bar-Gill and Gazal-Ayal 2006) and is an addition to the sparse literature discussing the beneficial effects of exoneree compensation (Fon and Schäfer 2007; Mandery et al. 2013). After we propose a plea-bargaining model in Section 2 to formalize our proposed mechanism, we discuss in Section 3 how our model interacts with some of the previous law and economics literature on plea bargaining and exoneree compensation. Section 3 also discusses potential effects of our mechanism on deterrence. We conclude in Section 4. The Appendix contains proofs of the lemmas and propositions.

2. Model

We assume that the government possesses an imperfect detection mechanism that leads to the prosecution of guilty (G) as well as innocent (I) individuals. To capture unobservable heterogeneities in defendants' attitudes toward pleading guilty versus going to trial, we assume that prosecuted individuals differ in their attitudes toward risk, σ ,⁷ and that their utility is represented by $u = M_\sigma$, where M represents the monetary equivalent of an individual's state⁸ and $\sigma \in (0, \infty)$.

Individuals are initially endowed with wealth w .⁹ The prosecution offers each individual a plea bargain that discounts the sanction (s) for the underlying offense by δs .¹⁰ Hence, the utility of a defendant who accepts a plea bargain is $[w - (1 - \delta)s]\sigma$. We assume that δ induces at least some very risk-averse innocent defendants to take the bargain option, since otherwise all defendants who take bargains are guilty, and therefore there is no type I error to be mitigated.¹¹

If an individual rejects the plea bargain, one of three things happens to him: he may be acquitted, convicted and later exonerated,¹² or convicted and never exonerated. The probabilities with which the defendant reaches these outcomes depend on whether he is innocent or guilty. In particular, an innocent individual is convicted with probability α_1 , whereas a guilty individual is convicted with probability $(1 - \alpha_2)$. Thus, α_1 and α_2 , respectively, denote the probability of wrongful

⁷ Absent this type of heterogeneity, one can completely separate out the guilty from the innocent at the plea-bargaining stage. Similar assumptions regarding differential risk attitudes can be found in the literature on plea bargaining starting with Grossman and Katz (1983) and continuing with Givati (2014). Incidentally, we contribute to the existing plea-bargaining literature by presenting a setup in which some innocent individuals accept plea bargains. Although this conforms to reality, virtually all of the existing models produce the contrary result. See Daughety and Reinganum (2015, p. 22n.36), which suggests that the only exception that the authors were able to locate is Reinganum (1988).

⁸ One may naturally wonder about the relationship between the monetary equivalent of a sanction and the actual sanction (when the sanction is nonmonetary). Block and Lind (1975) formalize this relationship by constructing a utility function (U) with two arguments ($W =$ wealth and $S =$ non-monetary sanction). They define $f(W, S)$ as the monetary equivalent of a sanction for a person with wealth W as follows: let $U(W^0, 0) = U(W, S)$; then $f(W, S) = W - W^0$. Hence, one can rewrite $U(W, S)$ as $U[W - f(W, S), 0]$. The sanction that we consider below, s , can be interpreted as the monetary equivalent of the nonmonetary sanction (that is, f). We relegate a brief discussion of an issue related to the dependency of f on W to note 13.

⁹ We assume that wealth is constant among individuals and capture heterogeneities through variations in their attitudes toward risk. A similar approach would be to assume variation in wealth along with a utility function that exhibits decreasing absolute risk aversion, which would generate the result that richer defendants are more likely to refuse plea bargains. Although our results would be preserved in this setup, we choose to focus on attitudes toward risk because they are unobservable, and therefore it is more difficult for prosecutors to discriminate between defendants on the basis of this source of heterogeneity.

¹⁰ We assume, for simplicity, that the discount rate is fixed. Legal scholars note that this assumption often holds. Bibas (2004, p. 2487), for instance, reviews existing practices and concludes "that many defendants reap the same, crude discount regardless of fine differences in guilt and proof."

¹¹ Formally, we assume that δ is such that expression (1) does not (and therefore expression [2] does not) hold around $\sigma = 0$. As can be inferred from the proof of lemma 2, this condition corresponds to $\delta > [w^{(1-\alpha_1)}(w + \psi - s)^{\alpha_1 \rho_1} (w - s)^{\alpha_1(1-\rho_1)} + s - w]/s$.

¹² Throughout this section, we assume that exonerated can take place only after the defendant serves his full sentence. This is only for expositional purposes, and in Section 3.1 we show that our results extend to the case in which a convict can be exonerated before he fully serves his sentence.

conviction (type I errors) and the probability of false acquittal (type II errors) through trial, and $1 > \alpha_1 + \alpha_2$. Exoneration (conditional on conviction) happens with a probability of ρ_1 if the person is innocent and a probability of $1 - \rho_2$ if he is guilty, and $1 < 1 - \rho_1 + \rho_2$. If he is exonerated, the government offers him an exoneree compensation of $\psi \leq s$.

Hence, an innocent individual refuses a plea bargain if

$$I(\sigma, \psi) \equiv \sqrt[\sigma]{(1 - \alpha_1)w^\sigma + \alpha_1\rho_1(w + \psi - s)^\sigma + \alpha_1(1 - \rho_1)(w - s)^\sigma} > w - (1 - \delta)s. \tag{1}$$

And guilty individuals refuse plea bargains if

$$G(\sigma, \psi) \equiv \sqrt[\sigma]{\alpha_2w^\sigma + (1 - \alpha_2)(1 - \rho_2)(w + \psi - s)^\sigma + (1 - \alpha_2)(\rho_2)(w - s)^\sigma} > w - (1 - \delta)s. \tag{2}$$

We express individuals' constraints with reference to I and G as in expressions (1) and (2),¹³ because it allows us to separate out the incentive effects of plea-bargain discounts, exoneree compensations, and attitudes toward risk. In particular, I and G capture the desirability of the trial option as a function of the attitude toward risk and exoneree compensation, whereas the right-hand side of each expression— $w - (1 - \delta)s$ —captures the value of plea bargains to defendants. The former is affected only by the level of exoneree compensation and attitudes toward risk, whereas the latter is affected only by δ . As expected, risk-loving defendants find the trial option more attractive than risk-averse defendants.

