

Signaling and Plea Bargaining's Innocence Problem

Russell D. Covey*

Abstract

The dominant theoretical model of plea bargaining predicts that, under conditions of full information and rational choice, criminal cases should uniformly be settled through plea bargaining. That prediction holds for innocent and guilty defendants alike. Because it is perfectly rational for innocent defendants to plead guilty, plea bargaining might be said to have an "innocence problem." Plea bargaining's innocence problem is, at bottom, the result of a signaling defect. Innocent defendants lacking verifiable innocence claims are pooled together with guilty defendants who falsely proclaim innocence. As a result, both groups of defendants are treated similarly at trial and in plea bargaining. Signaling defects, however, at least in theory, can be overcome. As economists have observed, a variety of signaling mechanisms reliably communicate nonverifiable information among parties. This Article considers one such signaling mechanism, subwagers, and attempts to demonstrate their capacity to signal the kind of private information typical in the plea bargaining context. Not only are subwagers theoretically possible, the Article argues that police interrogation already functions as such a subwager. The Article addresses the theoretical basis and empirical evidence supporting the hypothesis regarding interrogation's important impact on plea bargaining, and concludes that plea bargaining theory must be modified to take account of the effects of subwager-type signaling.

Table of Contents

| | |
|---|----|
| I. Introduction | 74 |
| II. Plea Bargaining's Innocence Problem | 77 |

* Associate Professor, Georgia State University College of Law. J.D., Yale Law School; M.A., Princeton University; A.B., Amherst College; rcovey@gsu.edu. My thanks to Janine Kim, Michael O'Hear, and Martin Pritikin for their valuable comments on an earlier draft of this Article.

| | | |
|------|--|-----|
| A. | Trial Shadow Theory | 77 |
| B. | Some Proposed Fixes and Their Limitations | 83 |
| 1. | Abolition | 83 |
| 2. | Screening and the Capacity/Accuracy Tradeoff | 86 |
| 3. | Improved Pretrial Discovery | 88 |
| III. | The Innocence Problem as a Signaling Defect | 91 |
| A. | Signaling Mechanisms | 95 |
| B. | The Two-Move Subwager Game | 103 |
| IV. | Interrogation and Sub-Wagers | 110 |
| A. | Cooperation as an Anti-Pooling Device | 110 |
| B. | Interrogation as a Subwager | 120 |
| C. | Implications | 124 |
| V. | Conclusion | 130 |

I. Introduction

Plea bargaining has an innocence problem. The dominant theoretical model of plea bargaining—so-called "trial shadow theory"—predicts that, once charged, innocent and guilty persons alike almost always act rationally by pleading guilty rather than contesting guilt at trial.¹ Indeed, it predicts that every case able to withstand minimal pretrial scrutiny should be resolved (assuming the parties are adequately informed about the evidence and act rationally) with a negotiated plea bargain.² By all accounts, those predictions are not far off the mark. The latest data available indicates that approximately 95% of all adjudicated felony criminal charges are disposed with through guilty pleas.³ When more than forty-seven out of every fifty dispositions are guilty pleas, it is safe to say the theory is on to something.

1. See Stephanos Bibas, *Plea Bargaining outside the Shadow of Trial*, 117 HARV. L. REV. 2464, 2464 (2004) ("The conventional wisdom is that litigants bargain toward settlement in the shadow of expected trial outcomes. In this model, rational parties forecast the expected trial outcome and strike bargains that leave both sides better off by splitting the saved costs of trial.").

2. See *id.* at 2466 ("By and large, though, scholars view the shadow of trial as the overwhelming determinant of plea bargaining.").

3. Douglas A. Berman & Stephanos Bibas, *Making Sentencing Sensible*, 4 OHIO ST. J. CRIM. L. 37, 42 (2006).

Plea bargaining's critics have responded to the avalanche of guilty pleas with a wide range of proposed reforms. Some have advocated total abolition of plea bargaining;⁴ others have urged shifting resources to early screening to prevent weak cases from being prosecuted.⁵ Still others have called for more liberal discovery requirements to increase the amount of information available to defendants at the time they enter guilty pleas.⁶ Although there are merits to all these suggestions, none do much to address the innocence problem—at least not without imposing costs upon society (in the form of making it harder to convict the guilty) that society seems ill-prepared to accept. In fact, some of these reforms—abolition and enhanced pre-plea discovery in particular—might well make innocent defendants worse off rather than better, or at least more likely to enter false guilty pleas.

The economic model of plea bargaining highlights the central dynamic driving the innocence problem: As long as plea prices are determined by trial shadows, anything that fails to cast a trial shadow is irrelevant to plea prices. But information that is irrelevant to trial outcomes might not be irrelevant to other important issues that bear upon just outcomes. Take, for instance, a defendant's honest but unverifiable belief that she is factually innocent of the crime charged.⁷ Although seemingly a highly salient piece of information, such private information cannot affect plea prices if the defendant has no way to communicate the information to the prosecutor or a jury. As a result, defendants who know they are innocent would appear to do no better in plea bargaining than defendants who know they are guilty.⁸ In game theory terms, pleading guilty pursuant to a utility-maximizing bargain is a "dominant strategy" for guilty and innocent defendants alike, irrespective of the content of private information.

4. See *infra* Part II.B.1 (discussing the claims of scholars who advocate the abolition of plea bargaining).

5. See *infra* Part II.B.2 (discussing the claims of scholars who advocate limiting prosecutorial discretion).

6. See *infra* Part II.B.3 (discussing the claims of scholars who advocate greater discovery rights for defendants pretrial).

7. I refer here to an "honest belief" of factual innocence, rather than knowledge, because some defendants might be honestly mistaken. The exclusion of private information from the calculation of plea price in that instance enhances accuracy, but I assume for purposes of this Article that honest mistakes about innocence are rare.

8. A contrary problem arises when the prosecutor holds private information about what is likely to happen at trial but prefers not to reveal it to the defendant precisely because of the impact that information will have on the plea price. This can happen, for instance, where an important eyewitness to the crime dies or otherwise becomes available. For further discussion, see *infra* notes 52–60 and accompanying text.

Plea bargaining's dominance as a strategy depends, of course, on the assumptions that one makes about the payoff structure and rules of the game. Under standard assumptions, private information about guilt and innocence is irrelevant to plea bargaining outcomes. Under other payoff and rule structures, however, private information can be made relevant through what economists refer to as "signaling."⁹ This Article discusses one such payoff and rule structure, which I refer to as a "subwager," that might permit incorporation of private information into plea pricing to improve the accuracy of the pricing system. In other words, under certain circumstances subwagers might increase sentence differentials between defendants who know they are innocent and those who know they are guilty.

Part II describes the innocence problem in greater detail and demonstrates why abolition, stronger prosecutorial charge screening, and enhanced discovery all fail to resolve it. Part III then considers the innocence problem as a signaling defect. Innocent defendants have private information that is relevant but not verifiable. Because they cannot reliably signal their innocence in a manner prosecutors can trust, that private information is effectively lost. Part III then demonstrates, using a simple card game as a model, how introduction of a second move—a subwager—might permit private information to have an impact on plea prices in a way that enhances the accuracy of plea bargaining. Finally, Part IV considers whether subwagers have any relevance to criminal procedure as currently practiced or as might be envisioned. It concludes that interrogation acts as a signaling subwager in some respects and reviews data gathered by interrogation researchers that support this conclusion. The analysis suggests that plea bargaining theory must be modified to reflect how pre-trial signaling impacts plea prices. Contrary to the predictions of the shadow of trial model, plea bargaining may in some ways actually improve upon trial outcomes. Signaling also helps explain some of the puzzles surrounding interrogation, including why courts tolerate police trickery and deception in interrogation and why so many suspects submit to interrogation notwithstanding the conventional wisdom that invoking the right to silence is the better strategy. Finally, the subwager concept suggests further ways that private information might be harvested to increase the accuracy and reliability of plea bargaining.

9. See Robert B. Ahdieh, *The Strategy of Boilerplate*, 104 MICH. L. REV. 1033, 1041 (2006) ("Signaling involves the communication of information that cannot be effectively communicated through explicit statements of intention or character.").

II. Plea Bargaining's Innocence Problem

A. Trial Shadow Theory

Trial shadow theory provides the dominant account of plea bargaining.¹⁰ According to trial shadow theory, plea bargaining is a voluntary dispute resolution mechanism that occurs, like settlements in civil cases, in the "shadow" of trials.¹¹ Plea bargains are the product of both parties' rational forecast of trial outcomes, and negotiations take place against this background.¹²

In a world where plea bargains are negotiated in trial shadows, the most important factors that determine the terms on which a plea bargain is negotiated—what I refer to as the "price" of a plea—are (1) the trial sentence anticipated if the case were tried and resulted in a conviction,¹³ (2) the likelihood that a trial will result in a conviction,¹⁴ and (3) the resource costs of trying the case.¹⁵ Assuming that prosecutors seek to maximize and defendants

10. See Bibas, *supra* note 1, at 2464 ("Plea-bargaining literature predicts that parties strike plea bargains in the shadow of expected trial outcomes.").

11. See generally Robert H. Mnookin & Lewis Kornhauser, *Bargaining in the Shadow of the Law: The Case of Divorce*, 88 YALE L.J. 950 (1979) (examining the impact of the legal system on negotiations and bargaining that occur outside the courtroom in the context of divorce law).

12. See Robert E. Scott & William J. Stuntz, *Plea Bargaining as Contract*, 101 YALE L.J. 1909, 1937 (1992) (stating that both prosecutor and defendant must generate an estimate of the likelihood that defendant will be convicted if the case goes to trial and "that estimate will determine the price that each will insist on as a condition of reaching a bargain"). Not all scholars of plea bargaining, of course, agree that plea prices are determined by trial shadows. See, e.g., Stephen Schulhofer, *Criminal Justice Discretion as a Regulatory System*, 17 J. LEGAL STUD. 43, 75 (1988) (criticizing trial shadow theory).

13. See Bibas, *supra* note 1, at 2465 (noting that one of the factors that determines the plea price is "the likely post-trial sentence").

14. Given any initial charge carrying a standard sentence upon conviction, the primary pricing variable is the strength of the evidence in the case. Albert W. Alschuler, *The Prosecutor's Role in Plea Bargaining*, 36 U. CHI. L. REV. 50, 58 (1966) ("The overwhelming majority of prosecutors view the strength or weakness of the state's case as the most important factor in the task of bargaining.").

15. Scott & Stuntz, *supra* note 12, at 1941 (noting that the need to conserve adjudication costs lowers the plea price, and "explain[s] why the parties might reach a bargain even in cases where both sides are virtually certain that the defendant will be convicted at any subsequent trial and have a good idea what the resulting sentence will be"). In most serious cases, the state will have to pay most resource costs because most criminal defendants are indigent and represented by appointed counsel. See CAROLINE WOLF HARLOW, U.S. DEP'T OF JUSTICE, BUREAU OF JUSTICE STATISTICS SPECIAL REPORT: DEFENSE COUNSEL IN CRIMINAL CASES 5 tbl.7 (2000), available at <http://www.ojp.usdoj.gov/bjs/pub/pdf/dccc.pdf> (showing that among defendants

seek to minimize sentences, the price of any plea should be the product of the anticipated trial sentence and the likelihood of conviction, discounted by some factor to reflect the resources saved by not having to try the case.¹⁶ A more refined plea pricing model might also include such additional factors as the parties' respective risk preferences, agency costs, and discount rates.¹⁷

Where plea bargain prices are based primarily on expected trial outcomes, a defendant's ability to negotiate a lenient plea bargain will largely depend on the strength of the evidence and the potential resource savings from a guilty plea.¹⁸ Prosecutors should be willing to "pay" greater discounts for guilty pleas where the evidence is weak or the resource costs of trial are greater relative to

charged with violent felonies in large metropolitan counties, 83% are poor enough to qualify for court-appointed counsel).

16. See Jennifer F. Reinganum, *Plea Bargaining and Prosecutorial Discretion*, 78 AM. ECON. REV. 713, 714 (1988) (defining expected punishment as "the product of the probability of conviction and the anticipated sentence upon conviction at trial"). This model does not incorporate what is sometimes referred to as "cooperation bargaining." See Eric Rasmusen, *Mezzanatto and the Economics of Self-Incrimination*, 19 CARDOZO L. REV. 1541, 1552 (1998) ("Cooperation bargaining occurs when the defendant has information to trade, and the gain from bargaining includes this information, which could be used in other trials."). Cooperation bargaining, of course, further lowers plea prices because prosecutors are willing to pay a premium for the information, testimony, or both, that the defendant provides.

17. William Landes developed one of the first and most influential economic models of settlement. See William M. Landes, *An Economic Analysis of the Courts*, 14 J.L. & ECON. 61, 61 (1971) (stating that "the decision to settle or go to trial depends on the probability of conviction by trial, the severity of the crime, the availability and productivity of the prosecutor's and defendant's resources, trial versus settlement costs, and attitudes toward risk"). Frank Easterbrook supplemented Landes' model by incorporating discount rates and agency costs into the model. See Frank H. Easterbrook, *Criminal Procedure As a Market System*, 12 J. LEGAL STUD. 289, 309 (1983) (noting that agency costs are not unique to the plea bargaining process). In addition, because guilty pleas are substantially cheaper than trials, and because the costs of trial usually (but not always) are imposed disproportionately on the state, the defendant can usually command an additional discount in exchange for his plea. See Russell D. Covey, *Fixed Justice: Reforming Plea Bargaining with Plea-Based Ceilings*, 82 TUL. L. REV. 1237, 1247 (2008) ("[T]he lion's share of process costs are borne by the state, and, as a result, prosecutors typically gain more than defendants by avoiding trial.").

18. See David Bjerk, *Guilt Shall Not Escape or Innocence Suffer? The Limits of Plea Bargaining When Defendant Guilt is Uncertain*, 9 AM. L. & ECON. REV. 305, 307 (2007) (explaining that plea prices increase as probability of guilt increases); Easterbrook, *supra* note 17, at 297 ("As the sentence on conviction and probability of conviction rise, so does the prosecutor's minimum demand."). For example, where the anticipated sentence upon conviction is ten years and there is a fifty percent likelihood of conviction, both prosecutors and defendants should, setting complicating factors aside, be preference-neutral to a plea bargained sentence of five years or a trial, and where the likelihood of conviction is twenty percent, the preference-neutral plea price is two years.

expected trial sentences.¹⁹ Assuming that guilty pleas are a more efficient way to resolve charges, however, prosecutors who seek to maximize punishment per unit of resource should always be willing to pay some discount reflecting those savings. It seems to follow that rational parties will uniformly prefer to resolve their cases through guilty pleas rather than trials, at least where they have reasonably complete information about admissible evidence.²⁰

Most significantly for innocent defendants, the evidentiary strength of a case should not impact the rationality of resolving a case through plea bargaining. In cases where the evidence of guilt is strong, only modest penalty reductions need be offered to induce rational defendants to plead guilty; when the evidence is weak, however, more significant inducements must be offered.²¹ Nonetheless, assuming that the prosecutor's discretion to bargain is unfettered²²

19. As a result, the model predicts that plea bargains will be most generous (and therefore most frequently accepted) in cases involving misdemeanors and other less serious offenses. The process costs expended by defendants will be particularly high relative to penalty costs where only minor penalties are involved. *See, e.g.*, MALCOLM FEELEY, *THE PROCESS IS THE PUNISHMENT* 30 (1992) (arguing that for defendants accused of petty crimes it is the cost of being caught up in the criminal justice system that is most bothersome). Process costs will also be increasingly important in cases in which the defendant is incarcerated before trial and the length of pretrial detention begins to approach the expected punishment upon conviction. *See* Josh Bowers, *Punishing the Innocent*, 156 U. PA. L. REV. 1117, 1138 (2008) (noting that for innocent defendants, the costs of pretrial detention are relevant if an offer exists that promises release in a period shorter than the pretrial interval). Process costs may also impose disproportionate burdens on white collar criminals and the minority of other defendants who pay for their own legal representation. In each of those cases, the defendant might even be willing to pay a premium over the expected trial sentence, rather than take a discount, to obtain the benefits of avoiding a drawn out trial process. Nonetheless, that dynamic does not alter the fact that both parties maximize their utility by settling the case through plea bargaining.

20. *See, e.g.*, Chantale LaCasse & A. Abigail Payne, *Federal Sentencing Guidelines and Mandatory Minimum Sentences: Do Defendants Bargain in the Shadow of the Judge?*, 42 J.L. & ECON. 245, 249 (1999) (stating that models based on minimizing expected sentence predict that all cases are resolved by plea and criticizing these models on that basis). For reasons that will be discussed further below, defendants will rarely possess significant private information about the trial evidence because prosecutors are responsible for developing the inculpatory evidence, and defendants have powerful incentives to reveal exculpatory evidence before trial.

21. *See, e.g.*, Alschuler, *supra* note 14, at 60 (noting "the practice of bargaining hardest when the case is weakest"); Welsh S. White, *A Proposal for Reform of the Plea Bargaining Process*, 119 U. PA. L. REV. 439, 451 (1971) (quoting one attorney who stated, "New York prosecutors often reduce their sentence recommendations by at least fifty percent if they believe that there is a fifty percent chance of a hung jury, and by a great deal more if they believe that there is a fifty percent chance of acquittal").

22. Of course, artificial constraints on bargaining can prevent parties from reaching a mutually beneficial settlement. As some commentators have observed, mandatory minimum sentences impede efficient "market resolution" of some cases, especially cases that are particularly weak (including those brought against the innocent), because they prevent the

and both parties understand the relevant pricing inputs and hold similar views regarding their values, they should be able to negotiate a mutually beneficial plea price even in cases where the defendant would likely be acquitted at trial.²³

In short, as long as the prosecutor is willing and able to discount plea prices to reflect resource savings, regardless of guilt or innocence, pleading guilty is the defendant's dominant strategy. As a result, non-frivolous accusation—not proof beyond a reasonable doubt—is all that is necessary to establish legal guilt.²⁴ This latter point forms the root of plea-bargaining's "innocence problem," which refers here not merely to the fact that innocent people plead guilty, but that the economics of plea bargaining drives them to do so.

