Beyond Plea Bargaining: A Theory of Criminal Settlement

Richard Lorren Jolly
J.J. Prescott
University of Michigan Law School, jprescott@umich.edu

Available at: https://repository.law.umich.edu/articles/2883

Follow this and additional works at: https://repository.law.umich.edu/articles

Part of the Criminal Law Commons, Criminal Procedure Commons, and the Dispute Resolution and Arbitration Commons

Recommended Citation

This Article is brought to you for free and open access by the Faculty Scholarship at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in Articles by an authorized administrator of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.
BEYOND PLEA BARGAINING: A THEORY OF CRIMINAL SETTLEMENT

RICHARD LORREN JOLLY
J.J. PRESCOTT

INTRODUCTION .............................................................................................................. 1049
I. THE INTERESTS AND CONSTRAINTS OF CRIMINAL SETTLEMENT ......................... 1059
   A. The Parties’ Interests in Criminal Settlement ......................................................... 1060
   B. The Judge as a Constraint on Settlement .............................................................. 1067
II. CRIMINAL SETTLEMENT AS PARTIAL SETTLEMENT ............................................ 1072
   A. Procedure-Modification Agreements .................................................................... 1073
   B. Issue-Modification Agreements ............................................................................ 1082
   C. Outcome-Modification Agreements ..................................................................... 1087
III. MODELING JUDICIAL REVIEW OF CRIMINAL SETTLEMENT .............................. 1095
   A. Constraining Judicial Discretion .......................................................................... 1097
   B. Judicial Shaping of Criminal Settlement ............................................................ 1104
CONCLUSION ............................................................................................................. 1114
Abstract: Settlement is a term rarely used in criminal law. Instead, people speak almost exclusively of plea bargaining—i.e., enforceable agreements in which a defendant promises to plead guilty in exchange for a prosecutor’s promise to seek leniency in charging or at sentencing. But a traditional plea agreement is just the most visible instance of a much broader class of possible criminal settlement agreements. In terms of their fundamentals, criminal settlements are indistinguishable from their civil counterparts: through either an atomized or comprehensive bargain, parties exchange what they have for what they want, advancing their respective interests in cost minimization, risk mitigation, and value maximization. Focusing only on a defendant’s promise to plead guilty discounts the diversity and complexity of the agreements into which defendants and prosecutors may and regularly do enter. This Article advances a comprehensive framework of criminal settlement—one that leverages incomplete or partial settlements as an analytical frame that stretches beyond mere plea bargaining. As in the civil context, criminal settlements need not resolve disputes outright but may instead limit or redefine a dispute in a way that the parties find mutually beneficial. A crucial difference in the criminal context, however, is that these bargains necessarily take place in the shadow of judicial discretion that regulates access to the state’s power to punish. Consequently, prosecutors and defendants may agree to reshape procedures, issues, and potential outcomes in order to constrain or influence judicial decision-making in ways that are congruent with their respective interests. But judges are not passive in this landscape. They, too, may act strategically to prompt the bargains we ultimately observe. By modeling this interplay between parties and judges, this Article fashions a more complete picture of the motivations, consequences, and policy implications of criminal settlement.
INTRODUCTION

The fundamental dynamics of settlement activity in the criminal domain are not unlike those that characterize the settlement of civil actions. Both contexts see parties bargaining toward mutual promises because they each believe they will be better off with an agreement than without one.\(^1\) Overwhelmingly—at least ninety-seven percent of the time at the federal level—the criminally accused offer up some or all of their procedural rights because they place greater value on the consequences they expect will follow from a prosecutor’s particular promises, which are usually to seek leniency at sentencing or to forgo certain charges, than they do on the option to exercise those rights.\(^2\) Prosecutors, for their part, usually prefer the potential outcomes associated with a defendant’s waiver of these procedural rights—for instance, the certainty of a conviction or savings in time and resources—more than they do the possibility of some alternative punitive result, usually one emerging from the default of a full criminal trial.\(^3\) Although outcomes are not denominated in dollars and the risk preferences of civil and criminal defendants may differ in systematic ways, criminal settlement behavior, just like civil settlement behavior, is at root driven by the goals of cost minimization, risk mitigation, and value maximization.\(^4\)