Lemma 1. The following inequalities hold: $I_\sigma > 0$, $G_\sigma > 0$.

Lemma 1 formalizes the intuitive result that a risk-averse person is more willing than a risk-seeking person to plead guilty. The next lemma builds on this observation.

Lemma 2.

i) For any $s > 0$, there exist $\sigma_I(\psi, \delta)$ and $\sigma_G(\psi, \delta)$ such that innocent (guilty) individuals take plea bargains if and only if their attitudes toward risk are such that $\sigma < \sigma_I$ ($\sigma < \sigma_G$).¹⁴

ii) Moreover, the threshold attitude toward risk for guilty individuals is greater; that is, $\sigma_G > \sigma_I$.

Lemma 2 exploits the observation that innocent individuals are more willing to go to trial because their probability of being convicted is lower. This implies that

¹³The second terms inside the roots of the expressions for I and G reveal the implicit assumption that s , the monetary equivalent of the sanction, does not depend on wealth. This need not be true in the general case, since the cross derivative of the monetary equivalent $f(W, S)$, discussed in note 8, need not be 0 (that is, $f_{SW} \neq 0$ is possible). This possibility can be accounted for by letting s in expressions (1) and (2) be context specific. Repeating our analysis with context-specific s terms yields the same results. Therefore, we use a single s to simplify the notation and the derivation of results.

¹⁴We assume that indifferent individuals refuse the plea offer.

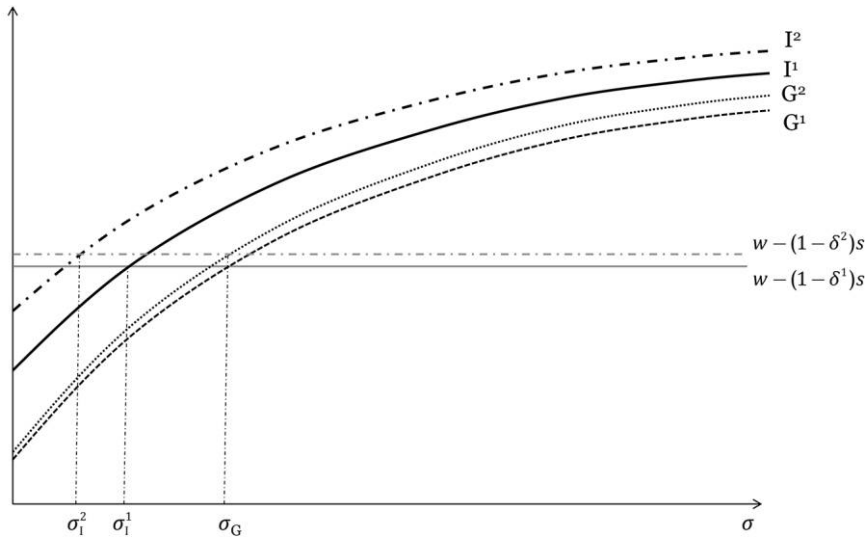


Figure 1. Simultaneous increases in ψ and δ that keep σ_G unchanged

innocent individuals’ threshold attitude toward risk for refusing plea bargains is lower than the corresponding threshold for guilty individuals. Proposition 1 relies on this result to identify exoneration schemes that incentivize innocent individuals to refuse plea bargains without affecting guilty individuals’ incentives. For purposes of propositions 1 and 2, it is assumed that the number of guilty and innocent individuals being prosecuted is fixed. The effects of compensation and plea-bargaining policies on deterrence are considered separately in Section 3.

Proposition 1. If the probability of false exoneration is sufficiently small,¹⁵ then for all regimes in which $\psi = 0$ and $\delta = \delta_N > 0$ —that is, no exoneree compensation is offered—one can construct a continuum of new regimes of the form $[\delta^*(\psi), \psi]$, where $\delta^*(\psi) > \delta_N$ for all $\psi > 0$ in which fewer innocent individuals take plea bargains—that is, $\sigma_1[\psi, \delta^*(\psi)] < \sigma_1(0, \delta_N)$ —and the number of guilty individuals taking plea bargains is unchanged—that is, $\sigma_G[\psi, \delta^*(\psi)] = \sigma_G(0, \delta_N)$.

The intuition behind proposition 1 can be illustrated with reference to Figure 1, which represents the relative returns for guilty and innocent individuals from going to trial (I and G as defined in inequalities [1] and [2]) and the return from pleading guilty ($w - (1 - \delta)s$). Figure 1 illustrates two schemes, denoted 1 and 2, in which the second scheme is generated through simultaneous increases in δ and ψ .