It should be noted that many of plea bargaining's defenders are not particularly troubled by this.²⁵ First, they point out that the innocence problem is not unique to plea bargaining but rather is a necessary consequence of imperfect trials.²⁶ If trials perfectly separated guilty defendants from innocent

parties from settling the case through agreement of a sufficiently light sentence to induce a rational defendant to relinquish his trial right. *See* Scott & Stuntz, *supra* note 12, at 1960–66 (arguing that when legislatures draft broad criminal statutes with mandatory minimum sentences, prosecutors have greater opportunity to overcharge and generate easy pleas). Mandatory minimums and maximums distort the plea bargaining market in ways similar to conventional price controls in economic markets by setting price floors and ceilings that bar transactions that otherwise would occur between willing bargainers. *Id.* at 1962–65. Abolition of mandatory minimum sentences might well increase the number of innocent defendants who plead guilty. *But see* Stephen J. Schulhofer, *Plea Bargaining as Disaster*, 101 YALE L.J. 1979, 1992 (1992) (arguing that repealing mandatory minimum sentencing laws would have little effect on the number of innocent defendants who plead guilty). The judicial power to reject apparently over-lenient plea bargains can have a similar distortive effect. *Id.* at 1995–96.

23. This is one of the major complaints leveled at plea bargaining. *See generally* Oren Bar-Gill & Oren Gazal Ayal, *Plea Bargains Only for the Guilty*, 49 J.L. & ECON. 353 (2006) (arguing for limitations on prosecutorial bargaining discretion to prevent prosecutors from inducing innocent defendants to plead guilty where evidence is weak).

24. This is because as long as the prosecutor can make a colorable threat of punishment, there should be some compromise plea bargain that the defendant should rationally prefer. Anecdotal evidence suggests this is how the system is perceived by defendants. As one defense lawyer explained in counseling his likely innocent client to plead guilty to a misdemeanor to avoid trial on serious felony charges, "Once a person is facing felony charges, the issue no longer is whether he did the crime; it's how to limit the damage." STEVE BOGIRA, COURTROOM 302: A YEAR BEHIND THE SCENES IN AN AMERICAN CRIMINAL COURTHOUSE 334 (2005).

25. *See, e.g.,* Scott W. Howe, *The Value of Plea Bargaining*, 58 OKLA. L. REV. 599, 603 (2005) (arguing that plea bargaining is valuable because it "maximizes deserved punishment" given resource constraints); Edward A. Ruttenburg, *Plea Bargaining Analytically—The Nash Solution to the Landes Model*, 7 AM. J. CRIM. L. 323, 353 (1979) (concluding that plea bargaining "benefit[s] defendants as a group").

26. *See* Frank H. Easterbrook, *Plea Bargaining as Compromise*, 101 YALE L.J. 1969,

defendants, no innocent defendant would plead guilty (at least as long as the process costs of trial did not exceed the costs of conviction) because every innocent defendant would go to trial and be acquitted. It is precisely because trials are flawed, and innocent persons are sometimes convicted, that innocent defendants might rationally perceive pleading guilty to be the utility-maximizing option.²⁷ Second, the charge that plea bargaining is unfair or unduly coercive is challenged by the observation that both prosecutors and defendants prefer plea bargaining.²⁸ After all, if either party believed that the bargain was not beneficial, it could simply decline it and go to trial. Notwithstanding the fact that some innocent defendants are induced to plead guilty as a result of the plea pricing mechanism, even innocent defendants are better off with plea bargaining, it is argued, than they would be without it.²⁹

These arguments have some persuasive force. Given that trials are imperfect, plea bargaining at least permits some mutually beneficial settlements that reflect rational criteria, and plea bargaining's outcomes might even be said to be just in the majority of cases. To the extent that evidence tends to be strongest against defendants who are "actually guilty," and weakest against defendants who are innocent, the pricing model seems consistent with conventional conceptions of justice.³⁰

1970 (1992) ("What disrupts this separation of the guilty from the innocent is not a flaw in the bargaining process but a flaw at trial.").

27. See *id.* at 1969 ("Persons at risk of unjust conviction may prefer a certain (but low) punishment in a plea bargain to the risk of conviction and high punishment after trial."); see also Easterbrook, *supra* note 17, at 320 ("Sometimes the evidence may point to guilt despite the defendant's factual innocence. It would do defendants no favor to prevent them from striking the best deals they could in such sorry circumstances.").

28. See Easterbrook, *supra* note 17, at 309 ("Defendants presumably prefer the lower sentences to the exercise of their trial rights or they would not strike the deals. Prosecutors also prefer the agreements; they may put the released resources to use in other cases, thus increasing deterrence.").

29. See Scott & Stuntz, *supra* note 12, at 1961 (describing the situation of an innocent defendant facing a life sentence upon conviction who is offered a discounted sentence of ten years in return for a guilty plea as "a sad one, but preventing the offer only makes it sadder"); Easterbrook, *supra* note 17, at 320 (arguing that even where evidence falsely implicates factually innocent defendants, "[i]t would do defendants no favor to prevent them from striking the best deals they could in such sorry circumstances"); Easterbrook, *supra* note 26, at 1975 (arguing that defendants should not be forced to exercise their trial rights as such a rule would force defendants "to take the risk of conviction or acquittal" even if they prefer "a certain but small punishment to a chancy but large one").

30. See William J. Stuntz, *The Uneasy Relationship Between Criminal Procedure and Criminal Justice*, 107 YALE L.J. 1, 48 (1997) (hypothesizing that the "odds of conviction . . . correlate strongly with innocence and guilt").

The strength of the evidence, however, will not always reflect the actual guilt or innocence of the defendant. Cases that turn on a confident witness's identification, for example, may look quite strong to a jury but include a large proportion of actually innocent defendants.³¹ Similarly, in cases in which the defense contends mistake, lack of mens rea, or some other hard-to-prove fact, the evidentiary problem will play a disproportionately large role in the determination of the plea price.³² In addition, the probability of conviction may turn on facts that are not strictly related to the strength of the evidence. For instance, all things being equal, a defendant with a criminal history faces a higher chance of conviction given the state's ability to impeach him with that criminal record if he takes the witness stand.³³

While replication of trial outcomes is the gold standard in the trial shadow model, the ideal of replicating trial outcomes points to the limitations of the model; if plea bargains are a reflection of expected trial outcomes, plea-bargained outcomes will, by definition, never be any more accurate than trial outcomes.³⁴ To the extent that trials themselves are imperfect vehicles for separating the innocent from the guilty, plea bargains cast in trial shadows can at best only replicate the error rate of the trial system.³⁵ Moreover, given the

31. See Easterbrook, *supra* note 26, at 1970 ("An eyewitness may have made a mistaken identification. The guilty person may have accused an innocent one to divert suspicion. It is hard to assess the worth of testimony. Juries give eyewitness testimony more credence than it deserves.").

32. See *id.* ("Guilt may depend on the defendant's mental state, which is elusive.").

33. See, e.g., Daniel Givelber, *Lost Innocence: Speculation and Data About the Acquitted*, 42 AM. CRIM. L. REV. 1167, 1190 (2005) (presenting data from a sample showing that juries were substantially more likely to convict defendants that they knew had criminal records compared with those that they knew did not). The history of the rise of plea bargaining shows that when states began allowing defendants to testify under oath in the 1860s and 1870s, defendants with criminal records faced increasing pressure to enter guilty pleas because their chances at trial were diminished: Either they testified and were subject to impeachment, or they declined to testify, in which cases juries almost certainly drew an adverse inference. See Jennifer L. Mnookin, *Uncertain Bargains: The Rise of Plea Bargaining in America*, 57 STAN. L. REV. 1721, 1729 (2005) (reviewing GEORGE FISHER, *PLEA BARGAINING'S TRIUMPH: A HISTORY OF PLEA BARGAINING IN AMERICA* (2003)).

34. To be clear, trial shadow theory does not predict that those who plead guilty and those who go to trial will receive, ex post, the same punishments. Rather, it predicts that defendants will be offered plea bargains that equal the ex ante expected trial punishment (discounted, of course, to reflect resource savings). See Bibas, *supra* note 1, at 2465 ("[T]he classic shadow-of-trial model predicts that the likelihood of conviction at trial and the likely post-trial sentence largely determine plea bargains.").

35. Like the shadow of trial and abolition theories, other intriguing approaches to assessing plea bargaining nonetheless share the flaw of presuming that trial outcomes are the gold standard of measuring the success of any plea bargaining system. For instance, Ronald

substantial uncertainties inherent in attempting to estimate probabilities of conviction based on the often skeletal pretrial evidentiary record, outcomes negotiated in the shadows of trial likely are a good deal less accurate than trial outcomes themselves.³⁶ Thus, both in theory and in practice, the trial shadow model depicts plea bargaining as at best a comparable substitution for trial, lacking any of the collateral benefits of trial (such as the creation of a public record of the crime and the evidence) but producing substantial savings in resources, and at worst, a vastly inferior substitute that we put up with, like a counterfeit Rolex watch, because it is much cheaper than the real thing.

B. Some Proposed Fixes and Their Limitations

1. Abolition

Critics of plea bargaining have amassed an extensive catalogue of problems with plea bargaining, leading some to conclude that it should be abolished altogether.³⁷ Although it might well be warranted, abolition of plea bargaining does not solve the innocence problem. In a world of mandatory trials, innocent defendants would continue to be falsely convicted of crimes. If all innocent defendants held out for trial rather than entered guilty pleas, fewer of them would be convicted. But those who were convicted would receive proportionately harsher sentences that do not reflect the possibility of trial error (i.e., innocence).

Moreover, the quality or accuracy of individual trials would not likely improve in a universal trial system. Even if trials of the same quality now given to defendants in the five to ten percent of cases currently tried were provided to all criminal defendants, with no diminution in outcome accuracy,³⁸ prohibition

Wright's "trial distortion theory" begins with the premise that "trial outcomes, by definition," are "truthful outcomes." Ronald F. Wright, *Trial Distortion and the End of Innocence in Federal Criminal Justice*, 154 U. PA. L. REV. 79, 107 (2005).

36. See Bibas, *supra* note 1, at 2494–95 (arguing that as a result of weak discovery rules, accelerated timing, agency costs, and strategic bargaining tactics like bluffing and puffery, parties often "bargain blindfolded," substantially diminishing the likelihood that bargains accurately reflect trial outcomes).

37. For a representative sampling, see generally Albert W. Alschuler, *The Changing Plea Bargaining Debate*, 69 CAL. L. REV. 652 (1981); Albert W. Alschuler, *Plea Bargaining and Its History*, 13 LAW & SOC'Y REV. 211 (1979); Schulhofer, *supra* note 22; Stephen J. Schulhofer, *Is Plea Bargaining Inevitable?*, 97 HARV. L. REV. 1037 (1984).

38. It has long been widely assumed that more trials would mean less accurate trials. See, e.g., Arnold Enker, *Perspectives on Plea Bargaining*, in TASK FORCE ON THE ADMIN. OF JUSTICE,

would not enhance accuracy over a system of plea bargaining that accurately represented trial outcomes, as the trial shadow model assumes.³⁹ And if, as seems more likely, the quality and accuracy of trial outcomes would decrease if there were a substantial increase in the number of cases tried, then the accuracy of convictions in a regime of universal trials would likely be substantially diminished.⁴⁰ The abolition of plea bargaining would therefore result in more innocent persons being wrongfully convicted and sentenced to harsher average sentences. Even if the trial penalty disappeared (with abolition, defendants would no longer worry that the decision to contest charges might result in a higher sentence), some proportion of actually innocent defendants would still lose at trial and receive the same sentence as their actually guilty counterparts.⁴¹

PRESIDENT'S COMMISSION ON LAW ENFORCEMENT AND ADMIN. OF JUSTICE, TASK FORCE REPORT: THE COURTS 112 (1967) (noting that even if personnel were increased sufficiently to permit more trials, "[a] lowering of standards in order to man the store adequately may well result in poorer justice").

39. Of course, actual plea bargains do not perfectly correspond to the shadow of trial model. See Bibas, *supra* note 1, at 2545 (noting that "[t]rials affect pleas, but so do many other influences unrelated to the merits"); William J. Stuntz, *Plea Bargaining and Criminal Law's Disappearing Shadow*, 117 HARV. L. REV. 2548, 2548 (2004) (arguing that "there are serious impediments to efficient bargaining in criminal cases"). But the solution to the imperfect fit between practice and theory is to improve practice so that it more closely approximates the theory, assuming the theory posits a model we wish to emulate.

40. Stephen Schulhofer has urged widespread adoption of bench trials as an efficient substitution for guilty pleas. See Stephen J. Schulhofer, *Is Plea Bargaining Inevitable?*, 97 HARV. L. REV. 1037, 1050–87 (1984) (describing effects of Philadelphia's ban on plea bargaining). Although bench trials may well provide a more preferable mechanism to resolve many criminal cases than guilty pleas, the traditional right to trial guaranteed by the federal Constitution, as well as in those of each state, is a right to trial by jury. See *Duncan v. Louisiana*, 391 U.S. 145, 149 (1968) (holding that the right to a jury trial is a fundamental right for serious offenses). Whatever a bench trial is, it is not the same as a jury trial, and the justice bench trials administer would undoubtedly vary widely depending on the manner in which those trials were conducted. See *id.* at 156 (noting that jury trials were thought critical by framers, in large part, "to protect . . . against judges too responsive to the voice of higher authority"). Indeed, it may be that the real trial error rate is as low as it is simply because the number of trials is as low as it is. The numerous safeguards that have been built into the modern jury trial—from the privilege against self-incrimination, the right to counsel, the right to confront and cross-examine hostile witnesses, and the right to introduce relevant expert testimony to the cumbersome jury selection practices used to maximize the chances of a neutral fact-finding body—would simply not be feasible if deployed in every case. That observation is banal, but the fact that the development of this elaborate truth-discovering process was accompanied every step of the way by an ever-expanding plea-bargaining system is critical. A plea-bargaining system backed up by a maximally accurate trial check may produce more accurate results overall than a system of more summary trials with a higher error rate.

41. Example: The standard post-trial sentence for crime *X* is ten years. The prosecutor routinely offers a 50% plea discount. There is a 10% possibility of conviction (POC). The unitary pricing model predicts the following plea bargain: $(10)(0.1)(0.5) = 6$ months.

And assuming the total punishment meted out remains constant, abolition of plea bargaining would do nothing more than concentrate losses among a smaller number of unfortunates. It would not, without more, improve the ex ante lot of the actually innocent defendant.⁴²

The problem is most apparent in weak cases. As noted above, an unfettered plea bargaining market establishes conditions in which defendants with strong factual or legal defenses nonetheless are uniformly expected to receive, and accept, a plea offer sufficiently discounted to make it rational for them to prefer a guilty plea, a dynamic that looks a lot like coercing innocent people to plead guilty.⁴³ But barring such persons from making good deals and

Assuming that the average POC of all defendants charged in the jurisdiction is 75%, the average plea bargain for *X* crime, under this formula, should be 3.75 years. Now consider an abolition regime in which total punishment was held constant. Assume that there are 100 defendants in the jurisdiction. In the plea bargaining regime, those 100 defendants all plead guilty, producing 375 years of jail sentences. In an abolition regime, assume that those 100 defendants all go to trial. Of them, twenty-five will be acquitted, and the other seventy-five will be sentenced to an average term of five years. For the sake of simplicity, now assume that ten of those seventy-five defendants had POC's of 10%. In the abolition regime, nine would be acquitted at trial, but the tenth convicted. However, that defendant will receive, not the discounted sentence, but the full average sentence: Five years. As a whole, this group of ten likely innocents fares exactly the same under the abolition regime as under the plea bargaining regime, with only the ex post costs redistributed among them.

Now, we might say, as abolitionists like Albert Alschuler have, that the true "proportionate" or "just" sentence is the average plea-bargained sentence: 3.75 years. If everybody went to trial, and upon conviction, were sentenced to the average plea-bargain sentence, then we would see a net reduction of punishment (from 375 years to 281.25 years). The likely innocents would then benefit, as they would see their expected sentence after trial (ETS) fall from six months to 4.5 months. But then everybody, including guilty defendants, would see a decrease in ETS. In a plea bargain regime, defendants with 50% and 90% POCs, respectively, would have ETSs of 2.5 and 4.5 years. In an abolition regime, their ETSs would fall to 1.87 and 3.38 years. The point is, reducing sentences for everyone might well be the right thing to do, but the benefits of doing so flow to all defendants, not just the innocent ones. Such benefits, moreover, can be achieved simply by reducing maximum trial sentences—there is no need to abolish plea bargaining to obtain that benefit.

42. This outcome changes if we assume that increasing the number of trials will change the mix of cases that prosecutors will actually pursue. If prosecutors are induced to prosecute fewer cases, they naturally will dismiss their weakest cases first, which should improve the lot of the likely innocents. But reducing the number of cases prosecuted has deterrence costs. The gains to likely innocent defendants must be balanced against those deterrence losses. Moreover, there are many ways to reduce the number of cases prosecuted, any one of which will have the same effect of encouraging prosecutors to drop their weak cases first. *See infra* text accompanying notes 45–51 (evaluating various proposals designed to screen out weak cases).

43. Steven Bogira relates the story of one defendant, charged as an accomplice with attempted murder, aggravated battery, and a hate crime, and facing six to thirty years if convicted. BOGIRA, *supra* note 24, at 332–33. Although there was strong evidence against his two co-defendants, very little evidence connected the defendant with the crime. *Id.* at 327–30.

requiring them to stand trial and face a small chance of a large penalty does not obviously improve their lot.⁴⁴ *Ceteris paribus*, abolition of plea bargaining simply requires defendants to roll the dice. Many defendants, innocent and guilty alike, would perceive the restriction of choice as an uncompensated loss, which not only imposes substantial risk on the defendant but also eliminates both the discount (and the incentive) for minimizing trial costs.

2. *Screening and the Capacity/Accuracy Tradeoff*

A less drastic way to ameliorate plea bargaining's innocence problem might be to encourage the dismissal of charges against all defendants where evidence of guilt falls below some minimum threshold.⁴⁵ Several scholars have urged various reforms intended to protect innocent defendants by inducing prosecutors to screen out cases in which the probability of conviction is low. Ronald Wright and Marc Miller have argued that "hard" prosecutorial screening of weak cases would diminish the pressure on prosecutors to plea bargain and would protect innocent defendants who otherwise might feel compelled to enter guilty pleas.⁴⁶ Similarly, proposals to impose a fixed plea discount, plea-based ceilings, or a partial ban on plea bargaining are all predicated at least in part on the idea that discouraging pleas in weak cases by limiting the size of the plea discount not only will force more of those cases to trial, and thus lead to more acquittals, but also induce prosecutors to dismiss those cases rather than try them.⁴⁷

Nonetheless, when offered an opportunity to plead guilty to a misdemeanor with a guarantee of no prison, the defendant took the deal. *Id.* at 334. The judge in the case commented that had it been him, innocent or not, "he 'probably would have jumped at' the plea offer." *Id.*

44. This is especially true at the outer margins, where outcome uncertainty is less a product of evidence and more a product of irrational decisionmakers or trial surprises. As one defense lawyer and former prosecutor explained, "Sometimes trials bring surprises—surprises that turn the flimsiest cases into convictions 'Police lie, witnesses lie. It's a fact of life. So you get the best deal you can and you get out of there.'" *Id.*

45. That is, below some minimum threshold that itself is higher than the mere probable cause that is constitutionally required to arrest or charge. Some scholars argue—largely based on assumptions of public confidence—that fewer convictions accompanied by stiffer sentences in such cases is preferable to many convictions with light sentences. *See* Wright, *supra* note 35, at 111 (arguing that it is "better to have confidence in the accuracy of two convictions with sentences near full strength, along with one dismissal or acquittal, than to negotiate for three convictions that produce weaker sentences and inspire weaker confidence").