Despite this core similarity in the interests that stimulate and shape settlement across legal domains, there is a substantial difference in the raw amount of autonomy that civil and criminal parties enjoy in structuring their bargains. Civil litigants can sculpt highly specific settlements from a near infinite menu of options.\(^5\) Outside of class actions and a few other scenarios involving conflicts of interest or substantial third-party effects, civil plaintiffs can voluntarily withdraw their complaint in exchange for anything (or

\(^1\) For an early discussion on how rational parties will negotiate agreements to their mutual benefit, see generally ANTHONY T. KRONMAN & RICHARD A. POSNER, THE ECONOMICS OF CONTRACT LAW (1979).


nothing)—or, more practically speaking, anything that can be the subject matter of a legal contract—and do so without a judge’s permission. This means, for example, that civil litigants can construct agreements in which the alleged harm is incongruous with the negotiated resolution. Nothing and no one can stop a competent private party from accepting a peppercorn to resolve a wrongful death suit. The bargaining field is extremely broad and involves just the two in dispute—the plaintiff and the defendant.

Parties in a criminal dispute do not enjoy such freedom, and thus criminal settlement negotiations occur in a quite different environment. A criminal conviction, which transforms a defendant into someone who is legally recognized as having committed a crime, requires court approval. Perhaps more importantly, the physical incarceration of a defendant—often the primary outcome of interest on both sides of the table—simply cannot be a valid term in a private contract, at least not one that the parties can lawfully perform and enforce without the involvement of a court and the satisfaction or waiver of certain procedural formalities before a duly sworn judge. More generally, prosecutors and defendants are unable to unilaterally implement any mutually preferred outcome, at least one considered “criminal,” even if that outcome would otherwise exactly correspond with an outcome that a court would enforce. A bona fide judge must review, approve, and impose the terms of any agreement. Contrast the typical criminal plea colloquy and subsequent sentencing before a judge in a face-to-face proceeding with the

---

6 See Sanford I. Weisburst, Judicial Review of Settlements and Consent Decrees: An Economic Analysis, 28 J. LEGAL STUD. 55, 55 (1999). In most circumstances, there is no judicial review of these private decisions.

7 See U.S. CONST. amend. V.

8 Imagine a scenario in which a citizen approaches the state and agrees to publicly announce criminal guilt of some sort and proceed to prison without ever involving a judicial actor—perhaps a Hunger Games-type situation in which someone agrees to go to prison as “tribute” for a friend or family member. Or imagine a civil agreement in which private parties agree that their dispute is best resolved by one of them spending some period of time in prison (even a private one). Such fanciful agreements would not be enforceable, and the parties could challenge them under a variety of federal and state constitutional provisions. In this sense, what makes a traditional plea agreement enforceable and distinct, at least as compared to other agreements between parties, is that it necessarily involves the review of someone performing a judicial role. And, as we will discuss below, although a party can waive certain procedures to which they are entitled, a party cannot waive all such procedures. The executive branch cannot criminally punish a defendant merely by agreement. See infra notes 83–88 and accompanying text.

9 See FED. R. CRIM. P. 11 (outlining the types of agreements and procedures by which courts implement criminal settlements). Judges can and do reject proposed agreements for many reasons, including that a proposal would not give the defendant their just desserts in light of their crimes, that it would not deter either the individual charged or the general public in the future, and that it would result in an unfavorable public reaction. See Abraham S. Goldstein, Converging Criminal Justice System: Guilty Pleas and the Public Interest, 49 SMU L. REV. 567, 572 n.15 (1996) (collecting and analyzing cases on guilty pleas).
typical private civil settlement in which the parties jettison the court entirely, with one party agreeing to pay the other a lump sum in exchange for the latter abandoning their claims.