An increase in the amount of exoneree compensation has a disproportionately

¹⁵ The probability of false exoneration is sufficiently small if it is smaller than the upper bound expressed in condition (A20) in the Appendix.

larger impact on the expected return from going to trial for innocent individuals. In particular, because the probability of postconviction exoneration for innocent individuals (ρ_1) is much larger than that for guilty individuals ($1 - \rho_2$), increasing the amount of exoneree compensation reduces the requisite risk tolerance to refuse a plea by a much greater amount for innocent individuals than for guilty individuals. This observation is reflected by a greater shift in I versus G in Figure 1. Moreover, any increase in the number of guilty individuals refusing pleas that is caused by greater exoneree compensation can be offset by a small increase in the discount offered through plea bargains. This is reflected in Figure 1 by a shift in the return from the plea option that keeps σ_G constant. Therefore, simultaneous increases in exoneree compensations and plea discounts keep guilty defendants' plea-bargaining incentives unchanged but induce more innocent individuals to refuse pleas—that is, σ_G remains constant while σ_I is reduced.

The conditions that imply this dynamic become more intuitive on further investigation. In particular, expression (A19) in the proof of proposition 1 (see the Appendix) reveals that these conditions can be expressed as $\pi[\alpha_1/(1 - \alpha_2)] > (1 - \rho_2)/\rho_1$, where $\pi(\psi) = \{w - [1 - \delta^*(\psi)]s\}/(w + \psi - s)^{\sigma_{GN} - \sigma_{IC}}$ can be interpreted as a discount or inflation factor. This condition basically requires the false discovery rate (FDR) in trials (that is, wrongful convictions over convictions) discounted or inflated by π to be greater than the FDR in exoneration proceedings (false exoneration over exoneration).¹⁶ Whether π is a discount or inflation factor, and how large or small it is, depends on ψ and $\delta_N = \delta^*(0)$. In particular, $\pi(0) = \{w - (1 - \delta_N)s\}/(w - s)^{\sigma_{GN} - \sigma_{IC}} > 1$, since, per lemma 2, $\sigma_{GN} - \sigma_{IC} > 0$. Thus, for small exoneree compensations to cause the dynamics described above, a sufficient condition is that the FDR in trials be at least as large as the FDR in exoneration proceedings.

For larger values of ψ , the condition is not as intuitive, since $\pi(s) = \{w - [1 - \delta(s)]s\}/w)^{\sigma_{GN} - \sigma_{IC}} < 1$. Results that are easier to interpret emerge when we impose the requirement that all individuals who accept pleas are risk averse. An investigation of expressions (1) and (2) reveals that this is equivalent to $\delta_N < \alpha_2$. In this case, $\sigma_{GN} - \sigma_{IC} < 1$, and therefore a sufficient condition is that $\underline{\pi}[\alpha_1/(1 - \alpha_2)] > (1 - \rho_2)/\rho_1$, where $\underline{\pi} \equiv 1 - [(1 - \delta_N)s/w]$. The term $(1 - \delta_N)s/w$ is the ratio of the plea sanction to wealth, and $\underline{\pi}$ is a discount rate that is decreasing in this ratio. When $(1 - \delta_N)s/w$ is small, $\underline{\pi}$ is close to 1, which implies that as long as the FDR at trial is somewhat larger than the FDR at exoneration proceed-

¹⁶ The expression discussed here contains ratios of the form false positives over total positives, where positives are what the procedures aim to prove—namely, conviction in trials and exoneration in exoneration proceedings. These ratios are different than false discovery rates (FDRs) but can easily be converted into them. In particular, if (false positives)/(total positives) = $a/b = x$, then the false discovery rate is $a/(a + b) = x/(x + 1)$. Hence, the corresponding FDR condition is $[\text{FDR}(\text{trial})]/[\text{FDR}(\text{exoneration})] > 1/[\pi(\psi)]\{[(1 - \rho_2)/\rho_1] + 1\}/\{[a_1/(1 - a_2)] + 1\} \equiv F(\psi, a_1, a_2, \rho_1, \rho_2)$. It follows that when $1/\pi$ is greater than, but close to, 1, Φ is even closer to 1, since $(1 - \rho_2)/\rho_1 < a_1/(1 - a_2)$. Hence, $1/\pi$ is a conservative estimate—one that does not rely on the specific α terms and ρ terms—for how large $[\text{FDR}(\text{trial})]/[\text{FDR}(\text{exoneration})]$ must be. Note 17 provides a simple example that highlights this feature of π .

ings,¹⁷ increases in ψ , regardless of how large, can be used to reduce the number of innocent individuals pleading guilty without affecting the incentives of guilty individuals.

The next proposition builds on these observations and identifies the sanction scheme that minimizes the number of guilty pleas by innocent individuals, assuming that the FDR conditions discussed above hold.

Proposition 2. When the probability of incorrect exoneration is sufficiently small,¹⁸ among regimes that keep guilty individuals' plea-bargaining incentives unchanged, that which offers maximal compensation—namely, $\psi = s$ and $\delta = \delta^*(s)$ —maximizes the expected utilities of all defendants and minimizes the number of innocent individuals who plead guilty and, therefore, the number of wrongful convictions.

Proposition 2 formalizes the main function of the mechanism that we propose—namely, the reduction of wrongful convictions through the use of exoneree compensation. Because the number of wrongful convictions is declining in the amount of compensation, offering the highest exoneree compensation possible minimizes the number of wrongful convictions. One may question whether this comes at the cost of reduced deterrence. In Section 3 we argue that our mechanism affects deterrence most likely in a positive manner, if at all.

3. Discussion and Extensions

3.1. Early Exoneration

In our analysis thus far, we have assumed that exoneration, if it takes place, occurs only after the defendant incurs the entirety of his sanction. But in practice, a convicted person may be exonerated before he serves his full sentence. This possibility can be incorporated into our analysis by defining a probability distribution over the states of the world that lead to exoneration at different points in time while the convict is serving his sentence. Here we demonstrate that this consideration does not meaningfully affect our results.