46. Ronald Wright & Marc Miller, *The Screening/Bargaining Tradeoff*, 55 STAN. L. REV. 29, 32 (2002).

47. The assumption here is that prosecutors will replace time-intensive weak cases with

These proposals, unlike abolition per se, would in fact marginally improve the position of some innocent defendants. Indeed, anything that increases the cost of prosecuting weak cases would reduce the number of convictions of innocent defendants, not only because any reduction in total convictions will necessarily reduce the number of convictions of innocent defendants, but also because innocent defendants presumably make up a larger fraction of weak cases.⁴⁸ But, while reducing the number of weak cases prosecuted may be a desirable goal, there are many ways to accomplish it that have nothing to do with plea bargaining. For instance, an evidentiary rule barring convictions on the basis of the uncorroborated testimony of a single eyewitness would also weed out some of the weakest cases on the docket.⁴⁹ Greater protections against the use of deceptive interrogation tactics would limit the number of false confessions.⁵⁰ More vigorous sufficiency of the evidence review by

more pleadable, and thus more efficient, strong cases. See, e.g., Russell D. Covey, *Fixed Justice: Reforming Plea Bargaining With Plea-Based Ceilings*, 82 TULANE L. REV. 1237, 1268–81 (2008) (describing a proposal to limit plea discounts through enforcement of plea-based ceilings); Oren Gazal-Ayal, *Partial Ban on Plea Bargains*, 27 CARDOZO L. REV. 2295, 2314 (2006) ("[T]he most substantial effect of the partial ban is not an increase in the rate of trials—it is a change in the prosecutor's choice of cases"); *id.* at 2324 (arguing that "[w]eak cases should be dismissed, not settled").

48. That still does not mean that most of the benefits of eliminating weak cases will flow to innocent defendants. Assuming that most persons charged in criminal cases are, in fact, guilty, it is probably also true that most defendants, even in weak cases, are guilty. Reducing a prosecutor's ability to charge weak cases, therefore, will still benefit many more guilty defendants than innocent ones.

49. See Steven B. Duke et al., *A Picture's Worth a Thousand Words: Conversational Versus Eyewitness Testimony in Criminal Convictions*, 44 AM. CRIM. L. REV. 1, 10 (2007) (stating that "the testimony of a single eyewitness, like the testimony of a single conversational witness, may suffice to convict without corroboration"). Courts have recognized the special dangers in single-eyewitness cases, but few jurisdictions bar convictions in such cases. See, e.g., *United States v. Telfaire*, 469 F.2d 552, 558 (D.C. Cir. 1972) (upholding conviction in a robbery case in which the sole evidence against the defendant was testimony of one eyewitness, but urging the trial court to give a special jury instruction cautioning jurors in such cases). Experimental research has cast the efficacy of cautionary jury instructions into serious doubt. See Richard A. Wise et al., *A Tripartite Solution to Eyewitness Error*, 97 J. CRIM. L. & CRIMINOLOGY 807, 832 (2007) (discussing a study showing that "*Telfaire* instructions either did not affect or even decreased the participant's sensitivity to the witnessing and identification conditions in the simulated trials").

50. See Miriam S. Gohara, *A Lie for a Lie: False Confessions and the Case for Reconsidering the Legality of Deceptive Interrogation Techniques*, 33 FORDHAM URB. L.J. 791, 794 (2006) (arguing that trickery and deception in interrogation is a significant cause of false confessions and should be restricted).

appellate courts would discourage prosecutors from bringing cases in which the evidence of guilt was marginal.⁵¹

There is no doubt that fewer innocent persons will be falsely convicted if the number of weak cases is decreased, but that is not a reason by itself to place limits on plea bargaining rather than, say, increase burdens of proof or define substantive crimes with greater precision. There are a wide variety of ways to winnow out weak cases, many of which should be considered. But any reform that focuses on reducing weak cases must frankly acknowledge the costs of doing so: Anything that makes it harder to convict some innocent suspects by raising the bar to convictions will also make it harder to convict guilty suspects as well. As a result, large numbers of guilty suspects will avoid prosecution and punishment. Undoubtedly, many reforms that would benefit innocent defendants—such as adopting more demanding evidentiary standards—have not been adopted because politicians, policymakers, and the general public do not wish to pay that price.

3. Improved Pretrial Discovery

Under present law, criminal defendants have little right to obtain meaningful pre-plea discovery.⁵² Unlike civil discovery, the "criminal discovery rules are not designed to inform a defendant fully of the case against him."⁵³ In busy urban systems, what discovery defendants do obtain often amounts to little more than the police reports describing their alleged crimes.⁵⁴

51. See Keith A. Findley & Michael S. Scott, *The Multiple Dimensions of Tunnel Vision in Criminal Cases*, 2006 WIS. L. REV. 291, 355 (2006) (criticizing the "exceptional deference appellate courts exhibit when reviewing sufficiency of the evidence to convict"). Sufficiency of the evidence claims are extremely hard for convicted defendants to win, even if they are, in fact, actually innocent. See Brandon L. Garrett, *Judging Innocence*, 108 COLUM. L. REV. 55, 112 (2008) (stating that, of sixty defendants who were ultimately exonerated and whose appeals included a sufficiency claim, only one prevailed on that ground). Indeed, to the extent that abolition of plea bargaining does reduce the innocence problem, it does so in precisely this way. Like hard screening and caps, abolition would increase incentives to screen cases more carefully because every additional case represents an incremental and fixed tax on scarce resources, which utility-maximizing prosecutors are assumed to wish to conserve for cases with the highest likelihood of success.

52. See *Weatherford v. Bursey*, 429 U.S. 545, 559 (1977) ("There is no general constitutional right to discovery in a criminal case . . .").

53. John G. Douglass, *Fatal Attraction? The Uneasy Courtship of Brady and Plea Bargaining*, 50 EMORY L.J. 437, 452–53 (2001).

54. See BOGIRA, *supra* note 24, at 35 (describing a prosecutor who provides "discovery" to public defenders, and "as usual, this 'discovery' consists mostly of police reports regarding

Even where significant discovery is mandated, the law often does not require that it be provided prior to guilty plea hearings. *Brady v. Maryland*,⁵⁵ for instance, requires prosecutors to turn over material exculpatory evidence at or before trial, but omits any mention of a requirement to provide exculpatory evidence prior to entry of a guilty plea.⁵⁶ As a result, the scope of prosecutorial obligations to produce such discovery during plea bargaining remains in doubt.⁵⁷ In recognition of this, several commentators have urged that defendants' rights to such pre-plea discovery be enhanced, arguing that enhanced discovery is critical to informed plea bargaining.⁵⁸ In a system in which the vast majority of criminal cases are resolved with guilty pleas, discovery obligations that are not triggered until the eve of trial—if at all—are likely to have only a minimal practical impact.

If the goal is to sharpen trial shadows and thereby bring plea bargain outcomes closer to trial outcomes, these proposals are unassailable. By more fully educating defendants about "[m]ajor procedural uncertainties, such as the questionable admissibility of significant evidence under the state's evidentiary

the alleged robbery"); *id.* at 158 (describing a "double-filing system" that permitted police detectives to shield notes containing exculpatory information in a "street file," while producing a sanitized "official file" in discovery, thereby evading *Brady* obligations).

55. See *Brady v. Maryland*, 373 U.S. 83, 87 (1963) ("We now hold that the suppression by the prosecution of evidence favorable to an accused upon request violates due process where the evidence is material either to guilt or to punishment, irrespective of the good faith or bad faith of the prosecution.").

56. *Id.* at 86–91.

57. In *United States v. Ruiz*, the Supreme Court cast substantial doubt on the claim that pre-plea bargaining discovery is mandated by the Constitution, holding that the Constitution does not require prosecutors to disclose impeachment information relating to any informants or other witnesses prior to the defendant's entry of a guilty plea, and prosecutors do not violate the Constitution when they predicate plea offers on waiver of discovery. *United States v. Ruiz*, 536 U.S. 622, 623–24 (2002). It remains an open question, however, whether defendants who plead guilty are entitled to any *Brady* material. See, e.g., *United States v. Coker*, 514 F.3d 562, 570 (6th Cir. 2008) (noting that "[p]rior cases have explicitly declined to decide whether and to what extent a federal defendant who pleads guilty may raise a *Brady* claim") (citing *United States v. Ross*, 245 F.3d 577, 583 (6th Cir. 2001)); see also *United States v. Matthews*, 168 F.3d 1234, 1242 (11th Cir. 1999) ("We do not need to decide today, however, whether a guilty plea waives a defendant's claim under *Brady*, or . . . whether the *Brady* material must be known to the prosecution before the plea or merely before sentencing.").

58. See, e.g., Douglass, *supra* note 53, at 517 (arguing for strengthened "pre-plea procedures aimed at promoting disclosure"); Steven L. Friedman, *Preplea Discovery: Guilty Pleas and the Likelihood of Conviction at Trial*, 119 U. PA. L. REV. 527, 529–35 (1971) (arguing for enhanced pre-plea discovery to enable defendants to better assess the likelihood of conviction). See generally Eleanor J. Ostrow, *The Case For Preplea Disclosure*, 90 YALE L.J. 1581 (1981).

rules,"⁵⁹ enhanced pre-plea discovery procedures would permit defendants to better calculate their chances at trial and thereby make more informed plea bargaining choices.⁶⁰ Where prosecutors have admissible evidence that exculpates the defendant, requiring pre-plea production of such evidence will permit some innocent defendants to reject or favorably renegotiate pleas.

Improved pre-plea discovery, however, will not always benefit innocent defendants—and certainly will not benefit only innocent defendants. In some cases, more comprehensive discovery of inculpatory evidence might actually increase the likelihood that innocent defendants will plead guilty. Trial shadow theory predicts that plea bargains should be reached where rational parties agree about case value and can share the resource savings of forgoing trial.⁶¹ By educating defendants about case values, robust discovery should bring the parties case valuations closer together and facilitate negotiated outcomes. That result, however, will sometimes cause innocent defendants to plead guilty when they otherwise might not have. For instance, an individual might know she is innocent of an alleged crime. Absent knowledge of the state's evidence, she might rationally contest the false allegations. If the state's case were predicated on the perjurious or mistaken testimony of an eyewitness to the crime, and she is convinced after receipt of discovery that a jury might credit that testimony, however, she might enter a guilty plea that she otherwise would have declined. Such innocent defendants are better off as a result of improved discovery in the sense that it induces them to negotiate a punishment-reducing plea. But the price of the plea they negotiate will incorporate the distorting effects of the false testimony and wholly ignore her private information of innocence. A better outcome would be a plea bargain discounted to reflect her private information of innocence or, better still, dismissal of the charges altogether.

At the same time, more robust discovery rules will permit some guilty defendants to obtain (arguably) unjustifiably good deals. Informing defendants about procedural obstacles like missing witnesses, unconstitutional searches

59. Thomas R. McCoy & Michael J. Mirra, *Plea Bargaining as Due Process in Determining Guilt*, 32 STAN. L. REV. 887, 933 (1980).

60. *See, e.g., id.* (arguing that pre-plea discovery enhances accuracy of pleas because many defendants do not know all relevant facts, or are compelled to plead guilty by large sentence differentials motivated by weaknesses in the case of which they are unaware); Kevin C. McMunigal, *Disclosure and Accuracy in the Guilty Plea Process*, 40 HASTINGS L.J. 957, 968–97 (1989) (same); Schulhofer, *supra* note 22, at 1998 (advocating more expansive pretrial discovery "so that negotiating parties could more accurately estimate ex ante the likelihood of conviction at trial").

61. *See* Bibas, *supra* note 1, at 2466 (describing the incentives for plea bargaining underlying trial shadow theory).

that undermine the admissibility of evidence, unlawful interrogation methods, and the like, of which they otherwise are unaware, helps guilty defendants as well as innocent ones avoid punishment.⁶²

In short, the enhancement of pre-plea discovery would help to conform plea bargain outcomes to trial outcomes and is therefore justified under trial shadow theory. Under that rubric, there is little justification for withholding such evidence from a defendant, since markets and negotiation work best under conditions of full information. But enhanced discovery may not increase the disparity between the plea rate of guilty and innocent persons, and it might even have the opposite effect. By making it easier to make accurate predictions about trial outcomes based on a shared evidentiary data set, enhanced pre-plea discovery may increase both the number of guilty defendants who hold out for trial and the number of innocent defendants who enter guilty pleas.

Optimally, the plea-bargaining system should not merely replicate expected trial outcomes, it should resolve cases with greater accuracy and normative validity than a system of trials alone. To achieve that goal, however, plea bargaining must itself generate information that would not otherwise be available. In the next Part, I will suggest that the innocence problem might best be addressed by modifying the bargaining process to better incorporate the defendant's private information into the price structure. Only by increasing the total amount of information in the system can the innocence problem be diminished.

III. The Innocence Problem as a Signaling Defect

As game theorists have observed, although there is no direct way to tell whether any disclosure of private information is truthful, sometimes one can indirectly infer the content of private information "from the actions that the informed party takes."⁶³ The communication of information in this way is referred to in economics as "signaling."⁶⁴ As one leading text on game theory

62. Courts routinely refuse to overturn guilty pleas where prosecutors withheld information highly relevant to the probability of conviction at trial, but not relevant to the defendant's substantive guilt of the charges. *See, e.g.,* *People v. Jones*, 375 N.E.2d 41, 42 (N.Y. 1978), *cert. denied*, 439 U.S. 846 (1978) ("[The] defendant was not denied due process when the District Attorney's office did not disclose during plea negotiations that it had received information that the complaining witness had died.").

63. DOUGLAS G. BAIRD ET AL., *GAME THEORY AND THE LAW* 122 (1994).

64. *See id.* at 123 (discussing the concept of signaling).

explains, "[s]ignaling takes place when those who possess nonverifiable information can convey that information in the way they choose their actions."⁶⁵

For signaling to occur, at least one party must have private information and an economic incentive to communicate it. Those conditions are typical of plea bargaining.⁶⁶ Defendants usually know if they are innocent or guilty, and innocent defendants have powerful incentives to communicate their private information to prosecutors. The ease by which guilty defendants can mimic nonverifiable innocence claims made by innocent defendants, however, prevents prosecutors from taking those claims seriously.⁶⁷ Observing this, Dean Robert Scott and Professor William Stuntz characterize plea bargaining's innocence problem as, at bottom, a signaling defect.⁶⁸

This is not to say that private information does not have an impact on the *disposition* of cases. The defendant's subjective knowledge of guilt or innocence affects the bargaining process and thus case outcomes, in several ways. First, defendants who perceive themselves to be guilty are probably more likely to discount their chances of acquittal at trial, and defendants who perceive themselves to be innocent may be more likely to believe they will be acquitted.⁶⁹ Moreover, as Scott and Stuntz argue, factual knowledge of guilt or innocence should improve the defendant's ability to forecast future evidentiary developments.⁷⁰ Innocent defendants know that newly discovered evidence will be more likely to exonerate than to incriminate them.⁷¹ In addition,

65. *Id.*

66. See IAN MOLHO, *THE ECONOMICS OF INFORMATION: LYING AND CHEATING IN MARKETS AND ORGANIZATIONS* 10 (1997) (noting that, absent an effective signal, the underlying method of conveying private information loses value where that information is otherwise nonverifiable).

67. See *id.* (comparing the denials of a criminal defendant to common lies that one hears often in everyday life). As Molho explains, using the classic "market for lemons" problem to illustrate, "the consequence of having 'liars' present in the population (who may try to pass off bad cars as good, say) is ultimately to degrade the information content of the communication ('this is a good car'), conceivably up to the point where it becomes meaningless." *Id.*

68. See Scott & Stuntz, *supra* note 12, at 1944 (discussing the relationship between signaling and plea bargaining, specifically the use of prior criminal conduct as a signal).

69. See McCoy & Mirra, *supra* note 59, at 924 ("[A]n individual defendant's knowledge of his own guilt or innocence seems likely to skew his perception of his chances of acquittal at trial.").

70. See Scott & Stuntz, *supra* note 12, at 1937 (explaining that "[w]hat [the Defendant] actually did and thought thus bears powerfully on any prediction about what evidence is likely to come out during the course of further investigation" and thus can be "critical" to the accuracy of predictions about probability of conviction).

71. See David Bjerk, *Guilt Shall Not Escape or Innocence Suffer? The Limits of Plea Bargaining When Defendant Guilt is Uncertain*, 9 AM. L. & ECON. REV. 305, 306 (2007) (arguing that "any new evidence that arises in the time leading up to trial generally strengthens

innocent defendants who testify at trial may be more convincing than guilty defendants. Their stories will naturally cohere because they need not fabricate exculpatory, and potentially disprovable, lies on the witness stand.⁷² Based on that information, they should calculate a more accurate, and lower, probability of conviction than the prosecutor. In addition, private information about guilt or innocence may affect a defendant's desire to go to trial. Some guilty defendants are no doubt moved by remorse to plead guilty irrespective of their expected chances at trial. Innocent defendants may also be more motivated to contest charges than guilty defendants among the general pool of defendants with the same statistical likelihood of conviction.⁷³ These factors suggest that private information affects how individuals respond to plea offers at any given price.⁷⁴ Innocent defendants, knowing they are innocent, may be more likely to reject plea offers, while guilty defendants might be more likely to accept those offers.⁷⁵

But none of these factors suggests that the private information has much effect on plea prices themselves. Prosecutors cannot infer anything reliable about the defendant's private information based on their responses to plea offers because too many other factors influence plea evaluation. As Dean Scott and Professor Stuntz observe, the signal sent by innocent defendants through

the case against a guilty defendant, but generally weakens the case against an innocent defendant").