To formalize the possibility of early exoneration, let $z_{j \in \{1, G\}}^i = z(\gamma^i)$ denote the probability with which a person will be exonerated after suffering a proportion γ_i of the monetary equivalent of his sanction. To make this case compatible with our previous analysis, we assume without loss of generality that $z_j^i > 0$ for $i \in \{1, 2, \dots, n\}$ —that is, that γ is a discrete random variable with $\gamma_i \in (0, 1]$ for $i \in \{1, 2, \dots, n\}$. In this case, expressions (1) and (2) become

¹⁷ If, for instance, the plea discount is $\frac{1}{3}$, then $s/w = 3/10$, $\alpha_1 = .05$, $1 - \alpha_2 = .40$, $\rho_1 = .1$, and $1 - \rho_2 = .01$. Then $\pi = .8$, which means that it is sufficient for $\alpha_1/(1 - \alpha_2)$ to be 1.25 times $(1 - \rho_2)/\rho_1$. The corresponding minimum FDR rate—that is, Φ as defined in note 16—is 1.25(1.1/1.125), which is approximately 1.22.

¹⁸ The probability of false exoneration is sufficiently small if it is smaller than the upper bound expressed in condition (A20).

$$I(\sigma, \Psi) \equiv \sqrt[\sigma]{(1 - \alpha_1)w^\sigma + \alpha_1 \sum z_1^i (w + \psi_i - \gamma_i s)^\sigma + \alpha_1 (1 - \rho_1)(w - s)^\sigma} > w - (1 - \delta)s \tag{3}$$

and

$$G(\sigma, \psi) \equiv \sqrt[\sigma]{\alpha_2 w^\sigma + (1 - \alpha_2) \sum z_G^i (w + \psi_i - \gamma_i s)^\sigma + (1 - \alpha_2)(\rho_2)(w - s)^\sigma} > w - (1 - \delta)s, \tag{4}$$

where $\sum_{i=1}^n z_j^i = \rho_1$ if $j = I$, $\sum_{i=1}^n z_j^i = 1 - \rho_2$ if $j = G$, and $\psi = (\psi_1, \psi_2, \dots, \psi_n)$ is the menu of compensations chosen by the government that are receivable by exonerees released after serving part of their sentences. Using these expressions and following steps almost identical to those in the proofs of lemmas 1 and 2 reveals that the intuitive properties described by these two lemmas are preserved. Similarly, we can repeat the steps in the proof of proposition 1 by replacing expression (2) with expression (4). In particular, the effect of increasing ψ_i for any $i \in \{1, \dots, n\}$ on the threshold attitude toward risk for innocent defendants is now given by

$$\frac{d\sigma_{I_C}}{d\psi_i} = -\frac{I_{\psi_i} - s\delta_{\psi_i}^*}{I_\sigma} < 0, \tag{5}$$

which holds when

$$I_{\psi_i}(\sigma_{I_C}, \psi) - G_{\psi_i}(\sigma_{G_N}, \psi) > 0, \tag{6}$$

because $I_\sigma > 0$ (as proven in lemma 1) and $s\delta_{\psi_i}^* = G_{\psi_i}$. Differentiating I and G with respect to ψ_i and substituting their derivatives reveals that expression (6) is equivalent to

$$\left[\frac{w - (1 - \delta_N)s}{w} \right]^{\sigma_{G_N} - \sigma_{I_C}} \frac{\alpha_1}{1 - \alpha_2} z_1^i > z_G^i. \tag{7}$$

Hence, increasing all possible exoneree compensations leads to better separation of guilty and innocent individuals, if a slightly modified version of condition (A20) holds: the probability of erroneously exonerating a defendant after he serves γ_i of his sentence (that is, z_G^i) must be sufficiently small for all i .

3.2. Deterrence Effects

In Section 3.1, we demonstrated how the number of wrongful convictions can be reduced by simultaneously increasing the amount of exoneree compensation and the discount for plea bargaining. However, as stated in proposition 2, this scheme increases the expected utility of guilty defendants. Initially it may seem that this would significantly reduce deterrence, since the expected cost associated with the commission of a crime is reduced. This conjecture is incorrect for three reasons.

First, as discussed in the literature, the probability of false exoneration is most

likely very low. As Gross (2008, p. 175, citing Gross et al. 2005) states,¹⁹ “Some exonerated defendants are no doubt guilty of the crimes for which they were convicted, in whole or in part, but the number is likely very small. It is extremely difficult to obtain this sort of relief after a criminal conviction in America, and it usually takes overwhelming evidence.” Therefore, any increase in guilty individuals’ expected utility from going to trial is quite small. This also implies that the simultaneous increase in the plea discount necessary to keep guilty individuals’ plea incentives constant is minimal. Hence, the reduction in the expected cost associated with committing crime is most likely very small.

Second, as Fon and Schäfer (2007) demonstrate, exoneree compensation can increase deterrence by reducing the cost of refraining from crime and thereby mitigating the negative effect of type I errors on deterrence. Although recent research demonstrates that the effect of type I errors on deterrence is not entirely clear,²⁰ Fon and Schäfer’s (2007) main point remains valid: any deterrence effect that is due to type I error is mitigated by exoneree compensation.

Third, as we discuss below, large exoneree compensations are likely to incentivize prosecutors to channel their efforts toward prosecuting strong cases.²¹ This is likely to cause more prosecutions of guilty individuals and fewer prosecutions of innocent individuals. Hence, deterrence is likely to be enhanced by an increase in the probability of conviction for guilty individuals.

These three observations reveal that the deterrent effect of our mechanism is far from being clearly negative. In fact, because the increases in the utility from crime are very small, our mechanism likely enhances deterrence through the two aforementioned reductions in type I and type II errors.