72. See Daniel J. Seidmann & Alex Stein, *The Right to Silence Helps the Innocent: A Game-Theoretic Analysis of the Fifth Amendment Privilege*, 114 HARV. L. REV. 430, 448–49 (2000) ("Accounts furnished by innocent suspects are generally more credible than those provided by guilty defendants."). Innocent defendants may also be better able to provide detailed accounts of the facts, which has been found to increase credibility. See Jansen Voss, *The Science of Persuasion: An Exploration of Advocacy and the Science Behind the Art of Persuasion in the Courtroom*, 29 LAW & PSYCHOL. REV. 301, 309 (2005) (finding that possessing "a command of details" increases one's credibility).

73. On the other hand, in practice, where defendants resist acceptance of plea offers based on assertions of actual innocence, the frequent response is not rejection of the offer, but rather pressure from the defendant's lawyer to take the discounted offer notwithstanding the accuracy of the defendant's assertion. See DEBRA S. EMMELMAN, JUSTICE FOR THE POOR: A STUDY OF CRIMINAL DEFENSE WORK 114 (2003) (noting that defense lawyers must "continue to advise their clients about the virtues of plea bargaining" even when the client desires trial).

74. Unbridgeable differences in case evaluation, therefore, might marginally improve the accuracy of the sorting process in plea bargaining. See McCoy & Mirra, *supra* note 59, at 924 (noting that "[defendants] who were guilty would be statistically overrepresented in the group who chose the plea and those who were innocent would be statistically overrepresented in the group who rejected the plea in favor of a trial").

75. See WELSH S. WHITE, LITIGATING IN THE SHADOW OF DEATH: DEFENSE ATTORNEYS IN CAPITAL CASES 169 (2006) (finding that innocent defendants are more likely to reject plea offers than guilty defendants).

their higher price demands is indistinguishable from the comparatively higher price demands of those who heavily discount the future or who are less risk-averse.⁷⁶ Since criminals likely are (almost by definition) less risk-averse and heavy discounters, the "innocence signal" implicit in the rejection of a plea offer is almost certain to be imperceptible to prosecutors.⁷⁷ Accordingly, prosecutors cannot adjust plea prices based on unverified defendant signals of innocence. The plea bargaining "game" thus results in a "pooling equilibrium"⁷⁸ in which defendant's private information is not incorporated into the price of the plea.⁷⁹

This might not be the case if there were a mechanism to leverage privately held information so that it had a more direct impact on plea pricing. Scott and Stuntz note that perjury sanctions, in theory, could be used to provide such leverage.⁸⁰ If guilty defendants who tried to send false innocence signals could be identified and adequately punished for lying, they would not falsely claim innocence. In this ideal world, only innocent defendants would claim innocence. Prosecutors could trust that signal and adjust their plea offers accordingly. Of course, as Scott and Stuntz point out, there is no practical way to enforce truth-telling with perjury sanctions. First, the only way to verify the truthfulness of an innocence claim would be to conduct what is, in essence, a trial.⁸¹ Doing so obviates any resource savings from plea-bargaining and therefore nullifies its benefits. Guarantees that are too costly to enforce will not be enforced. Second, the penalties for perjury would have to approach or

76. See Robert E. Scott & William J. Stuntz, *A Reply: Imperfect Bargains, Imperfect Trials, and Innocent Defendants*, 101 YALE L.J. 2011, 2012 (1992) ("[It is likely that] innocent defendants as a class are significantly more risk averse than guilty defendants as a class, [so] a prosecutor's failure to internalize a defendant's private information will cost the prosecutor nothing because the defendant, even if innocent, will take the deal anyway.").

77. See *id.* (arguing that innocent defendants, on the whole, are more risk-averse than guilty defendants).

78. See Scott & Stuntz, *supra* note 12, at 1911 (noting that "strategic impediments to efficient bargains lead to a pooling of guilty and innocent defendants," which in turn "leads predictably to innocent defendants being offered (and taking) the same deals as guilty ones").

79. Not only does the defendant's private knowledge not affect plea bargaining prices, it does not affect trial outcomes in the event the case is tried. In fact, precisely because the defendant lacks any way to verify his private information at trial, private information is irrelevant to a plea pricing system predicated on anticipating trial outcomes. See Seidmann & Stein, *supra* note 72, at 458 (noting that "mere augmentation of private information—that is, of any information with uncertain credentials—could not increase the factfinder's accuracy").

80. See Scott & Stuntz, *supra* note 12, at 1943 (considering the effectiveness of raising the penalty for lying).

81. *Id.*

exceed those of the substantive crime charged to have any likelihood of inducing truth-telling, an unrealistic prospect.⁸² Third, only some innocence claims would be verifiable even with trial-like verification methods. Because nonverifiable innocence claims have no independent value to prosecutors, an enhanced perjury regime would still be gameable by savvy guilty defendants at least some of the time. For all of those reasons, enforced truth-telling is not itself a practical way to verify innocence claims.⁸³ Scott and Stuntz thus conclude that there is no feasible way to incorporate private information into plea pricing.⁸⁴ But, perhaps Scott and Stuntz's conclusion is overly pessimistic. In the next section, this Article considers ways in which an alternative signaling mechanism might be devised.

A. Signaling Mechanisms

The literature on signaling provides numerous examples of signaling mechanisms that can, under certain assumptions, effectively communicate private nonverifiable information between parties based on players' actions. Michael Spence's leading analysis of signaling illustrates the basic features of an effective signaling mechanism.⁸⁵ Spence models a labor market in which previously obtained educational credentials function as informational signals regarding otherwise unverifiable applicant attributes.⁸⁶ The model assumes that there are two types of workers: high-productivity workers and low-productivity workers.⁸⁷ Employers are willing to pay higher salaries to high-productivity workers than to low-productivity workers, and all workers, of course, prefer

82. *See id.* at 1943–44 (explaining that penalties for lying must be set so high to work that they "would conflict with basic criminal law norms of proportionality"); *see also* Seidmann & Stein, *supra* note 72, at 460 ("To deter perjury, the punishment for lies must exceed the punishment imposed for the underlying substantive crime."). On the other hand, it could be argued that for perjury sanctions to impact a defendant's choice to lie, the added punishment need not outweigh the punishment for the charged crime. It would be enough if the marginal risk of extra punishment for perjury tipped the balance in favor of accepting a discounted guilty plea rather than hazard the risk of conviction and enhanced punishment.

83. *See* BAIRD ET AL., *supra* note 63, at 122 (stating that private, nonverifiable information cannot be compelled through legal rules).

84. *See* Scott & Stuntz, *supra* note 12, at 1943 ("None of these devices [to incorporate private information into plea pricing] can be relied upon by prosecutors and defendants to overcome their information barriers.").

85. *See generally* Michael Spence, *Job Market Signaling*, 87 Q.J. ECON. 355 (1973).

86. *Id.* at 361–68.

87. *Id.* at 361.

high salaries to low ones.⁸⁸ To win high-salary job offers, applicants must convince employers that they are high-productivity (or "good") types and not low-productivity (or "bad") types.⁸⁹ All applicants tell employers that they are good types, so without further information, employers disregard all such statements.⁹⁰ Good types thus need to find a way to reliably signal their type to employers, and employers need a way to reliably separate good types from bad.⁹¹ Spence's model then demonstrates how educational attainment could serve as a signal of an employee's type in such a job market.⁹²

In developing the job market signaling model, Spence makes the following assumptions about education. First, Spence assumes that obtaining an education is costly—it requires substantial investment.⁹³ Second, the costs of obtaining the educational attainment necessary to constitute the signal are assumed to be differentially distributed.⁹⁴ That is, the costs of obtaining the education signal are assumed to be high for low-productivity workers but low for high-productivity workers. Because of these differential costs, it is rational for high-productivity workers—but not low-productivity workers—to invest in education.⁹⁵ Third, these differential costs are logically related to productivity.⁹⁶ If educational attainment were not positively correlated with high productivity, employers would lose interest in hiring such persons, and educational attainment would fail as a signal.⁹⁷ Finally, education can serve as an effective signaling device only if the increase in salary sufficiently compensates high-productivity workers so that the benefits of obtaining the education are not outweighed by the costs.⁹⁸ As Spence notes, however,

88. *Id.* at 362.

89. *Id.*

90. *Id.* at 356.

91. *See id.* at 357 (discussing "the endogenous market process whereby the employer requires (and the individual transmits) information about the potential employee, which ultimately determines the implicit lottery involved in hiring, the offered wages, and in the end the allocation of jobs to people and people to jobs in the market").

92. *See id.* at 361–68 (analyzing education as a signal).

93. *Id.* at 358.

94. *Id.* at 362.

95. *Id.*

96. *Id.* at 358.

97. *See id.* at 367 (noting that where there is an "absence of any correlation between educational costs and productivity, education could *never* be an effective signal, in the sense of conveying useful information, in an equilibrium" where such conditions exist).

98. *Id.* at 362.

depending on the specific values assumed, job market signaling may lead either to separating, semi-pooling, or pooling equilibria.⁹⁹

Spence's job-market signaling model functions based on past player commitments,¹⁰⁰ but signaling mechanisms need not be backward looking; they could turn on present conduct as well. In capital markets, for instance, investors might have difficulty distinguishing good and bad investment risks. Some firms make high profits, others do not, but all firms hold themselves out as high-profit companies, and it might be difficult or costly to verify such claims. If high dividend payments cost less for a company with high profits than for a company with low profits, however, then payment of high dividends might effectively signal that the firm is a high profit company.¹⁰¹ The signal is reliable because of the differential costs: It cannot easily be copied by low profit firms.¹⁰²

Indeed, it turns out that there are a wide variety of ways in which signaling costs can produce effective signaling mechanisms. Reliability in some cases might be indicated by the payoff structure confronting potential signalers. Reputation, for example, exhibits such a payoff structure.¹⁰³ Developing a

99. *See id.* at 361–68 (analyzing the effect of different values for education as a job market signal).

100. *See id.* at 360 (noting that employers' valuation of a specific signal is based on its past performance).

101. *See* Joseph E. Stiglitz, *Information and Capital Markets*, in FINANCIAL ECONOMICS: ESSAYS IN HONOR OF PAUL COOTNER 96 (Cathryn M. Cootner, Paul H. Cootner & William F. Sharpe eds., 1982) (developing a signaling model in which payment of dividends signals profitability to uninformed investors). Similarly, it has been argued that the financial structure of a firm may serve as a signal of expected profits, where companies with high debt rates signal confidence in high future cash flows sufficient to profitably pay off the debt, and high equity ratios signal lower expected profits. *Id.* at 120–21 (citing Hayne E. Leland & David H. Pyle, *Information Asymmetries, Financial Structure, and Financial Intermediation*, 32 J. FIN. 371 (1977); Stephen A. Ross, *The Determination of Financial Structure: The Incentive-Signalling Approach*, 8 BELL J. ECON. 23 (1977)).

102. Reputation is also a signaling mechanism that functions by sending information to others about one's type. Parties will seek to maintain a reputation for being a good type so that others will transact with them under conditions of uncertainty. When the long-term interests of a market participant exceed potential short-term gains, the need to maintain a good reputation, and thereby protect one's access to the market, provides strong guarantees to potential transaction partners against short-term cheating. Even though a claim is nonverifiable at the time of transaction, eventual discovery of the claim's falsity will undermine the party's ability to make credible claims in the future with the deceived party or, more generally, with all other parties who have access to reputational information. *See* BAIRD ET AL., *supra* note 63, at 159–87 (discussing reputation as a signal in the context of repeating games).

103. *See* Ronald J. Gilson & Robert H. Mnookin, *Disputing Through Agents: Cooperation and Conflict Between Lawyers in Litigation*, 94 COLUM. L. REV. 509, 534–50 (1994) (discussing the incentives for lawyers to develop a cooperative reputation as a signal for clients).

good reputation requires players to visibly forgo opportunities to cheat. Players that forgo those opportunities lose out on the short-term benefits they otherwise might have earned. As a result, reputation is costly to develop and maintain. Precisely because the good-reputation player has invested substantial resources in the cultivation of reputation—an investment she logically will seek to protect—her signals are perceived as trustworthy (at least as long as the short-term benefits of deception do not exceed the sunk costs of cultivating a reputation).¹⁰⁴ Because of this last point, signaling models in which reputation provides an effective signal assume repeat play.¹⁰⁵

More broadly, any credible commitment of resources that "locks in" a player's strategy functions as a reliable signal. A manufacturer that makes large capital investments in production capacities, for example, signals to potential competitors its commitment to compete in a particular market since, having made the investment, it would be plainly irrational subsequently to abandon it. Such sunk-cost commitments can effectively dissuade competitors from encroaching into the market, because competitors know that the incumbent manufacturer's committed investment limits potential profits available from the entry of a new market competitor.¹⁰⁶ Similarly, sunk-cost signals can be exchanged between manufacturers and consumers of products through advertising. For instance, assume that a company makes an inexpensive, disposable razor and is considering an expensive national advertising campaign to promote the razor. To justify the advertising costs, consumers must make repeat purchases of the razors, and consumers will only buy the product a second time if the product in fact is high quality. If the economics of the decision to advertise is common knowledge, the very act of advertising signals the manufacturer's private information that the razor is high quality. Because advertising is only economical to the seller if a consumer in fact makes repeat purchases, and the consumer will only make repeat purchases if the product is high quality, the consumer can interpret the seller's willingness to advertise as a signal that the seller's product is high quality.¹⁰⁷ After all,

104. All bets are off, economically speaking, when the stakes rise sufficiently high. Indeed, this is the basic ploy used by card sharks and pool sharks who lull their prey with false ineptness in small stakes early rounds, cultivating an appearance (reputation) that makes others willing to play and then breaching in the "money game."

105. See BAIRD ET AL., *supra* note 63, at 167 (noting that the incentive to cultivate reputation disappears when the competitors "act as if they are playing a one-shot game").

106. See *id.* at 58–62 (using the example of an established cement plant in a small town and an upstart company to demonstrate the usefulness of signaling through sunk cost commitments).

107. See MOLHO, *supra* note 66, at 95 (noting that under the specified conditions "it is

rational manufacturers would not choose to advertise if they knew their razors were low quality, and knowing this, consumers can buy the razor with greater assurance that their small investment will be rewarded. The buyer's marginal willingness to take a small risk (buy a razor of unproven quality) is increased by the buyer's observation of the seller taking a larger risk in aggressively marketing the product.

Although signals must always be costly, an effective signaling mechanism need not be based on sunk-cost investments. Some signaling mechanisms work because they are backed by enforceable promises about future conduct. A warranty is a kind of insurance contract that shifts the risk of product failure, at least in part, to the producer. In commercial markets, for instance, warranties signal a seller's private information about the quality of a particular product to buyers who may be unable independently to test or verify quality. It may be relatively cheap for the seller to offer a warranty at the time of sale, but it is costly to have to perform on the warranty at a future date. The signaling value of the warranty derives from the fact that future costs are borne differentially. If the costs of offering a warranty are inversely proportional to the quality of the warranted product, then only the producers of high-quality goods will offer the warranties. Buyers need not seek additional information about the quality of the product and can assume that any producer who offers a warranty is selling a high-quality product. At least under those conditions, the warranty effectively separates high- and low-quality producers.¹⁰⁸

In short, there are a variety of ways to construct an effective signaling mechanism. Risks can be transferred among parties (warranties), prior investments that reflect private information about type can be made (education), or actions consistent only with type, and that are difficult to fake, can be currently taken (high-dividend payments). In many instances, signaling is possible only because the players are engaged in multiple rounds of play (reputation, advertising).

Although they differ in operative structure, effective signaling mechanisms appear to share four important characteristics. First, signaling mechanisms do not work unless the signals sent are reliable. For a signal to be reliable, it must be costly.¹⁰⁹ Cost-free signals can be sent by anyone interested

possible to construct a separating equilibrium where only the sellers of high quality brands advertise").

108. *See id.* at 30 (discussing the role of warranties as providing "a credible way of informing the buyer about quality" and thus an example of a potentially effective signaling device); *see also* BAIRD ET AL., *supra* note 63, at 124 (same).

109. *See* AMOTZ ZAHAVI & AVISHAG ZAHAVI, *THE HANDICAP PRINCIPLE*, at xiv ("[I]n order

in communicating the message, and as such, are worthless. Second, not only must signals be costly to send in general, the cost of the signal must vary with its content.¹¹⁰ It is the differential cost of signaling that guarantees the signal's reliability.¹¹¹ Third, signals cannot be arbitrary. The nature of the signal must be logically related to the content sought to be communicated. That is, it must be well suited to "convey the specific message it carries."¹¹² Fourth, signaling must be economical. If signaling costs exceed signaling benefits even for those whose signaling costs are cheapest, then all parties will be better off forgoing signaling.¹¹³

to be effective, signals have to be reliable; in order to be reliable, signals have to be costly."). The Zahavis developed an evolutionary theory of signaling. *See id.* ("The investment that animals make in signals is similar to the 'handicaps' imposed on the stronger contestants in a game or a sporting event: for example . . . the extra weight the swifter race horse must carry . . .").

110. *See* Assaf Likhovski, *Argonauts of the Eastern Mediterranean: Legal Transplants and Signaling* 8 (Tel Aviv University School of Law Working Paper Series, 2008), available at <http://ssrn.com/abstract=1098179> ("In order to prevent those not possessing the quality from mimicking the signal, signals may sometimes be costly, that is there will be a link between the signal's cost and the quality it signifies in a way which would make the signal too costly to produce by deceivers.").

111. *See* ZAHAVI & ZAHAVI, *supra* note 109, at 57–59 (explaining that if the "cost of a signal is reduced to the extent that every individual can use it equally well, then the signal can no longer reveal differences in the quality or motivation of individuals"); *id.* at 59 ("If there is no cost, nothing prevents cheaters from using a signal to their benefit and to the detriment of the receivers, and that signal will lose its value as a signal."); *see also* Spence, *supra* note 85, at 358 (noting that a negative correlation of signaling costs and the subject that is signaled, which in Spence's study is the productive capability of employees, "is a prerequisite for effective market signaling to take place"). The prerequisite of differential costs is referred to in the signaling literature as the "single crossing property." As Molho explains,

[T]he single crossing property is that from any given situation it costs more for low quality types to increase the signal than it does for high quality types, whether the signal is a warranty on a used car, education, or whatever. Without this property the signal is bound to fail to communicate information, since there would be nothing to prevent low quality types from imitating high quality types in the signals they send.