3.3. *The Selection-of-Cases Effect*

Bar-Gill and Gazal-Ayal (2006, p. 353) focus on the budget constraint of prosecutors to identify what they term the “selection of cases” effect of “restricting the permissible sentence reduction in a plea bargain.” They observe that if prosecutors cannot offer substantial discounts in return for a guilty plea, defendants in weak cases will not be willing to plea bargain. As a result, a prosecutor who is pressured by budget constraints will have an incentive not to charge defendants when the probability of conviction is low, and hence the defendant is unlikely to accept a plea-bargain offer. Therefore, prosecutors cannot make a credible threat of pursuing cases against defendants for whom the probability of being convicted at trial is low. Thus, prosecutors have to substitute potentially weak cases with

¹⁹ See also Fon and Schäfer (2007, p. 278n.21): “For convenience, we assume that only the wrongfully convicted has a chance of getting a retrial. This is justified on the grounds that usually a (successful) retrial is only possible if substantial new evidence in favor of the convicted shows up. This is unlikely if the criminal was rightfully convicted”; Gazal-Ayal and Tor (2012, p. 351): “To overcome the challenge of identifying truly innocent defendants, we turned to the one group whose innocence is nearly certain: defendants who were wrongfully convicted and later exonerated.”

²⁰ See Lando (2006), Garoupa and Rizzolli (2012), and Lando and Mungai (2014) for debate on the effects of type I errors on deterrence.

²¹ As we discuss below, this point is due to Bar-Gill and Gazal-Ayal (2006).

strong cases ex ante to better administer their resources and to secure a high success rate.

This selection-of-cases effect identified by Bar-Gill and Gazal-Ayal (2006) is generated by exoneree compensation as well. The primary effect of exoneree compensation is to reduce the expected cost of going to trial, relative to taking plea bargains, for the innocent and to thereby increase the number of innocent individuals who are willing to go to trial when prosecuted. Prosecutors who know that they cannot make a credible threat against such defendants are unlikely to bring cases against them in the first place.

These selection effects are likely to increase the value of postconviction exoneree compensation further through two interrelated channels. First, because fewer innocent individuals will be prosecuted, the number of type I errors is likely to be reduced. Second, deterrence will be increased; because prosecutors will be devoting more resources to the prosecution of the truly guilty, they will thereby be increasing the sensitivity, in a statistical sense, of the law enforcement system.

3.4. *Postplea Exoneree Compensation*

In Section 2, we demonstrated how one can achieve better separation of innocent and guilty individuals in the plea-bargaining stage by increasing the posttrial compensation made available to exonerees, but we did not consider post-plea-bargaining exoneree compensation. We would like to clarify that we are not suggesting that postplea compensation may not have a function or a value. In general, such compensation may have fairness-related benefits and, perhaps, some administrative costs resulting from an increased number of exoneration proceedings initiated by guilty individuals. Our objective in this section is not to debate the normative desirability of postplea exoneree compensation but to demonstrate that the utility of our mechanism is unaffected by considerations related to such compensation.

First, and perhaps most important, our mechanism is independent of the availability of the existence of postplea compensation. The two types of compensation are conceptually separable, as demonstrated by the split among the 30 states that have some form of exoneration statute; currently, postplea compensation is either completely²² or partially²³ unavailable in the District of Columbia and 12 of the 30 states²⁴ with exoneration statutes, whereas the remaining states do not place conditions based on whether the exoneree has pleaded guilty. Hence, independent of the level of postplea compensation available, posttrial compensation

²² For example, Iowa Code Ann., sec. 663A(6)–(8) (West 2014); N.J. Stat. Ann., secs. 52:4C-1–7 (West 2013); Okla. Stat. tit. 51, sec. 154 (2014); and Ohio Rev. Code Ann., secs. 2305.02 and 2743.48 (LexisNexis 2014), require that the person did not plead guilty.

²³ For instance, the Nebraska statute (Neb. Rev. Stat., sec. 29-4601–4608) hinders compensation for those who have pleaded guilty unless the guilty plea was coerced by law enforcement officers.

²⁴ These states are California, Iowa, Massachusetts, Nebraska, New Jersey, New York, Ohio, Oklahoma, Virginia, Washington, West Virginia, and Wisconsin (2014 Compensation Chart, Innocence Project internal document, available from the authors on request).

can be increased simultaneously with plea-bargain discounts, as mentioned in Section 2, to achieve better separation of innocent and guilty individuals.

Despite the conceptual separability of postplea and posttrial compensation mechanisms, one may wonder whether simultaneous increases in the amount made available through the two forms of compensation are likely to induce the type of separation discussed in Section 2. The answer is definitely yes, if one is willing to increase posttrial compensation by more than postplea compensation. To achieve the type of separation considered in our model, all that needs to be achieved is an increase in the innocent individual's expected utility of going to trial that is greater than the increase in his expected utility of pleading guilty. This can easily be achieved by increasing the posttrial compensation by more than the postplea compensation for exonerees.

Finally, even if one focuses only on compensation schemes that offer equal amounts of compensation to those who have pleaded guilty and those who were wrongfully convicted through trial, empirical considerations suggest that increases in compensation will have a larger effect on the expected utility of the trial option for innocent individuals. Three empirical considerations are relevant.