MOLHO, *supra* note 66, at 79 n.7.

112. ZAHAVI & ZAHAVI, *supra* note 109, at 229. Zahavi and Zahavi elaborated on this point with the following:

The Handicap Principle states that the receiver of a signal has a stake in the signal's reliability, or accuracy, and will not pay attention to it unless it *is* reliable. Thus signals are not arbitrary; rather, each signal is the one best suited to reliably convey the specific message it carries. It follows that there must be a logical connection between the message and the signal.

Id.

113. *See* Spence, *supra* note 85, at 367 ("[E]ffective signaling depends upon not only the

It should also be noted that, in the language of game theory, different signaling regimes can give rise to different types of equilibria. An optimal signaling regime produces a separating equilibrium, that is, one in which receivers could differentiate signalers with perfect accuracy and treat each according to type.¹¹⁴ But separating equilibria are not the only possible type of equilibrium. Also possible are semi-separating (or semi-pooling) equilibria—in which some separating (and some pooling) occurs although the separating is not complete—and pooling equilibria—in which the signaling mechanism fails and the signalers wind up being "pooled," i.e., treated in a nondifferentiated manner.¹¹⁵

The critical question here is whether the conditions characteristic of plea bargaining permit the operation of a signaling mechanism that produces at least a semi-separating equilibrium. Certainly, many of the above-noted strategies are not readily adaptable to plea bargaining.¹¹⁶ Take the pre-commitment strategy. Criminal records function a bit like educational degrees in the job market and in that sense perform important signaling functions in plea bargaining. A criminal defendant's innocence claim almost certainly diminishes in credibility as the extent of the defendant's criminal history increases, particularly where the conduct reflected in that criminal history is similar to the newly alleged crime. But criminal history is too blunt an instrument to accurately signal private information about guilt or innocence in particular cases. If anything, it signals propensity to commit crime rather than conduct with respect to a particular historical incident. Indeed, the rules of evidence generally bar admission of criminal history at trial precisely because

negative correlation of costs and productivities, but also upon there being a 'sufficient' number of signals within the appropriate cost range.").

114. See MOLHO, *supra* note 66, at 78 (introducing the concept of a separating equilibrium in the context of job market signaling).

115. See *id.* (discussing the various levels of separation and pooling equilibria in the context of job market signaling).

116. Price itself is a signaling device. See *id.* at 43 (explaining how price can signal quality of products when information is imperfect and transaction costs are significant) (citing Asher Wolinsky, *Prices as Signals of Product Quality*, 50 REV. ECON. STUD. 647 (1983)). Bebchuk considers the signaling impact of settlement price in the context of settlement of civil disputes in which two parties must decide whether to settle or litigate a contract dispute. Lucian Arye Bebchuk, *Litigation and Settlement Under Imperfect Information*, 15 RAND. J. ECON. 404, 408 (1984). As Bebchuk explains, if we assume that both parties will agree to settle if the settlement price is lower than their expected costs after trial, but that the trial outcome is not certain because one of the parties possesses private information that will not be revealed until trial, a separating settlement offer will be possible, in which defendants with positive private information will reject the offer and defendants with negative private information will accept it. *Id.*

fact-finders all too easily, and thus potentially falsely, treat propensity as a strong, perhaps overwhelming, signal of guilt.¹¹⁷ Criminal history signals one's "type" in the sense of one's propensity to lawfulness, but it does not signal one's culpability as to specific charges.¹¹⁸ Unlike employers seeking employees with the propensity for hard work, criminal adjudication requires conclusions not about propensity but about specific historical instances.¹¹⁹

The warranty model also seems inapt. As Scott and Stuntz observe, the closest analogy to warranty in plea bargaining is enforced truth-telling through perjury sanctions.¹²⁰ The threat of harsh future penalties for perjury in theory could work as a kind of prosecutorial insurance policy, but enforced truth-telling is not workable in plea bargaining for the reasons identified by Scott and Stuntz.¹²¹ To obtain performance on the warranty, the buyer must adduce a demonstrably defective product before the seller will be willing to perform. That is relatively simple to do when a lawn mower or a car stops working. It is much less feasible in the event of an innocence claim that can only be tested by

117. See *Blakely v. Washington*, 542 U.S. 296, 318–19 (2004) (O'Connor, J., dissenting) (observing that character evidence such as criminal history "has traditionally been off limits during the guilt phase of criminal proceedings because of its tendency to inflame the passions of the jury").

118. This is why, as Josh Bowers explains, recidivist offenders are not only more likely to commit future crimes but are also more likely to be falsely accused of committing them. Josh Bowers, *Punishing the Innocent*, 156 U. PA. L. REV. 1117, 1119–21 (2008). Moreover, unlike employment decisions, the stakes in a criminal case involving serious charges are simply too high to expect much utility from reading a criminal history signal. First-time offenders have powerful incentives to misrepresent their private information of personal guilt, at least if charges are serious, and prosecutors know it.

119. For similar reasons, a defendant's reputation will not likely provide much valuable signaling information in the plea bargaining context. A criminal defense lawyer's concern with reputation, however, might make possible a different kind of signaling device. As Frank Easterbrook has noted, criminal defense attorneys might invest substantial capital in developing a reputation for truthfulness with prosecutors, and when both sides know this is true, their occasional unverifiable representations that a particular client is "innocent" could provide a useful and trustworthy signal to prosecutors. See Easterbrook, *supra* note 26, at 1971 ("A good part of the practice of many defense lawyers, especially in the period before indictment, is supplying information to prosecutors."). This assumes that criminal defense lawyers have private information not possessed by prosecutors. Typically, however, defense lawyers have no private information, merely what their clients give them—which they no more than prosecutors can verify. Defense attorneys intent on acting as intermediaries between innocent clients and prosecutors will thus have to develop their own mechanisms to distinguish between true and false innocence claims.

120. Scott & Stuntz, *supra* note 12, at 1942–44.

121. See *supra* notes 80–84 and accompanying text (explaining the reasons why enforced truth-telling does not work in the plea-bargaining context).

trial. In short, it will be too difficult to collect future payment for present breaches.

But reliance on criminal history and the threat of perjury sanctions are not the only possibilities. As I explain below, the key to creating an effective signaling mechanism in the context of plea bargaining is to reconceptualize it as a multiple-move game rather than a single-move game. As it turns out, strategies that are dominant in a single-move game might not be dominant in a game containing multiple moves. In the next section, this Article illustrates how a single-move game in which private information is irrelevant to settlement prices can be transformed with the addition of a second move in a way that makes the private information relevant.

B. The Two-Move Subwager Game

In this subpart, this Article seeks to illustrate a type of signaling mechanism with applications to plea bargaining. It does this by examining the impact of rule changes on a simple two-player card game, which works as follows. Gambler (*G*) is dealt four cards from a standard deck of playing cards, two of which are red and two are black. The cards are placed in a pile. *G* "wins" if the bottom card in his pile is red and "loses" if the bottom card is black. Thus, if *G* turns over the bottom card and it is red, he gets to keep his money. If the bottom card is black, however, he must pay the House (*H*) \$20. *H*, though, values not having to play out games, and is willing to accept 50% less from players to save "dealing costs." Accordingly, *G* can opt out of the game by paying a \$5 fee to *H*.¹²² One need not be a mathematician to see that a rational player seeking to minimize losses should pay the fee (and get out of this no-win game as quickly as possible). Gamblers who follow that strategy will, over the long run, fare better than gamblers who choose to play out the game in the hope of "winning."

Thus far, even though only half of the Gamblers would lose, there is no differentiation of strategy among would-be winners and losers. That outcome potentially changes when private information is introduced. Imagine the same rules as above, except that now *G*—and only *G*—is entitled to look at (and keep) the top card.¹²³ *G* may not show the card to *H*. Although *G* does not

122. Expected Payout (EP) = $20(.5) = 10$, discounted by $50\% = 5$.

123. In this game, the player gets to look at the top and not the bottom card. This information helps him assess his chances at the game, but it does not give him certain knowledge as to how the game will play out. This rule is intended to mimic the position that

know how his hand will play out, he has additional private information unavailable to *H*. If the top card in *G*'s possession is black, *G* now knows that his chance of winning (i.e., that the bottom card is red) is two-thirds. What can *G* do with this new information? He might try to renegotiate the settlement price, based on the fact that he now knows his chance of losing is only one-third rather than one-half. After all, the true "market value" of his settlement is now lower, given the lower odds that he will have to pay full price if the game is played out. Under the assumptions used so far, the "real" market price of settlement should be \$3.33.¹²⁴ But *G* will be unable to extract that price through direct negotiation. *H* won't drop its price simply because *G* represents that the card he holds is black. After all, if it credited his claim, every player would have an incentive to claim the top card was black, and without an ability to corroborate the claim, the House would be unable to differentiate honest and dishonest reports.

Alternatively, *G* could play out his hand. The private information assures *G* that the odds of winning are better than initially reckoned. However, *G* has no rational incentive to play out, because the risk of playing and losing is still too great (i.e., *G*'s expected payoff if the game is played out is \$6.66).¹²⁵ That is worse than the \$5.00 settlement offer currently on the table. Thus, even though *G* has private information that permits him to more accurately assess his chances of prevailing in the game, that private information does him no good. Settling by paying the fee remains the dominant strategy.

G's problem is that he is unable to directly communicate his private information to *H* in a way that gives *H* reason to act on the information. What is needed, of course, is some signaling mechanism by which *G* could indirectly, but reliably, communicate his private information to *H*. This mechanism can be created by adding a second move to the game in the form of a "subwager." The subwager might be set up as follows: *H* and *G* would agree that, although *G* still cannot show his "up card" to *H*, the remaining three cards would be shuffled and one turned over.¹²⁶ If that card turns up red, *H* must settle the case for a greatly

defendants face in plea bargaining, in which their private information of factual guilt or innocence is relevant to their assessment of what will happen at trial, but not a reliable indicator of trial outcomes.

124. $EP = (1/3) \times (20) = \6.66 . If *G*'s *EP* is now \$6.66, then *G* can pay 1/2 of \$6.66 (\$3.33) as a settlement price.

125. $EP = (\text{probability that } G \text{ loses}) \times (\text{amount } G \text{ will pay if he loses}) = (1/3) \times (20) = \6.66 .

126. The shuffling must be presumed not to change the actual outcome of the game. Perhaps, prior to shuffling, a third party would note the color of the down card, which would be revealed if the game was played out.

discounted price, say, a quarter of the initial settlement offer: \$1.25. If the card turns up black, however, *H* will greatly increase the settlement price to, say, \$10.

In this two-move game, *G*'s decision to accept the subwager will turn on his private information. That is, if *G*'s privately held card is black, *G* will have a strong incentive to accept the wager, since *G* knows that two-thirds of the remaining cards are red.¹²⁷ If *G*'s privately held card is red, however, *G* will have an equally strong incentive to decline the wager and settle for the initial offer price, since *G* knows that the odds are that he will lose if the game is played out. This can be verified by looking at how expected payoffs change once the private information and second move opportunities are introduced. If *G*'s up-card is black, he knows that he will win the subwager in two of three plays (because the remaining three cards include two reds and one black), and thus that his expected payoff if he settles after the subwager is \$4.16, an almost 20% improvement over the initial settlement offer.¹²⁸

Gamblers with positive private information will do even better than this, however, because those that lose the subwager will still know that their chances of prevailing if the game is played out are better than initially forecast. With this knowledge, they know that their real expected payoff from playing out the game at that point is only \$6.33, rather than the \$10 initially forecasted in the absence of the private information. If that adjustment is made, then *G*'s real expected payoff, if he accepts the wager, is only \$3.63.¹²⁹

If *G*'s privately held card is red, on the other hand, he will do markedly worse by accepting the wager, since now he knows there is a two-thirds chance that his bottom card will be a losing black card. In that case, his expected payoff if he accepts the subwager is \$7.09.¹³⁰ Because in this latter case the initial settlement offer of \$5 is better than the expected payoff from taking the subwager, *G*s with unfavorable private information rationally should decline the subwager and settle at the initial offer price.

Understanding this, *H* can confidently make differential settlement offers to *G*s depending on their responses to the subwager. The subwager mechanism

127. And thus that the bottom card—which determines whether *G* wins or loses—is likely to be a winning red card.

128. $ESUB = P(wB)(HO) + 1 - P(wB)(AO) = (.667)(1.25) + (.333)(10) = 4.16$, where $ESUB$ = expected price if subwager is accepted, $P(wB)$ is probability of winning if a black card is the private up-card, HO is the House's offer in the event that *G* wins the subwager, and AO is the alternative offer in the event *G* loses the subwager.

129. That is, either he settles for the \$1.25 if the card comes up, as expected, red, or he plays out if the card unexpectedly comes up black, in which case he has a two-thirds chance of winning (and thus his expected payout is only \$6.33).

130. $ESUB = 1 - P(wB)(HO) + P(wB)(AO) = (.333)(1.25) + (.667)(10) = 7.09$.

reliably signals the private information of *G*s and thus improves the accuracy of the game. Without the subwager, every gambler should accept the initial settlement offer. But only half of those *G*s would actually have "lost" had they played out the game.¹³¹ In the single-move game, therefore, would-be losers and winners alike would settle on the same terms. In the two-move game, however, private information has an impact on settlement prices.

The two-move game produces more accurate results in the sense of lowering the expected payoff of gamblers who likely would have won had the game played out, and increasing the expected payoff of gamblers who likely would have lost. That result should not be surprising. After all, by modifying the rules to enable *G* to view one of the cards, we have increased the amount of information in the system. But under the single-move game the same information could have been made available and yet, as we saw, have no effect on strategies. Setting up a second move allows that additional information to be factored into the payoff structure of the game.

This simple game is intended both to illustrate the point that it is possible to change a game's rules to make previously irrelevant private information relevant to the pricing mechanism and to provide a crude parallel model of the pricing dynamics at work in plea bargaining.¹³² In the model, the House represents the prosecutor, gamblers represent defendants, and the dollar figures represent the respective punishment imposed on defendants to resolve their cases. Defendants can pay the Prosecutor a "fee," that is, accept a negotiated punishment, or they can "play out," that is, go to trial. Defendants in criminal cases often have private information about their "type," that is, whether or not

131. The game is error free because it is always possible to verify with 100% certainty whether the bottom card is red or black.

132. A more complicated game that more realistically captures the dynamic of plea bargaining might factor in mutual incentives to avoid "playing out" hands. In such a game dealing would not be cost-free. Instead, if *G* chooses to play out the game, *H* must incur a \$15 "dealing" fee. Now, if *G* plays out the game and loses, *H* nets \$5. If *G* wins, however, then *H* loses \$15. *G*'s expected payoff from the game, therefore, is \$5. Rational *H*s will be no more interested in playing out the game than will rational *G*s. Everyone will be better off if *G* simply pays *H* the \$5 fee rather than playing out the game. In that situation, *G*'s payout is \$5, and *H*'s is \$5.

Now imagine that the parties could bargain over the fee. *G* is better off with a negotiated settlement rather than playing the game out as long as the fee does not exceed \$10. *H* is better off negotiating than playing as long as the fee exceeds \$5. There is a wide range of negotiated fees that would make both parties better off. Assuming the two sides split the difference, the fee would be set at \$2.50.

I do not develop this more complex model here because it is unnecessary for purposes of illustrating the signaling function of subwagers, which are equally effective in communicating private information in both the simple and the complex versions of the game.

they are factually guilty. Defendants, however, do not have any private information about whether the legal system will find them guilty, although their private information enhances their ability to predict outcomes. To replicate that dynamic, the game allows players to view one of their cards, providing information that enhances the player's knowledge of the probabilities, but still not know with certainty what will happen if the game plays out. The game, moreover, satisfies the basic criteria for reliable signaling:¹³³ The signaling is costly because making a bad subwager leads to predictably worse outcomes than making a good subwager; the costs of the signaling are differentially distributed; the signal is logically related to the message (whether the player is likely to win or lose the game); and accepting the subwager is economical for the relevant subgroup.¹³⁴

There are, it should be acknowledged, several problems with using the card-color game to analyze plea bargaining as it occurs in practice.¹³⁵ One of the biggest is the problematic quality of private information regarding guilt or innocence. Most scholars have assumed that private information of guilt or innocence enhances the information-holder's ability to assess the probability of conviction.¹³⁶ This is predicated primarily on the assumption that the private-information-holder's knowledge can help him predict whether as-yet-undiscovered or undisclosed evidence will be inculpatory or exculpatory. In

133. See *supra* notes 109–13 and accompanying text (describing the basic criteria necessary for reliable signaling).

134. Although I think it is conceptually useful to think of the subwager as a signaling device, technically speaking, some might describe it instead as a "screening" device. See generally Joseph Stiglitz, *Sorting Out the Differences Between Signaling and Screening Models* (Nat'l Bureau of Econ. Research, Technical Working Paper No. 93, 1990) (defining games in which the informed move first as signaling games and games in which the uninformed move first as screening games). The distinction is not relevant to the analysis presented here.

135. The game's assumptions, of course, diverge from plea bargaining in real life in a variety of ways. The game assumes that parties have full information about the probabilities and make rational decisions based upon them, whereas in real life the parties rarely, if ever, have full information, may frequently disagree about the probabilities, and often act irrationally. Similar criticisms are frequently made of any attempt at economic modeling of human decisionmaking. The typical rejoinder is that analytical techniques employed by economists and game theorists based on assumptions such as rational choice do not presuppose that those assumptions are always, or even usually, consistent with empirical reality. Rather, they attempt to uncover logical relationships useful for constructing predictive theories and testable hypotheses.

136. See Bjerck, *supra* note 18, at 306–07 (assuming defendants can better predict the probability that they will be convicted in arguing that guilty defendants will accept longer plea bargains than innocent defendants); Scott & Stuntz, *supra* note 12, at 1936–37 (explaining that a defendant, especially an innocent defendant, can better assess the probability of conviction than the prosecutor because the defendant knows more about what new evidence may or may not be found later).

the card-color game, the private information held by the player (regarding attributes of his hand) increases the player's ability to predict the outcome of the game.

The assumption that private knowledge of guilt and innocence is helpful in predicting trial outcomes, however, will not always hold true. In some cases, the parties can be reasonably confident that no yet-to-be-discovered evidence exists and that the universe of admissible evidence is a closed set.¹³⁷ In such cases, the private information has no predictive value. In other cases, private information may lead to incorrect predictions about future evidence. Although it may sometimes be true that truthful witnesses are more credible, some empirical data suggests that the opposite may also sometimes be true.¹³⁸ Because witnesses who lie often carefully rehearse their stories, their accounts are delivered in an (artificially) coherent and confident fashion.¹³⁹ Fact-finders often misperceive such coherence as an indication of sincerity.¹⁴⁰ In contrast, characteristics common to truthful testimony, including imperfect recall of details, inconsistency, and a less-confident delivery, can be misinterpreted by fact-finders as signposts of falsity.¹⁴¹ In short, the degree to which factual innocence improves the odds of acquittal, and thus the degree to which private information regarding factual guilt will improve a criminal defendant's ability to estimate trial outcomes, varies from case to case. It may be significant in some cases and trivial or nonexistent in others.

Second, the payoff structure makes a critical difference in determining whether private information is transmitted. As the card-color game demonstrates, at least under some conditions a two-move game can provide

137. Imagine, for example, a simple assault or robbery case, in which the prosecution evidence consisted solely of the victim's eyewitness identification of a suspect, who asserts as an alibi that he was home, alone, sleeping at the time the crime occurred.

138. See Jeremy A. Blumenthal, *A Wipe of the Hands, A Lick of the Lips: The Validity of Demeanor Evidence in Assessing Witness Credibility*, 72 NEB. L. REV. 1157, 1194 (1993) (reviewing social science experimental data indicating that speaker cues that are "popularly believed to manifest deception are qualitatively and quantitatively different from those which are actually observed during deception").

139. See Margaret Talbot, *Duped: Can Brain Scans Uncover Lies?*, NEW YORKER 52, 52 (July 2, 2007) ("A liar's testimony is often more persuasive than a truth-teller's. . . . People who are afraid of being disbelieved, even when they are telling the truth, may well look more nervous than people who are lying."). See generally Charles F. Bond & Bella M. Depaulo, *Accuracy of Deception Judgments*, 10 PERSONALITY & SOC. PSYCHOL. REV. 214 (2006) (analyzing extensive research results regarding an individual's ability to discern whether another person is lying or telling the truth).

140. Talbot, *supra* note 139, at 52.

141. *Id.*

greater price accuracy than a single-move game. Players with positive private information can utilize it to obtain better outcomes than players who either lack the information or have negative private information. But minor changes in the payoff structure can destroy these beneficent effects.

If we assume that players have access to private information whether or not they have an opportunity to make the subwager, the private information itself will perform an anti-pooling function, but only if the initial plea discount does not exceed a certain magnitude.¹⁴² A gambler who learns that the color of the top card is black also knows that the odds that the bottom card is red have risen from one-half to two-thirds. That information permits the gambler to revise his estimate of the cost of playing out from \$10 to \$6.66. However, as in the example above, where the settlement discount is 50%, the expected cost of playing out after incorporation of the private information still exceeds the initial settlement offer of \$5. Under those conditions, players will settle rather than play regardless of the nature of the private information. If the settlement discount narrows, however, the private information can tip the decision point towards rejection of the initial settlement offer. A 25% settlement discount would be large enough to entice gamblers with no private information (or negative private information) to settle, but not big enough to entice gamblers with positive private information to settle. They would be better off refusing the deal and playing out.¹⁴³

In sum, while an effective signaling mechanism can be designed in theory by using the subwager concept, that result occurs only when certain conditions are satisfied. Still, subwagers can improve the accuracy of plea prices compared with trial shadow pricing models alone by incorporating heretofore unaccounted private information into the pricing mechanism. But to be effective, the subwager must be based on information that has a meaningful relationship to the actual guilt or innocence of defendants, and must be embedded within a set of bargaining rules calibrated to make the subwager outcome relevant to optimal plea-bargaining strategies. In the next Part, the Article will discuss the use of subwagers in practice. As it turns out, at least

142. Bjerk makes this point in positing that a defendant's private knowledge of guilt should make him "willing to accept a longer plea bargain sentence than an otherwise similar appearing innocent defendant." Bjerk, *supra* note 18, at 307.

143. The ability to "capture" private information by enforcing a narrower plea/trial sentence differential provides a powerful justification for enforcing strong limitations on the differential. For a discussion of how such limits might be practically and effectively enforced, see Russell D. Covey, *Fixed Justice: Reforming Plea-Bargaining with Plea-Based Ceilings*, 82 TUL. L. REV. 1237 (2008).

one feature of current criminal procedure—interrogation—already functions, coincidentally or not, much like the subwager device described above.

IV. Interrogation and Sub-Wagers

In this Part, I argue that the procedural mechanism best suited to perform the separating function necessary to achieve the goal of more accurate plea-pricing is interrogation. Building on the insights of Seidmann and Stein's path-breaking game-theoretic analysis of the privilege against self-incrimination,¹⁴⁴ the Article argues that the decision to cooperate or not cooperate in the pretrial investigation satisfies (in a partial way) the essential prerequisites of a signaling mechanism. Submitting to interrogation is costly to criminal suspects, and that cost is imposed differentially. Interrogation is not cost-free for innocent suspects, but innocent suspects' costs are less—sometimes far less—in submitting to interrogation than those of guilty suspects. Moreover, the signal produced in interrogation—cooperation in the interrogation room—is logically related to the subject of the signal: The suspect's guilt or innocence. And interrogation is, or at least can be, an economically rational move for innocent suspects.

A. Cooperation as an Anti-Pooling Device

In a fascinating analysis of the privilege against self-incrimination, Seidmann and Stein apply game theory to interrogation.¹⁴⁵ In so doing, the authors attempt to refute the classical assumption that only the guilty benefit from the right to silence.¹⁴⁶ They argue "that the right to silence helps to distinguish the guilty from the innocent by inducing an anti-pooling effect that enhances the credibility of innocent suspects."¹⁴⁷ Guilty suspects, they argue, face a difficult strategic position, which they liken to "zugzwang," a concept in chess signifying a situation in which all available moves worsen a player's

144. See generally Seidmann & Stein, *supra* note 72. For critical commentaries on the Seidmann and Stein thesis, see Stephanos Bibas, *The Right to Remain Silent Helps Only the Guilty*, 88 IOWA L. REV. 421 (2003); Gordon Van Kessel, *Quieting the Guilty and Acquitting the Innocent: A Close Look at a New Twist on the Right to Silence*, 35 IND. L. REV. 925 (2002).

145. See generally Seidmann & Stein, *supra* note 72.

146. See generally *id.*

147. *Id.* at 433.

strategic position.¹⁴⁸ As Seidmann and Stein explain, a guilty suspect in the interrogation room is in zugzwang.¹⁴⁹ Telling the truth is not a good strategy for the guilty suspect because doing so entails confessing to a potentially serious crime.¹⁵⁰ Lying is also treacherous. If the suspect is caught lying, his guilt will be both obvious and easier for the prosecutor to prove.¹⁵¹ Silence, however, is certain to signal guilt and may improve the prosecutor's ability to prove it if the suspect's silence can be brought to the jury's attention.¹⁵²

The choice confronting suspects in the interrogation room can be depicted in the following two-by-two matrix:

| | | |
|------------|---------------|--------------------|
| | Exoneration | Self-Incrimination |
| Disclosure | 0(<i>b</i>) | -20(<i>nb</i>) |
| Silence | <i>X</i> | -10 |

For innocent defendants, the payoff associated with truthful disclosure of the facts as they know them is high: If their account of innocence is believed, the police will exclude them from further investigation, and they will be exonerated.¹⁵³ If their account inculcates them in the crime, however, or is verifiably false, their decision to disclose in the interrogation room will be a bad one, and they will self-incriminate. Silence, on the other hand, promises no chance of exoneration. The suspect who refuses to answer his interrogator's questions will look guilty in the interrogator's eyes, and his silence will probably lead to further investigation.¹⁵⁴ This outcome is plainly worse than exoneration, but much better than self-incrimination.

Seidmann and Stein argue that the key contribution of the right to silence is to create a meaningful opportunity for strategic decisionmaking, which in turn is a precondition for the existence of a signaling opportunity.¹⁵⁵ When

148. *Id.* at 441–50.

149. *Id.*

150. *Id.* at 443–44.

151. *Id.* at 445–46.

152. *See id.* at 444 ("Silence or false responses in the face of criminal accusations usually signal guilt.").

153. In the matrix, *b* = believed, and *nb* = not believed. An *X* denotes that a particular outcome is impossible, given a particular choice. The numerical value represents the utility or disutility resulting from the choice.

154. *See* Seidmann & Stein, *supra* note 72, at 444 ("[I]f [a suspect] were to opt for silence . . . the police and prosecution would infer his factual guilt and would therefore devote more time and effort to securing his conviction.").

155. *See id.* at 467–68 (comparing interrogation strategies of guilty suspects when there is

exposed lies and silence are equally harmful to the suspect's position, the best strategy for a rational suspect is to lie; whereas silence is both damning and certain, lying is only incriminating if the lies are exposed.¹⁵⁶ It follows, then, that rational guilty defendants will lie as long as silence is incriminating. A guilty suspect's lies, however, impose costs, or externalities, on innocent suspects.¹⁵⁷ Because innocent suspects' nonverifiable but true innocence claims are indistinguishable from the nonverifiable and false innocence claims of guilty suspects, both sets of innocence claims will pool.¹⁵⁸ Fact-finders will not be able to distinguish the true innocence claims from the false ones. As a result, all innocence claims will be discounted, which makes innocent defendants as a group worse off.

Robust rules safeguarding the right to silence, however, change the incentive structure of guilty suspects. Because silence does not carry the same incriminating consequences as unsuccessful lying, guilty suspects have an incentive not to lie but to invoke the right to silence instead.¹⁵⁹ To make silence a truly rational option, however, the right must be made as robust as possible, a need reflected in cases like *Griffin v. California*¹⁶⁰ and *Doyle v. Ohio*¹⁶¹ that extend the right to silence as far as, and perhaps beyond, its practical if not its logical limits.¹⁶² The goal is to make silence the best strategy for guilty defendants but not innocent ones. Where that happens, the marginal value of an unverifiable innocence claim increases to the benefit of innocent suspects who can cheaply make such claims.¹⁶³

and is not a right to silence).

156. *Id.* at 468.

157. *Id.* at 469.

158. *Id.* at 468–69.

159. *Id.* at 469.

160. *See Griffin v. California*, 380 U.S. 609, 613 (1965) (barring comment on defendant's silence at trial).

161. *See Doyle v. Ohio*, 426 U.S. 610, 617–18 (1976) (prohibiting adverse inferences based on defendant's silence at custodial interrogation).

162. Observers have frequently pointed out the implausibility of jurors actually ignoring a defendant's trial silence. Some argue that *Griffin* extends the privilege well beyond its justifiable purposes. *See, e.g.*, Charles Hobson, *The Minimalist Privilege*, 1 N.Y.U. J.L. & LIBERTY 712, 713 n.16 (2005) (arguing that cases like *Griffin* demonstrate the "privilege has strayed from its moorings"). These critics may be correct as a matter of history, but they are missing the functionality of a robust privilege.

163. *See Seidmann & Stein, supra note 72*, at 469 (explaining that juries will only believe innocent suspects' claims of innocence under a system in which guilty suspects would have no incentive to lie).

Seidmann and Stein's analysis provides a neat explanation for the privilege, and one that may be, as they contend, more intrinsically persuasive than the deontological defenses that have been advanced by the privilege's defenders.¹⁶⁴ The plausibility of Seidmann and Stein's account, however, depends largely on the plausibility of their assumption that false statements made by lying suspects "pool" with true statements made by innocent suspects in a way that "might impair the credibility of statements given by innocent suspects."¹⁶⁵ This claim appears problematic in one important respect. If Seidmann and Stein mean that a nonverifiable innocence claim has diminished credibility *at trial* as a result of the pooling effect, that claim seems conspicuously implausible.¹⁶⁶ Juries do not consider the credibility of innocence claims in the aggregate, but usually only the credibility of the innocence claim proffered by one defendant on trial.¹⁶⁷ Jury trial is generally not a comparative business, and jurors do not go about considering or assessing the "market value" of nonverifiable exculpatory statements. Jurors also have limited ability to compare innocence claims intertemporally because they are, by and large, one-shot actors who lack the opportunity to assess credibility with sufficient frequency to reflect any conceivable externality effect.¹⁶⁸

But if the externality argument is implausible as to jury trials, it seems quite plausible with respect to police and prosecutors, who are repeat players and who undoubtedly do assess a suspect's credibility in light of their

164. *See id.* at 503–04 (contending that the anti-pooling rationale for privilege "outscores its competitors by both justifying and coherently explaining each branch of Fifth Amendment jurisprudence").

165. *Id.* at 457–58.

166. Seidmann and Stein seem to imply that juries rationally respond to the market value of exonerating statements at several points in their article. *See, e.g., id.* at 469 ("By remaining silent, the guilty suspect therefore separates herself from the suspicious (and all other) innocent suspects. As a result, the jury draws a favorable inference from any exculpatory statement, and innocent suspects (who alone make such statements) are thus acquitted, regardless [of other evidence in the case].").

167. Occasionally, juries might have to make a comparative judgment about the credibility of competing innocence claims. For example, they might be faced with such a situation if multiple defendants are charged with the same crime and tried together, and each defendant points the finger at his codefendants. *See, e.g., Zafiro v. United States*, 506 U.S. 534, 535–36 (1993) (discussing trial of four defendants for possession of narcotics found in a car and a box in an apartment when each defendant disclaimed knowledge of the drugs and argued that the drugs belonged to one of the other defendants).

168. *See* Van Kessel, *supra* note 144, at 953 ("[T]oday's jurors often sit for only one or two trials and have little or no market savvy or experience. Today's jurors rarely return to the market place of exonerating statements.").

comparative experience with other criminal suspects.¹⁶⁹ It would not be at all surprising if police and prosecutors' ability to evaluate the truthfulness of a suspect's innocence claim marginally decreased where suspects were forced to manufacture exculpatory lies in response to questioning.¹⁷⁰ Some guilty defendants who otherwise would have remained silent would incriminate themselves absent a right to silence, but others would lie successfully, or at least in a manner that left the question of their credibility in doubt. And these additional plausible but nonverifiable innocence claims would pool with those made by actually innocent defendants, as Seidmann and Stein's analysis predicts.¹⁷¹

The shift in the pooling argument's focus from jurors to police and prosecutors has interesting consequences for plea bargaining. Since the "externality" problem is not likely to change the way a jury would assess the credibility of any particular defendant's innocence claim, the existence of a general right to silence would not seem to have any effect on individual defendants' probability of conviction. Does it then follow that the right to silence has no impact on plea prices? The trial shadow model would seem to suggest that it does not because plea prices only reflect expected trial outcomes.¹⁷² Indeed, in a world in which the only thing that matters is expected trial outcomes, submitting to interrogation looks like an act of ignorance or insanity. After all, the defendant who submits to interrogation at best does not harm his case; any self-serving or exculpatory statements will be inadmissible at trial.¹⁷³

169. See *id.* (pointing out that police and prosecutors are more likely to be aware of a changing market in exonerating statements by defendants than trial jurors).

170. See TOM BUCKE ET AL., HOME OFFICE RESEARCH STUDY 199, THE RIGHT OF SILENCE: THE IMPACT OF THE CRIMINAL JUSTICE AND PUBLIC ORDER ACT 1994, at 34–36 (2000) (U.K.), available at <http://rds.homeoffice.gov.uk/rds/pdfs/hors199.pdf> (showing no increase in confessions in Great Britain despite an increase in the number of suspects answering questions following changes in the law permitting evidentiary use of silence).

171. See Seidmann & Stein, *supra* note 72, at 503 (arguing that, absent a right to silence, all guilty suspects would lie during interrogation, and thus, fact-finders would be unable to distinguish between true and false testimonies).

172. See Scott & Stuntz, *supra* note 12, at 1937 (stating that both the prosecutor and the defendant must generate an estimate of the likelihood that the defendant will be convicted if the case goes to trial, and "that estimate will determine the price that each will insist on as a condition of reaching a bargain").

173. In a few rare situations the refusal to cooperate with police can be used at trial as evidence of guilt. Pretrial silence thus sometimes increases the likelihood of conviction, and thus reduces the plea price. See *Fletcher v. Weir*, 455 U.S. 603, 607 (1982) (finding that adverse inferences may be drawn from a defendant's post-arrest silence prior to receipt of *Miranda* warnings); *Jenkins v. Anderson*, 447 U.S. 231, 238 (1980) (finding that adverse

Unless one is prepared to characterize the vast majority of suspects as insane or ignorant, something else must be leading them to submit to interrogation. That something, I think plainly, is signaling. Suspects in the interrogation room are faced with a strategic problem that functions, essentially, like a subwager. They can accept the subwager and cooperate with the interrogation, or they can decline the subwager by invoking their right to silence.

The rewards and risks are distributed similarly to those in the card-color game. Suspects who successfully exculpate themselves effectively bargain their way out of further prosecution by convincing the police or prosecutor that the costs of prosecution exceed any possible benefits. Suspects who talk and incriminate themselves, in contrast, greatly cheapen the price of their pleas,¹⁷⁴ in some cases (where they confess) by virtually eliminating outcome uncertainty and thereby reducing the value of the plea to the value of the resource savings of foregoing trial.¹⁷⁵ Suspects who do not talk at all marginally decrease the value of their pleas, but not because the silence can be used against them at trial. Rather, their plea values decrease because their silence signals that they have something to hide.¹⁷⁶

inferences may be drawn from a defendant's pre-arrest silence). The point of the robust law protecting the privilege, however, is to minimize silence's evidentiary value.

174. Cassell and Hayman report empirical data showing that "defendants whom police successfully questioned were less likely to receive concessions in plea bargaining." Paul G. Cassell & Bret S. Hayman, *Police Interrogation in the 1990s: An Empirical Study of the Effects of Miranda*, 43 UCLA L. REV. 839, 911 (1996). The data indicated that defendants who cooperated with interrogators and succeeded were three times more likely to win charge bargaining concessions in the district studied than defendants who made incriminating statements. *See id.* (finding that 30.6% of confessing defendants pled guilty without a reduction in offense level, while only 9.4% of suspects who were interrogated but did not incriminate themselves did the same).

175. This obvious consequence is borne out by empirical evidence showing, not surprisingly, that suspects who confess during interrogation plead guilty in very large numbers. *See* PAUL SOFTLEY, HOME OFFICE RESEARCH STUDY 61, POLICE INTERROGATION: AN OBSERVATIONAL STUDY IN FOUR POLICE STATIONS, at 45 (1980) (U.K.), *available at* <http://www.homeoffice.gov.uk/rds/pdfs05/hors61.pdf> (reporting that "[o]nly two among 57 suspects who had made a confession pleaded not guilty, whereas 12 of 34 suspects who had not made a confession pleaded not guilty").