First, a person who has pleaded guilty in the past is likely to lose credibility in exoneration proceedings. Therefore, given heightened requirements in exoneration proceedings, the probability of a postplea exoneration is much smaller than the probability of a posttrial exoneration. As Simon (2012, pp. 227–28n.17) states, “Overturning a conviction is close to impossible for inmates who were convicted based upon their pleas.” Second, recent scholarship (Gazal-Ayal and Tor 2012; Tor, Gazal-Ayal, and Garcia 2010) demonstrates an innocence effect in plea bargaining: “Innocents are less willing to accept plea offers than guilty defendants” (Tor, Gazal-Ayal, and Garcia 2010, p. 97). One plausible cause for this effect is innocent individuals discounting the difference between the expected utilities associated with the plea option and being convicted at trial. This implies that an increase in the expected value of trial that is due to increased exoneree compensation is likely to be valued more than similar increases associated with the expected value of pleading guilty. Third, because individuals who have pleaded guilty have less to gain from being exonerated than do people convicted through trial, they are less likely to initiate exoneration proceedings.

These three considerations imply that an increase in posttrial compensation is likely to increase the expected utility of the trial option by more than similar increases in postplea exoneration would increase the expected utility associated with pleading guilty. Therefore, the separating function identified in our model can be obtained even in cases where the posttrial and postplea compensation are constrained to be equal.

4. Conclusion

Policy organizations and scholars, most notably the Innocence Project, urge states to adopt statutes that either enable exoneration or increase the amount of

compensation available to exonerees. Although there are many fairness-related effects associated with these proposed statutory reforms, the potential for increased exoneree compensations to reduce wrongful convictions has been unnoticed or overlooked. In this article, we have proposed a mechanism that relies on large exoneree compensations that reduce the number of wrongful convictions without affecting guilty individuals' incentives and are unlikely to affect deterrence. Accordingly, our analysis adds another item—a reduction in wrongful convictions—to the list of benefits that are associated with large exoneree compensations.

Appendix

Proofs of Lemmata and Propositions

Proof of Lemma 1

Let

$$k_1 \equiv (1 - \alpha_1), \quad k_2 \equiv \alpha_1 \rho_1, \quad k_3 \equiv \alpha_1 (1 - \rho_1),$$

$$K_1 \equiv w, \quad K_2 \equiv w + \psi - s, \quad K_3 \equiv w - s, \quad \text{and} \quad f(\sigma) \equiv \sum_{i \in \{1, 2, 3\}} k_i K_i^\sigma. \tag{A1}$$

Then it follows that

$$I = f(\sigma)^{1/\sigma}, \text{ and, therefore, } I_\sigma = -\frac{1}{\sigma^2} \ln [f(\sigma)] f(\sigma)^{1/\sigma} + \frac{1}{\sigma} f(\sigma)^{(1/\sigma)-1} f'(\sigma). \tag{A2}$$

Hence, $I_\sigma > 0$ if and only if

$$\sigma f'(\sigma) > f(\sigma) \ln [f(\sigma)]. \tag{A3}$$

Differentiating f and using the power and product rules of logarithms on both sides, we have

$$\sum_{i \in \{1, 2, 3\}} k_i \ln [(K_i^\sigma)^{K_i^\sigma}] > \ln [f(\sigma)^{f(\sigma)}]. \tag{A4}$$

Next, let $g(x) \equiv \ln(x^x)$. It follows from the definition of $g(x)$ and condition (A17) that $I_\sigma > 0$ if

$$\sum_{i \in \{1, 2, 3\}} k_i g(K_i^\sigma) > g \left(\sum_{i \in \{1, 2, 3\}} k_i K_i^\sigma \right). \tag{A5}$$

But $g(x)$ is convex, and therefore, because of Jensen's inequality, the above inequality holds. Hence, $I_\sigma > 0$. By following analogous steps, we can show that $G_\sigma > 0$ too. Q.E.D.

Proof of Lemma 2

i) Let

$$k_1 \equiv (1 - \alpha_1), \quad k_2 \equiv \alpha_1 \rho_1, \quad k_3 \equiv \alpha_1 (1 - \rho_1),$$

$$K_1 \equiv w, \quad K_2 \equiv w + \psi - s, \quad K_3 \equiv w - s, \quad \text{and} \quad f(\sigma) \equiv \sum_{i \in \{1, 2, 3\}} k_i K_i^\sigma.$$

Then, it follows from l'Hôpital's rule that

$$\lim_{\sigma \rightarrow \infty} [\ln(I)] = \lim_{\sigma \rightarrow \infty} \left\{ \frac{d \ln[f(\sigma)] / d\sigma}{d\sigma / d\sigma} \right\} = \lim_{\sigma \rightarrow \infty} \left\{ \frac{d \ln[f(\sigma)]}{d\sigma} \right\} = \lim_{\sigma \rightarrow \infty} \left[\frac{\sum_{i \in \{1, 2, 3\}} k_i K_i^\sigma \ln(K_i)}{f(\sigma)} \right]. \tag{A6}$$

Dividing the numerator and the denominator by K_1^σ yields

$$\lim_{\sigma \rightarrow \infty} [\ln(I)] = \lim_{\sigma \rightarrow \infty} \left[\frac{k_1 \ln(K_1)}{k_1 + k_2 (K_2/K_1)^\sigma + k_3 (K_3/K_1)^\sigma} + \frac{k_2 \ln(K_2) (K_2/K_1)^\sigma + k_3 \ln(K_3) (K_3/K_1)^\sigma}{k_1 + k_2 (K_2/K_1)^\sigma + k_3 (K_3/K_1)^\sigma} \right]. \tag{A7}$$