176. Silence also decreases the value of the silent suspect's plea by increasing the investigatory burden on the state. Information that it might have obtained cheaply from the suspect now can only be obtained by investing additional investigatory resources in the case. *See id.* at 41 (describing police questioning as providing "a short cut to the successful conclusion of an investigation" based on interview research findings that "[i]n over 30 per cent [sic] of cases investigating officers said that, if the suspect had refused to answer their questions, they would have tried to obtain additional evidence"). As with all investments, the state will expect a return, and will be willing to pay less for the plea as a result. This conclusion is

Thus, apart from any information obtained in the course of the interrogation, the very act of choosing to submit to interrogation or to remain silent reveals valuable private information. By treating the signal as an important piece of information that is relevant to plea price, the interrogation subwager should, in theory, have a partially separating effect on a pool of defendants that trial shadow theory would fail to distinguish—namely, suspects who cooperated with interrogators and made no incriminating statements and suspects who refused to cooperate and, therefore, also made no incriminating statements. As a result, prosecutors should be expected to treat noncooperating suspects differently than cooperating suspects, even where the objective post-interrogation probabilities of conviction are identical.

A review of empirical data generated in several frequently cited studies supports this analysis. First, there is some reason to believe that the decision to cooperate with interrogators correlates with guilt and innocence. What little empirical research there is exploring the reasons why suspects waive their right to silence and talk to police questioners suggests that the decision is often based on a belief that the police will presume them guilty if they refuse to cooperate.¹⁷⁷ The presumption that non-cooperation signals guilt is made by the innocent even more than the guilty, as was evidenced in one study that found that innocent persons will waive their *Miranda* rights and talk to the police at significantly higher rates than guilty persons.¹⁷⁸ This data supports the view that the decision to cooperate itself is indicative, although certainly not dispositive, of a suspect's personal assessment of guilt or innocence.

supported by the empirical data. Cassell and Hayman's data, for instance, indicates that defendants who refuse to cooperate in interrogation are less likely to obtain a plea concession than suspects who submit to questioning and don't incriminate themselves. Cassell & Hayman, *supra* note 174, at 911. This research indicates that 15.4% of silent suspects pleaded "to charges at the same level as initially filed," while only 9.4% of unsuccessfully questioned suspects pled to the same charges. *Id.*

177. Viewed in light of its signaling implications, the consistently high rate at which suspects voluntarily submit to questioning notwithstanding the right to silence is more easily understandable. See, e.g., Timothy O'Neil, *Why Miranda Does Not Prevent Confessions: Some Lessons from Albert Camus, Arthur Miller and Oprah Winfrey*, 51 SYRACUSE L. REV. 863, 865 (2001) (attempting to answer the question "why after all these years do suspects persist in waiving *Miranda* and confessing to the police?").

178. See Saul M. Kassin & Rebecca J. Norwick, *Why People Waive Their Miranda Rights: The Power of Innocence*, 28 L. & HUM. BEHAV. 211, 215 (2004) (finding in a mock interrogation following an incident in which half of the subjects committed a mock crime and half did not, thirty-six percent of the "guilty" subjects waived their *Miranda* rights, while eighty-one percent of the "innocent" subjects waived rights and talked to police).

Second, data gathered in two leading studies of interrogation supports the hypothesis that a suspect's decision to cooperate has tangible consequences for plea bargaining. Richard Leo conducted what is probably the leading American investigation of police interrogation and its consequences.¹⁷⁹ In the study, Leo observed 122 police interrogations in person and an additional 60 videotaped interrogations.¹⁸⁰ He then traced the impact of interrogation on case outcomes.¹⁸¹ In particular, he attempted to determine the comparative significance of invoking or waiving *Miranda* rights on case outcomes, and if there was a waiver and subsequent interrogation, of the consequences of a successful interrogation (i.e., incriminating statements were obtained) versus an unsuccessful interrogation.¹⁸² A second, widely cited, study was conducted by Paul Cassell and Bret Hayman.¹⁸³ In the Cassell and Hayman study, the researchers participated in screenings of more than 200 felony charge filings.¹⁸⁴ Based on these screenings, the researchers tracked charges filed against 219 suspects.¹⁸⁵ Although the authors reach different substantive policy conclusions, data from both studies bear out the intuition that significant consequences turn on the decision to invoke or waive the right to silence. First, Leo found that suspects who invoked their right to silence or to a lawyer were charged at a slightly higher rate (73%) than suspects who agreed to answer police questions (69%).¹⁸⁶ This supports the conclusion that refusing to answer police questions increases police suspicion.¹⁸⁷

Even more significant are the outcomes of suspects who agreed to answer police questions but who did not make any incriminating statements during interrogation (who I will refer to as "cooperating suspects") compared with

179. See generally Richard A. Leo, *Inside the Interrogation Room*, 86 J. CRIM. L. & CRIMINOLOGY 266 (1996).

180. *Id.* at 268.

181. See *id.* at 269 ("I analyze the effects of police interrogation practices, *Miranda* warnings, and incriminating statements on the subsequent stages of the criminal process, such as the adjudication of guilt, case disposition, and sentencing.").

182. *Id.* at 287–89.

183. See generally Cassell & Hayman, *supra* note 174.

184. *Id.* at 842.

185. *Id.* at 852.

186. *Id.* at 287. The four percent disparity in outcomes, however, was not statistically significant. *Id.* Nonetheless, the data does support a statistically significant finding that refusing to cooperate with interrogators does not improve the likelihood of not being charged. *Id.* at 288.

187. A recent study of the British experience with the right to silence came to the same conclusion. See DAVID WOLCHOVER & ANTHONY HEATON-ARMSTRONG, WOLCHOVER AND HEATON-ARMSTRONG ON CONFESSION EVIDENCE 632–33 (1996) (finding that suspects who remain silent are more likely, in borderline cases, to be charged).

those of suspects who refused to submit to interrogation or who terminated an ongoing interrogation ("noncooperating suspects"). Here, the disparity appears more dramatic. Although Leo's study did not directly contrast the outcomes of these two groups, a review of his published data shows that noncooperating suspects were convicted in 53% of the cases,¹⁸⁸ whereas cooperating suspects were convicted only 43% of the time.¹⁸⁹

The impact was even more significant on charging decisions. In Leo's study, whereas 73% of noncooperating suspects were charged,¹⁹⁰ only 56% of cooperating suspects were charged.¹⁹¹ Cooperating successfully with interrogators thus appears to improve the likelihood of avoiding charges by more than 33%.

Cassell and Hayman's data shows a comparable disparity, finding that whereas noncooperating suspects were ultimately convicted in 70.6% of cases, cooperating suspects were convicted in only 49.3% of cases.¹⁹² Cassell and Hayman also found that prosecutors filed charges in 81.0% of cases in which suspects invoked their rights, compared with 74.0% when suspects successfully cooperated.¹⁹³ Even more striking is the disparity revealed by Cassell and Hayman's data in dismissal rates. Whereas the authors found that noncooperating suspects had charges dismissed or entered pleas in abeyance¹⁹⁴ only 7.7% of the time—the lowest rate of all categories, including confessing suspects—30.2% of cooperating suspects won dismissals or entered pleas in abeyance.¹⁹⁵

188. Leo, *supra* note 179, at 288 tbl.11.

189. *Id.* at 299 tbl.16.

190. *Id.* at 288 tbl.10.

191. *Id.* at 299 tbl.16. Cooperating defendants also had their cases dismissed at a high rate (fifty-three percent). *Id.* I could not determine, however, from Leo's published data what the dismissal rate was in those cases in which suspects invoked their *Miranda* rights. Similar data was reported in Paul Softley's study of police interrogation in the U.K. See SOFTLEY, *supra* note 175, at 45 (reporting that "[a]mong those who had neither confessed nor made an admission, one half were released and one half were prosecuted").

192. Cassell & Hayman, *supra* note 174, at 915 tbl.17.

193. *Id.* at 908 tbl.15. It is not clear whether the difference is statistically significant. See *id.* at 908 n.317 ("Comparing those successfully questioned to all others produced a definite trend that just missed significance at the .05 level, but was significant at the .10 level (*pr* [the likelihood of such a correlation occurring randomly] = .054). Comparing those successfully questioned to those unsuccessfully questioned was chi-square significant at the .05 level."). In comparison, prosecutors filed charges in 87.5% of cases in which they obtained incriminating statements. *Id.* at 908 tbl.15.

194. See *id.* at 909–10 ("Under such a plea, a defendant enters a plea of guilty, but charges are dropped if the defendant does not commit any new offenses within a set period of time.").

195. *Id.* at 912 tbl.16. The authors state, however, that these disparities are not statistically

In short, the data from both studies shows that accepting the interrogation subwager has a significant impact on the outcome of the case. Suspects who successfully cooperate with interrogators are more likely to see charges dismissed, more likely to receive a charge concession in plea bargaining, and less likely to be convicted if their case does in fact go to trial. Although the difference in outcomes between cooperating and noncooperating suspects might ultimately be caused by other factors, the data is consistent with the hypothesis that prosecutors treat cooperation itself as a potential signal of innocence and take that information into account in charging and plea bargaining. The tangible benefits of successful cooperation, moreover, help explain why suspects choose to submit to interrogation at such high rates notwithstanding *Miranda*. The data also show that the costs incurred by guilty suspects who gamble on the subwager are high. Suspects who take the subwager and lose (that is, confessing suspects) incur the greatest costs. Losing subwagerers are far less likely to have charges dismissed or to receive a charge concession in plea bargaining, and are most likely to be convicted if their cases go to trial. Finally, suspects who decline the gamble fall somewhere in the middle, being less likely than successfully cooperating suspects to receive a charge concession in plea bargaining, but more likely to receive such a concession than suspects who took the interrogation subwager and lost.

The subwager hypothesis predicts not only that suspects who refuse to cooperate in interrogation should be treated less leniently by prosecutors, but also that they should be more likely to take the moderate discount contained in the initial plea offer than defendants who take the interrogation subwager. As to this prediction, the data is not consistent. Leo's data appears to contradict that prediction, showing that cooperating suspects resolved their cases through guilty pleas twice as often as noncooperating suspects.¹⁹⁶ Cassell and Hayman's data, on the other hand, is consistent with the hypothesis. Their data showed that noncooperating suspects resolved their cases with guilty pleas at

significant. *Id.* at 913.

196. See Leo, *supra* note 179, at 289 tbl.12 (finding that defendants who waived their rights resolved cases through plea bargaining 48.51% of the time, as opposed to the 24.32% plea resolution rate among those who invoked their *Miranda* rights). Of course, there are other factors that affect these rates. Most important is Leo's finding that individuals with criminal histories invoked their *Miranda* rights more frequently than those without criminal histories. See *id.* at 286 ("[A] suspect with a felony record in my sample was almost four times as likely to invoke his *Miranda* rights as a suspect with no prior record and almost three times as likely to invoke as a suspect with a misdemeanor record."). Criminal history might well impact prosecutorial plea-bargaining strategies. Unfortunately, Leo's published data does not permit comparison of the plea offers made to waiving and invoking defendants.

the highest rate (92.3%), and went to trial the least (0.0%).¹⁹⁷ Cooperating suspects, in comparison, resolved their cases through plea bargaining at a 65.7% rate, and those who did received larger plea discounts.¹⁹⁸ Nearly 40% of suspects who were questioned unsuccessfully received plea discounts of two to three charge levels, whereas only 7.7% of noncooperating suspects received the larger discount.¹⁹⁹ Finally, the vast majority (69.2%) of noncooperating suspects in the Cassell and Hayman study settled their cases for what appears to be a modest discount—one charge level—whereas slightly less than half of the cooperating suspects settled for the modest discount (48.4%).²⁰⁰ Cassell and Hayman's data shows exactly the pattern expected if interrogation operates as a subwager: Suspects are presented with the choice to accept the subwager. Those that do and win get the best outcomes. Those that do and lose get the worst, and those who decline to accept the subwager settle their cases for a modest, intermediate-sized discount. The small sample size in the Cassell and Hayman study, and the inconsistencies between the two studies as to plea rates of noncooperating suspects, however, mean that further empirical research in this area is necessary before any strong claims can be made about whether, and under what circumstances, the interrogation subwager translates into predictable plea discounts.

B. Interrogation as a Subwager

For any procedural device to provide the signaling benefits of the subwager model discussed above, it must satisfy the preconditions of signaling mechanisms generally. Recall again the criteria. Effective signals must be costly, costs must be differentially distributed, the distribution must be related to the quality signaled, and the decision to signal must be economical.²⁰¹ How does interrogation fare on these criteria?

The answer, I think, is mixed. First, submitting to interrogation is costly. Every suspect who agrees to answer questions risks making an admission that might be used against him in court. Even innocent suspects incur costs by submitting to interrogation, and those costs might be substantial in some

197. Cassell & Hayman, *supra* note 174, at 912 tbl.16.

198. *Id.*

199. *Id.*

200. *Id.*

201. *See supra* notes 109–13 and accompanying text (describing the basic criteria necessary for reliable signaling).

cases.²⁰² Second, there is clearly some cost differential between innocent and guilty suspects. Although innocent suspects incur costs by submitting to interrogation, guilty suspects incur higher costs because it is much easier for an innocent suspect to tell a consistent credible story than it is for guilty suspects. The costs, however, are not perfectly distributed. Because answering questions is ultimately voluntary, guilty suspects have some ability to monitor the progress of the interrogation and prevent revelation of incriminating information. That ability, however, is limited. Once questioning begins, the suspect relinquishes substantial control over the shape and direction of the questioning.²⁰³ Interrogators, moreover, can probably gauge with decent accuracy when suspects are not being forthright in their responses. So even when incriminating information is withheld during the interrogation, it will remain much harder for guilty suspects to falsely convince interrogators that they have nothing to hide.

Still, there is no doubt that some guilty suspects do trick interrogators into believing their false exculpatory stories. Guilty suspects talk in the majority of cases and innocent suspects sometimes do not.²⁰⁴ Estimates regarding interrogation success rates vary from as low as 32% to as high as 76%.²⁰⁵ Leo found that incriminating statements were obtained in 76% of cases in which defendants submitted to interrogation,²⁰⁶ Cassell and Hayman report a success

202. For example, a suspect in a murder case undoubtedly strengthens the state's case by truthfully explaining to authorities that he killed the victim in self-defense because he admits presence at the crime scene and the actus reus of the offense.

203. For that reason, interrogation will work best as the basis for an evidentiary subwager where the defendant contends that he or she is factually innocent and commits to that position early in the investigative process.

204. Leo's data found that only 21.7% of suspects invoked their rights after receiving *Miranda* warnings. Leo, *supra* note 179, at 276 tbl.3.

205. Leo surveys a variety of studies, most of which estimated successful interrogations occurring at a lower rate. *See id.* at 280–81 (citing Lawrence S. Leiken, *Police Interrogation in Colorado: The Implementation of Miranda*, 47 *DENV. L.J.* 1, 13 (1970) (reporting 32% success rate); David W. Neubauer, *Confessions in Prairie City: Some Causes and Effects*, 65 *J. CRIM. L. & CRIMINOLOGY* 103, 104–06 (1974) (reporting 46% success rate); Michael L. Wald et al., *Interrogations in New Haven: The Impact of Miranda*, 76 *YALE L.J.* 1519, 1566 (1967) (reporting 51% success rate); James W. Witt, *Non-Coercive Interrogation and the Administration of Criminal Justice: The Impact of Miranda on Police Effectuality*, 64 *J. CRIM. L. & CRIMINOLOGY* 320, 325 (1973) (reporting 67% success rate); Evelle Younger, *Results of a Survey Conducted in the District Attorney's Office of Los Angeles County Regarding the Effect of the Miranda Decision Upon the Prosecution of Felony Cases*, 5 *AM. CRIM. L.Q.* 32, 35 (1966) (reporting 50% success rate)).

206. Leo, *supra* note 179, at 280.

rate of only 42.2%.²⁰⁷ There is, of course, no way to tell what percentage of interrogations produced either false positives or false negatives.²⁰⁸ The possibility that a guilty suspect will successfully withstand interrogation, however, is probably sufficiently robust that many guilty suspects willingly take the risk, while the hazards of falsely confessing or otherwise making incriminatory admissions are sufficiently high that at least some innocent defendants decline to submit to interrogation notwithstanding the potential benefits of doing so. This, of course, waters down the signal's value. The extent to which this watering down occurs probably varies from case to case, just as the apparent cost of submitting to interrogation varies depending on the circumstances surrounding the interrogation. Particularities relating to the nature of the crime, the suspect, and the evidence available to the police prior to interrogation will all greatly impact the suspect's chances of maintaining a false but credible story. Although these factors may well be known to police and prosecutors, the fact that they will vary from case to case makes generalization difficult about the signaling value of a suspect's decision to submit to interrogation. Still, although interrogation is far from a perfect science, it does tend to produce relevant and, not infrequently, dispositive evidence, and thus the decision to submit to it at least some of the time will have signaling value. In addition, although interrogators do not obtain incriminating statements in every interrogation, empirical success rates are less important than perceived success rates. Subwagers based on interrogation will still signal the suspect's private knowledge of guilt or innocence as long as the suspect believes interrogation will reveal the truth.²⁰⁹

207. See Cassell & Hayman, *supra* note 174, at 871 (concluding, additionally, that the post-*Miranda* success rate was lower than the pre-*Miranda* success rate based on earlier studies of the effectiveness of interrogation).

208. The DNA exoneration cases document the fact that some innocent defendants falsely confess to crimes they did not commit. There also are undoubtedly many cases in which innocent suspects made incriminating statements that, though not full confessions, were used to help convict the defendant at trial.

209. For example, assume that lie-detector tests are no better than tarot cards at revealing whether an individual is telling the truth, and that their results are inadmissible in court. A subwager conditioning a more lenient price for successful questioning under the lie-detector would still effectively separate guilty and innocent defendants as long as the defendants themselves believed the lie-detector results were accurate and admissible in court because only subjectively innocent suspects would take the subwager. This tactic is, in fact, sometimes used. See *Smith v. Baldwin*, 510 F.3d 1127, 1130 (9th Cir. 2007) (describing a situation where the results of a lie detector test were inconclusive, but significance was still assigned to the test taker's responses). The court described the situation as follows:

While burglarizing the home of Emmett and Elma Konzelman, either Smith or his criminal companion, Jacob Edmonds, bludgeoned eighty-seven-year-old Mr.

Third, the signal is not arbitrary. Although the costs of sending a credible innocence signal may not always be evenly differentiated, the decision to cooperate with interrogators is directly and intuitively related to the suspect's guilt or innocence.

Finally, interrogation appears to satisfy the criterion that signaling be economical. That is, signaling costs for innocent suspects seem sufficiently low, and the payoff for successful signaling sufficiently high, that signaling is a rational choice. Not only do some suspects find it economical to attempt to signal innocence by submitting to interrogation, most do.²¹⁰ Indeed, so many suspects submit to interrogation that a puzzle arises. If interrogation acts as a subwager in the manner discussed and suspects are properly informed about the consequences of their choices, most innocents should talk and most guilty suspects should remain silent. But data show that the vast majority of suspects, including guilty suspects, submit to interrogation.²¹¹ Why do so many guilty suspects apparently make such bad bets? There are a variety of possible answers. One possibility is that suspects are in fact foolish or ignorant. Guilty suspects talk because they are unaware that there is a better strategy. Silence may not provide enough benefit to guilty suspects, or it may not be apparent to them that it does.

Alternatively, silence may be too beneficial to suspects. The strong legal protections for silence extended by *Miranda* may have caused police to feel (correctly or not) that they cannot make cases without interrogation evidence. The benefits of attempting to signal innocence by submitting to interrogation are outweighed by the potential costs of doing so. Where that is true and known to suspects, however, there is no marginal benefit in signaling even for innocent suspects. In such cases, signaling is uneconomical, and rational suspects will not signal. The perverse result of this dilemma, however, might be that police feel compelled to apply strong psychological pressure to get suspects to talk. Given that pressure, guilty suspects may talk not because they

Konzelman to death with a three-foot long crowbar. After Edmonds told police that Smith killed Konzelman, the prosecution offered Edmonds a plea deal contingent on his passing a polygraph examination. The results of the polygraph test were inconclusive, but the examiner opined that Edmonds had answered the questions truthfully and Edmonds entered the plea deal in exchange for his testimony against Smith.

Id.

210. See Cassell & Hayman, *supra* note 174, at 860 tbl.3 (finding that only 16.3% of suspects in the study invoked their *Miranda* rights and refused questioning); Leo, *supra* note 179, at 276 tbl.3 (noting that 78.29% of suspects waived their *Miranda* rights).

211. See *supra* note 210 and accompanying text (providing examples of such data).

believe it is in their rational interests, but because they believe that have no real choice. Everyone talks, and again no signaling results. Although the extent of pressure placed on suspects undoubtedly varies, it is probably safe to say that interrogation functions as a signaling mechanism some of the time, but that it is dependent on a confluence of factors to create the proper conditions. This is perfectly consistent with signaling theory. As Spence concluded about signaling in the labor market, whether a signaling mechanism produces a separating or pooling equilibrium depends upon the specific parameters that define the signaling model.²¹²

C. Implications

If the analysis here is correct, plea bargain theory must be modified to account for the substantial likelihood that plea bargain prices are influenced by signaling. Pretrial interrogation creates an obvious dilemma for guilty suspects, and the defendant's response to interrogation generates a signal permitting some plea price differentiation to occur among guilty and innocent defendants, a prediction supported by the empirical data. Even apart from the substantive information obtained from interrogation, the choice to submit to interrogation serves as an important signal to police and prosecutors that is manifested in plea prices. In short, there is good reason to believe that plea bargains are negotiated not only in the shadows of trial, but also in the shadows of strategic choices made by suspects prior to trial and, especially, in the interrogation room. Moreover, because interrogation-room subwagers take place in a far less legally regulated environment than trial, police and prosecutors are in a better position than jurors to acquire signaling information. They can then put that information to use for purposes of plea bargaining, making bargaining outcomes marginally more accurate than conventional trial shadow theory would predict.

Indeed, plea-bargaining outcomes in a regime where interrogation effectively functions as a signaling device may well be superior to those that would be generated were plea bargaining abolished. Because signaling occurs between prosecutors and suspects, and because that signaling information is hard to communicate to juries, the impact on plea prices can easily be overlooked by plea-pricing models dependent on trial shadows. Unless prosecutors had some way to communicate to juries the signals they receive,

212. See Spence, *supra* note 85, at 368 (noting that multiple equilibria are a "distinct possibility" of a signaling model).

which the rules barring comment on a defendant's pretrial silence prevent, abolition of plea-bargaining would have the perverse result of decreasing the total amount of information available to assess guilt and innocence.²¹³

The signaling analysis also has a variety of practical implications. First, if interrogation already functions informally as a subwager with signaling effects on plea prices, the stakes involved in the debate over acceptable interrogation practices are even higher (if they can be any higher) than supposed. Even interrogations that fail to produce useful or admissible evidence have an impact on case outcomes if police and prosecutors' charging decisions are affected not only by what suspects say in the interrogation room, but also by their willingness to submit to interrogation in the first instance.

Equally important, accurate plea-bargaining outcomes depend on ensuring that interrogation is effective, on one hand, but that it does not produce false-positives, on the other.²¹⁴ Innocent defendants must believe that cooperation with investigators is a rational course. Confidence can be enhanced by increasing suspects' perceptions that police interrogation will be conducted fairly for purposes of discovering truth. Innocent suspects must be confident that truthful answers will not be misconstrued or taken out of context and used to falsely convict them. At the same time, interrogation cannot be rendered so toothless that guilty suspects are encouraged to hazard it. Rather, from a signaling perspective guilty suspects need to be encouraged to opt for silence and a modest plea deal rather than attempt to falsely exculpate themselves, a strategy that is costly both to them and to actually innocent defendants. Walking this tightrope is obviously difficult, but there are some steps—such as videotaping interrogations—that would seem well suited to addressing both goals. Additionally, a renewed commitment to conforming interrogations not

213. Indeed, comparing the model to data regarding suspect cooperation rates suggests that contrary to Seidmann and Stein's conclusions regarding the anti-pooling effects of a strong right to silence, signaling effectiveness might be increased by weakening that right. Rather than maximize the value of silence, legal rules might better induce an anti-pooling effect by decreasing pressure on police to override a suspect's preference for silence. If jurors could be informed of the defendant's silence, then police might feel less compelled to elicit confessions. Innocent suspects would then have an incentive to talk, while guilty suspects would still fear incrimination. In any case, a very delicate balance is required to preserve the signaling mechanism. Silence can be neither so costly as to make it equivalent to confession (in which case everyone will talk), nor so cheap that everyone opts for it, making it overly difficult for police to convict the guilty.

214. This is the problem of managing signaling costs so they satisfy Spence's cost criteria. See Spence, *supra* note 85, at 367 ("Therefore, effective signaling depends not only upon the negative correlation of costs and productivities, but also upon there being a 'sufficient' number of signals within the appropriate cost range.").

only with the letter, but the spirit, of *Miranda* might also enhance the signaling power of interrogation by maximizing the voluntariness of the interrogation choice. Seen in this light, even staunchly pro-law-enforcement advocates might appreciate that *Miranda*'s contribution to plea bargaining is more nuanced than its critics suppose.

Second, the need to obtain reliable signals from criminal suspects helps explain some highly criticized features of criminal procedure. For example, as discussed above, many writers have called for discovery reforms to strengthen defendants' access to pre-plea discovery.²¹⁵ But the net benefits of enhanced discovery must be evaluated not only with respect to improving bargaining efficiency, but also in light of the need to maximize the accuracy of the system's signaling mechanisms. Current discovery rules limit the ability of guilty defendants to make fully informed plea bargain decisions in a strategic sense, but that effect is a direct consequence of the goal of maximizing the informational value of guilty pleas. Indeed, with the notable exception of *North Carolina v. Alford*,²¹⁶ the Supreme Court's plea bargaining cases are overwhelmingly predicated on the notion that guilty pleas are confessions, not strategic choices.²¹⁷ Seen in this light, providing criminal defendants more

215. See *supra* notes 52–62 and accompanying text (summarizing critics' arguments for greater defendant access to pre-plea discovery and explaining the costs and benefits of such discovery for defendants).

216. See *North Carolina v. Alford*, 400 U.S. 25, 33 (1970) (permitting defendant charged with capital murder to plead guilty while maintaining factual innocence and acknowledging that "reasons other than the fact that he is guilty may induce a defendant to so plead" (quoting *State v. Kaufman*, 2 N.W. 275, 276 (Iowa 1879))).

217. See, e.g., *Bordenkircher v. Hayes*, 434 U.S. 357, 363 (1978) (stating that defendants in plea bargaining "are presumptively capable of intelligent choice in response to prosecutorial persuasion, and unlikely to be driven to false self-condemnation"); *Brady v. United States*, 397 U.S. 742, 758 (1970) (finding confessions that result from plea bargaining generally voluntary, partially based on the view that plea bargaining does not increase the likelihood that defendants "would falsely condemn themselves," and based on "expectations . . . that pleas of guilty are voluntarily and intelligently made by competent defendants with adequate advice of counsel"). As such, guilty pleas routinely have been held to waive a defendant's claim to procedural rights. See *United States v. Broce*, 488 U.S. 563, 573 (1989) (finding a guilty plea voluntary despite counsel's failure to point out a potential defense); *Tollett v. Henderson*, 411 U.S. 258, 267 (1973) (finding a guilty plea voluntary despite counsel's failure to find a potential constitutional infirmity in grand jury proceedings); *McMann v. Richardson*, 397 U.S. 759, 770 (1970) (finding no grounds to withdraw a guilty plea where counsel "misjudged the admissibility" of a "confession"); *Brady*, 397 U.S. at 757 (finding a guilty plea voluntary even though defendant "misapprehended the quality of the State's case" and the "likely penalties"). In light of the defendant's supposedly truthful and accurate confession, courts have ruled that procedural rights designed to protect the accuracy of the truth-finding process are rendered effectively obsolete, and therefore subject to forfeiture. In *United States v. Ruiz*, for example, the Court concluded that Ruiz was not entitled to disclosure of exculpatory impeachment evidence under

information during plea-bargaining would tend to dilute the reliability of their admissions of guilt.²¹⁸ Although the Court's assumption that guilty pleas are reliable and non-strategic is empirically questionable, the aspiration it reflects is consistent with the signaling analysis presented here.²¹⁹ If a defendant is forced to make an independent assessment of trial outcomes in the absence of reliable information about the evidence that will actually be introduced at trial, she inevitably must rely on private information to make it. As such, her untutored plea reveals relevant and otherwise unobservable information about her guilt or innocence.²²⁰

Weak judicial regulation of police deception in interrogation serves a similar function. As with all signaling mechanisms, interrogation will most effectively provide a separating signal for innocent and guilty suspects where the cost differential of sending the innocence signal, that is, accepting the interrogation subwager, is greatest.²²¹ The risk that guilty suspects can effectively game the interrogation subwager is minimized by the suspect's

Brady because such evidence is relevant to the "fairness of trial," not the voluntariness of a guilty plea. See *United States v. Ruiz*, 536 U.S. 622, 629 (2002) (concluding that the Constitution does not require pre-guilty plea disclosure of impeachment information).

218. As one commentator observed, if the litigation process is thought of as a "search for truth," then a guilty plea conviction "is not impaired by any weakness in the trial evidence which the prosecutor could have adduced had the issue of guilt been contested." H. Richard Uviller, *Pleading Guilty: A Critique of Four Models*, 41 *LAW & CONTEMP. PROBS.* 102, 114 (1977).

219. There is ample anecdotal evidence that defendants routinely plead guilty to offenses they did not commit. We know for a fact that innocent people plead guilty. Although we do not know how often, given that the incentives inherent in the plea bargaining system are structured in a way that make guilty pleas the rational strategy irrespective of the defendant's guilt or innocence, its frequency may be substantial. As Daniel Givelber notes, the problem of convicted innocents is "invisible" for a variety of reasons, including that the legal system is not structured to adjudicate actual innocence (only if the defendant is guilty or not-guilty), and there are no formal pronouncements made or judicial opinions written after evidence of innocence does surface. Daniel Givelber, *Meaningless Acquittals, Meaningful Convictions: Do We Reliably Acquit the Innocent?*, 49 *RUTGERS L. REV.* 1317, 1323–25 (1997). Even more fundamentally, it is conceptually difficult, if not impossible, to determine who has been falsely convicted because doing so would require access to a guilt-innocence sorting mechanism superior to the mechanisms officially used—trial and guilty pleas—and no such mechanisms exist.

220. The point is not to increase the number of guilty pleas or to mimic probable trial outcomes, but to force defendants to rely on a different data set—the defendant's privately held information—which is not accessible to investigators (and might never become accessible to a jury)—rather than the state's evidence in estimating the likelihood of conviction. The ability to withhold pretrial discovery strategically provides a further opportunity to structure a more effective subwager. Contingent discovery regimes are increasingly common.

221. See *infra* notes 222–24 and accompanying text (discussing use of deceptive interrogation tactics and their effects on guilty and innocent suspects).

uncertainty as to the direction questioning might take. The law and practice of interrogation have evolved in ways that help police maximize uncertainty. For instance, police have developed numerous interrogation tactics based on deception, tactics (including lying about and concealing evidence) which the Supreme Court has consistently approved,²²² and which researchers have found are frequently used.²²³ Although quite arguably unfair, such tactics undoubtedly make it difficult for guilty suspects to gauge whether a false exculpatory account is consistent with facts already known to police, and therefore increase the difficulty of maintaining a coherent and plausible but untruthful account of innocence.²²⁴ Deceptive interrogation tactics thus discourage guilty suspects from hazarding interrogation and make silence a more appealing option. And precisely because silence is more appealing to guilty suspects, innocent suspects who can cheaply cooperate in the interrogation room benefit. The increase in signaling costs for guilty suspects, but not innocent ones, directly facilitates their self-separation.

Third, the analysis suggests ways that additional private information might be productively harvested from criminal defendants. Although in most cases interrogation will precede charging and therefore any information obtained from it will already have been incorporated into the plea price, there may be some cases in which circumstances permit a more conscious exploitation of the subwager device. In some cases, charging precedes arrest, such as in cases arising from grand jury investigations. In others, charging occurs absent interrogation because police did not bother to interrogate, because the defendant initially invoked his *Miranda* rights, because the defendant was represented at the time of arrest or shortly thereafter and the attorney prevented interrogation, or because the police otherwise lacked information or opportunity to conduct an early interrogation. In those cases, the defendant's cooperation in interrogation might explicitly be made a part of bargaining over plea prices. Even if an initial interrogation was conducted, interrogation might be useful as a subwager device where new evidence justifies additional, post-charge questioning.

222. See, e.g., *Illinois v. Perkins*, 496 U.S. 292, 300 (1990) (affirming a conviction based on a confession induced from a prisoner by an undercover police officer posing as a fellow inmate planning to escape).

223. See Leo, *supra* note 179, at 278 (characterizing "confront[ing] suspect[s] with false evidence of guilt" as a tactic that was used often during the course of the interrogations observed by the researcher).

224. See Seidmann & Stein, *supra* note 72, at 443 (explaining that a guilty defendant's effort to bluff innocence may backfire if police discover the bluff, thereby furnishing "further evidence of his guilt").

Finally, subwager devices might be constructed around defendants' investigational cooperation more broadly. Here I mean something different than the kind of investigational cooperation that prosecutors frequently negotiate for to help them investigate and prosecute crime. Defendants, of course, already receive substantial plea discounts for cooperating with police and prosecutors. But the usual plea discount follows provision of useful inculpatory information—that is, defendants can get a plea break by admitting guilt and implicating others.²²⁵ Like discounts for "acceptance of responsibility," the irony of typical cooperation discounts is that they are only available to guilty defendants. Because, by definition, innocent defendants are not in a position to provide useful inculpatory information to law enforcement, they may as a group receive worse plea deals than their guilty brethren. Defendants who agree to make open-ended commitments to affirmatively cooperate with prosecutors and law enforcement agents by submitting to interrogation and by providing other assistance in investigating the facts of their case, and against whom prosecutors do not obtain additional incriminating evidence as a result of that cooperation, should be rewarded with a different type of cooperation discount, one that is comparable to that awarded to defendants for acceptance of responsibility. Such investigational cooperation discounts would signal defendants' subjective belief in their own innocence and would, in effect, replicate the subwager dynamic, further helping to improve the pricing accuracy of the plea bargaining process. By providing incentives to innocent defendants to truthfully proclaim innocence, they would also minimize the perverse incentives of wrongfully convicted defendants to falsely confess guilt to limit their sentencing exposure or enhance their chance of parole.²²⁶

225. See, e.g., U.S. SENTENCING GUIDELINES MANUAL § 5K1.1 (2007) (authorizing sentence reduction based on government submission of a letter acknowledging provision by defendant of "substantial assistance" in investigating or prosecuting criminal activity).

226. See Daniel S. Medwed, *The Innocent Prisoner's Dilemma: Consequences of Failing to Admit Guilt at Parole Hearings*, 93 IOWA L. REV. 491, 497 (2008) (noting that proclaiming innocence at parole hearings hinders the possibility of parole, placing wrongfully convicted prisoners in a "true 'prisoner's dilemma'"); *id.* at 535 (noting that defendant's "acceptance of responsibility" is required at plea hearings by federal rules governing guilty pleas and is rewarded with sentencing concessions under the Sentencing Guidelines).

V. Conclusion

Plea bargaining is part of a complex system of criminal justice with dynamic and interdependent parts; its outcomes are a product of a wide array of factors, including draconian sentencing schemes and enormous prosecutorial charge discretion that may create fundamental bargaining disparities and, as a result, unfair bargains. An additional source of unfairness, however, is implicit in the basic economic model underpinning plea bargaining: The risk that bargaining, which occurs in the shadow of trials, misprices the guilty pleas of innocent defendants. Plea bargaining's pricing model misprices pleas in part because it fails to incorporate a critically important piece of information—the defendant's subjective knowledge of guilt or innocence—into the pricing mechanism.

Effective signaling mechanisms, however, promise the possibility of mitigating that defect. At least in theory, subwagers might be used to increase the accuracy of negotiation even though one party never receives private information directly and would not have credited the information had it come straight from the party who did have it. In practice, interrogation appears to function as a subwager device, signaling guilt or innocence to police and prosecutors in ways that likely increase the accuracy of resulting plea bargains. Although it will never provide a perfect separating signal, the analysis of subwagers suggests both that they do increase the amount of available information, and that there are opportunities to further exploit the signaling possibilities of subwagers in plea bargaining.