Hence, $\lim_{\sigma \rightarrow \infty} [\ln(I)] = \ln(K_1)$, which implies that $\lim_{\sigma \rightarrow \infty} I = K_1 = w$. The same steps can be used to show that $\lim_{\sigma \rightarrow \infty} G = K_1 = w$. Next, by manipulating equation (A6), we have that

$$\lim_{\sigma \rightarrow 0} [\ln(I)] = \lim_{\sigma \rightarrow 0} \left[\frac{\sum_{i \in \{1, 2, 3\}} k_i \ln(K_i)}{\sum_{i \in \{1, 2, 3\}} k_i} \right] = \lim_{\sigma \rightarrow 0} \left[\sum_{i \in \{1, 2, 3\}} k_i \ln(K_i) \right], \tag{A8}$$

since $\sum_{i \in \{1, 2, 3\}} k_i = 1$. Hence,

$$\lim_{\sigma \rightarrow 0} I = w^{(1-\alpha_1)} (w + \psi - s)^{\alpha_1 \rho_1} (w - s)^{\alpha_1 (1-\rho_1)}. \tag{A9}$$

By using almost identical steps, we can show that

$$\lim_{\sigma \rightarrow 0} G = w^{\alpha_2} (w + \psi - s)^{(1-\alpha_2)(1-\rho_2)} (w - s)^{(1-\alpha_2)\rho_2}. \tag{A10}$$

It is assumed, as stated in the first paragraph of Section 2, that $\lim_{\sigma \rightarrow 0} I < w - (1 - \delta)s$ for the relevant δ . Finally, as demonstrated in lemma 1, $I_\sigma > 0$, $G_\sigma > 0$. Hence, the intermediate-value theorem implies that there exist σ_j for $J \in \{I, G\}$ such that $J(\sigma_j, \psi) \geq w - (1 - \delta)s$ if and only if $\sigma \geq \sigma_j$.

ii) Moreover, $I(\sigma, \psi) > G(\sigma, \psi)$ for all $\sigma > 0$, since $1 > \alpha_1 + \alpha_2$. To see this, note that per expressions (1) and (2), this condition holds whenever

$$(1 - \alpha_1 - \alpha_2)w^\sigma > [(1 - \alpha_2)(1 - \rho_2) - \alpha_1 \rho_1](w + \psi - s)^\sigma + [(1 - \alpha_2)\rho_2 - \alpha_1(1 - \rho_1)](w - s)^\sigma \tag{A11}$$

and

$$\begin{aligned}
 (1 - \alpha_1 - \alpha_2)w^\sigma &\geq (1 - \alpha_1 - \alpha_2)(w + \psi - s)^\sigma \\
 &> [(1 - \alpha_2)(1 - \rho_2) - \alpha_1\rho_1](w + \psi - s)^\sigma \\
 &\quad + [(1 - \alpha_2)\rho_2 - \alpha_1(1 - \rho_1)](w - s)^\sigma,
 \end{aligned}
 \tag{A12}$$

since $\psi \leq s$. That $I_\sigma > 0$, $G_\sigma > 0$, and $I > G$ for all $\sigma > 0$ together imply that $\sigma_G > \sigma_I$. Q.E.D.

Proof of Proposition 1

Expression (2) implies that $\sigma_G(\psi, \delta)$, the cutoff attitude toward risk for guilty individuals, is given by

$$w - (1 - \delta)s = G(\sigma_G, \psi). \tag{A13}$$

Denote by $\sigma_{G_N} = \sigma_G(0, \delta_N)$ the threshold attitude toward risk for guilty individuals when there is no exoneree compensation. Next, let $\delta^*(\psi)$ be the discount rate, as a function of ψ , that produces the same threshold attitude toward risk. It follows from expression (2) that

$$\delta^*(\psi) = \frac{G(\sigma_{G_N}, \psi) - (w - s)}{s}. \tag{A14}$$

That $\delta^*(\psi) \in [\delta_N, 1)$ can be seen by noting that $\delta_\psi^* > 0$ (since $G_\psi > 0$), $\delta^*(0) = \delta_N$, and $\delta^*(s) = [G(\sigma_{G_N}, s) - (w - s)]/s < 1$ (since $G(\sigma, s) < w$ for all σ).

Next, define σ_{I_c} , the threshold attitude toward risk for innocent defendants when the compensation level is ψ and the discount rate is $\delta = \delta^*(\psi)$, as

$$H(\sigma_{I_c}, \psi) = I(\sigma_{I_c}, \psi) - w + [1 - \delta^*(\psi)]s = 0. \tag{A15}$$

The implicit-function theorem can be used to identify how an increase in the exoneree compensation affects σ_{I_c} :

$$\frac{d\sigma_{I_c}}{d\psi} = -\frac{\partial H / \partial \psi}{\partial H / \partial \sigma_{I_c}} = -\frac{I_\psi - s\delta_\psi^*}{I_\sigma} < 0, \tag{A16}$$

which holds if $I_\psi > s\delta_\psi^*$, since $I_\sigma > 0$, as proved in lemma 1. Because $s\delta_\psi^*(\psi) = G_\psi$, this condition is equivalent to

$$I\psi(\sigma_{I_c}, \psi) - G_\psi(\sigma_{G_N}, \psi) > 0. \tag{A17}$$

Differentiating I and G with respect to ψ and substituting the derivatives reveals that condition (A17) is equivalent to

$$(w + \psi - s)^{\sigma_{I_c} - \sigma_{G_N}} \frac{I^{1 - \sigma_{I_c}}}{G^{1 - \sigma_{G_N}}} \frac{\rho_1 \alpha_1}{1 - \alpha_2} > 1 - \rho_2. \tag{A18}$$

After substituting in $I(\sigma_{1c}, \psi) = G(\sigma_{G_N}, \psi) = w - [1 - \delta^*(\psi)]s$ (as noted in equations [A15] and [A16]), this condition becomes

$$\left\{ \frac{w - [1 - \delta^*(\psi)]s}{w + \psi - s} \right\}^{\sigma_{G_N} - \sigma_{1c}} \frac{\rho_1 \alpha_1}{1 - \alpha_2} > 1 - \rho_2. \quad (\text{A19})$$

A more intuitive condition that implies condition (A19) is

$$\left[\frac{w - (1 - \delta_N)s}{w} \right]^{\sigma_{G_N} - \sigma_{1c}} \frac{\rho_1 \alpha_1}{1 - \alpha_2} > 1 - \rho_2, \quad (\text{A20})$$

since $\delta^*(\psi) \geq \delta_N$ and $\psi \leq s$. It is worth noting that the left-hand side of condition (A20) is bounded from below, because $\bar{\sigma}$, defined as $I(\bar{\sigma}, 0) = w - [1 - \delta(s)]s$, constitutes an upper bound for σ_{1c} . Thus, when $1 - \rho_2$ is sufficiently small, it follows that $\sigma_1[\psi, \delta^*(\psi)] < \sigma_1(0, \delta_N)$ and $\sigma_G[\psi, \delta^*(\psi)] = \sigma_G(0, \delta_N)$ for all $\psi > 0$. Q.E.D.

Proof of Proposition 2

i) The utility of individuals who take (refuse) plea bargains is increasing in ψ (δ). Hence, these utilities are maximized when ψ and δ are at their maximal levels.

ii) The number of innocent individuals pleading guilty is increasing in σ_p , since individuals with $\sigma < \sigma_1$ plead guilty. As shown in the proof of proposition 1, σ_1 is decreasing in ψ . Therefore, the number of innocent individuals pleading guilty is decreasing in ψ and is minimized when $\psi = s$. Moreover, individuals who take pleas are convicted with certainty, whereas trial implies a probability of wrongful conviction. Therefore, the number of wrongful convictions is increasing in the number of guilty pleas taken by innocent individuals and is minimized when $\psi = s$. Q.E.D.

References

- Armbrust, Shawn. 2004. When Money Isn't Enough: The Case for Holistic Compensation of the Wrongfully Convicted. *American Criminal Law Review* 41:157–82.
- Bar-Gill, Oren, and Oren Gazal-Ayal. 2006. Plea Bargains Only for the Guilty. *Journal of Law and Economics* 49:353–64.
- Bibas, Stephanos. 2004. Plea Bargaining outside the Shadow of Trial. *Harvard Law Review* 117:2464–2547.
- Blackstone, William. 1765. *Commentaries on the Laws of England*. Oxford: Clarendon Press.
- Block, Michael, and Robert Lind. 1975. An Economic Analysis of Crimes Punishable by Imprisonment. *Journal of Legal Studies* 4:479–92.
- Bowers, Josh. 2008. Punishing the Innocent. *University of Pennsylvania Law Review* 156:1117–79.
- Daughety, Andrew, and Jennifer Reinganum. 2015. Informal Sanctions on Prosecutors

- and Defendants and the Disposition of Criminal Cases. Working paper. Vanderbilt University Law School, Nashville, TN.
- Fon, Vincy, and Hans-Bernd Schäfer. 2007. State Liability for Wrongful Conviction: Incentive Effects on Crime Levels. *Journal of Institutional and Theoretical Economics* 163:269–84.
- Garoupa, Nuno, and Matteo Rizzolli. 2012. Wrongful Convictions Do Lower Deterrence. *Journal of Institutional and Theoretical Economics* 168:224–31.
- Gazal-Ayal, Oren. 2006. Partial Ban on Plea Bargains. *Cardozo Law Review* 27:2295–2349.
- Gazal-Ayal, Oren, and Avishalom Tor. 2012. The Innocence Effect. *Duke Law Review* 62:339–401.
- Givati, Yehonatan. 2014. Legal Institutions and Social Values: Theory and Evidence from Plea Bargaining Regimes. *Journal of Empirical Legal Studies* 11:867–93.
- Gross, Samuel. 2008. Convicting the Innocent. *Annual Review of Law and Social Science* 4:173–92.
- Gross, Samuel, Kristen Jacoby, Daniel Matheson, Nicholas Montgomery, and Sujata Patil. 2005. Exonerations in the United States, 1989 through 2003. *Journal of Criminal Law and Criminology* 95:523–60.
- Grossman, Gene, and Michael Katz. 1983. Plea Bargaining and Social Welfare. *American Economic Review* 73:749–57.
- Innocence Project. 2009. When the Innocent Plead Guilty. January 26. <http://www.innocenceproject.org/when-the-innocent-plead-guilty>.
- Lando, Henrik. 2006. Does Wrongful Conviction Lower Deterrence? *Journal of Legal Studies* 35:327–37.
- Lando, Henrik, and Murat Mungan. 2014. The Effect of Type-I Error on Deterrence. Public Law Research Paper No. 687. Florida State University, College of Law, Tallahassee.
- Mandery, Evan J., Amy Schlosberg, Valerie West, and Bennet Callaghan. 2013. Compensation Statutes and Post-exoneration Offending. *Journal of Criminal Law and Criminology* 103:553–84.
- Norris, Robert J. 2012. Assessing Compensation Statutes for the Wrongly Convicted. *Criminal Justice Policy Review* 23:352–74.
- Reinganum, Jennifer. 1988. Plea Bargaining and Prosecutorial Discretion. *American Economic Review* 78:713–28.
- Simon, Dan. 2012. *In Doubt: The Psychology of the Criminal Justice System*. Cambridge, MA: Harvard University Press.
- Tor, Avishalom, Oren Gazal-Ayal, and Stephen Garcia. 2010. Fairness and the Willingness to Accept Plea Bargain Offers. *Journal of Empirical Legal Studies* 7:97–116.