In a similar vein, once criminal charges are filed, prosecutors and defendants must operate in a legislative and constitutional milieu that dictates specific punitive outcomes (or ranges of outcomes) for specific harmful acts, and, like Hotel California, such surroundings may prove difficult to leave.10 The parties cannot agree to dismiss the case without judicial approval.11 Although approval is not easily withheld, the power of judicial discretion looms.12 Moreover, the crime of conviction and the punishment that

---

11 See FED. R. CRIM. P. 11(c) (requiring judicial review of proposed settlements); id. R. 48(a) (requiring prosecutors to obtain “leave of court” before dismissing charges).
12 In Rinaldi v. United States, the Supreme Court noted that “[t]he words ‘leave of court’ were inserted in [Federal] Rule [of Criminal Procedure] 48(a) without explanation.” 434 U.S. 22, 29 n.15 (1977) (per curiam). Although those words “obviously vest some discretion in the court,” the Court recognized that “the circumstances in which that discretion may properly be exercised have not been delineated.” Id. The Court then explained that the requirement “is apparently meant to protect a defendant against prosecutorial harassment” but acknowledged courts have also used it “to deny a Government dismissal motion to which the defendant has consented if the motion is prompted by considerations clearly contrary to the public interest.” Id. Contra generally Thomas Ward Frampton, Why Do Rule 48(a) Dismissals Require “Leave of Court”? 73 STAN. L. REV. ONLINE 28, 32 (2020), https: //review.law.stanford.edu/wp-content/uploads/sites/3/2020/06/73-Stan.-L.-Rev.-Online-Frampton.pdf [https://perma.cc/7EVH-VLLM] (arguing that the animating purpose of Rule 48(a) “had almost nothing to do with the rights of the accused” and contending that the rule was rather “understood as vesting district judges with the power to limit unwarranted dismissals by corruptly motivated prosecutors”). The tension here derives from the separation of powers between the judiciary and executive. As one judge for the U.S. District Court for the District of South Carolina explained in 1860 when denying leave: “It is true that the court has no power to command the prosecuting officer to proceed in a criminal case if he is unwilling to do so. It is equally true that when the court permits an entry to be made in its minutes of the entry of a nolle prosequi it adopts and justifies that proceeding.” United States v. Corrie, 25 F. Cas. 658, 669 (C.C.D.S.C. 1860) (No. 14,869); see also United States v. Krakowitz, 52 F. Supp. 774, 784 (S.D. Ohio 1943) (denying dismissal and noting that “when the prosecuting officer . . . does not seek to enter a nolle prosequi solely upon his own responsibility, but by motion asks the sanction of the Court as his justification . . . such motion seeks the active exercise of the sound judicial discretion of the Court” because “more than a mere ministerial or administrative act is involved; a judicial determination is sought, and the Court may grant or withhold its sanction as the circumstances may require”).

In 2020, the scope of this judicial power took center stage following the Trump administration’s unopposed motion to dismiss charges brought against General Michael Flynn, where the United States charged him with, and he pleaded guilty to, willfully and knowingly making false statements to the Federal Bureau of Investigation under 18 U.S.C. § 1001. In re Flynn, 973 F.3d 74, 76 (D.C. Cir. 2020) (en banc) (per curiam). District Court Judge Emmet Sullivan, expressing concern that an improper political directive motivated the government’s decision to drop the charges against Flynn, set a briefing schedule and scheduled argument on the government’s motion. Id. at 77 (discussing the briefing schedule); id. at 98 (Rao, J., dissenting) (expressing con-
a judge must impose are often tightly linked, and the perceived severity of the two are almost invariably positively correlated under the law. The parties cannot settle a case with a formal finding indicating acute criminality and the imposition of a negligible punishment—a murder charge cannot be “resolved” with the defendant paying a thousand dollar fine. This constraint on bargaining remains even if such a monetary sanction would best match the particular prosecutor’s and defendant’s respective risk tolerances and expectations regarding the outcome of the dispute.

Certainly, some of these constraints are the direct consequence of legislative enactments that define crimes while simultaneously listing their potential punishments. Yet parties regularly waive, ignore, or define away their legislative rights and remedies in noncriminal contexts. And, even when the law strictly prohibits certain combinations of settlement terms (say, by imposing mandatory minimum or maximum sentences), one can imagine a world in which prosecutors and criminal defendants embrace bargains involving “fictions” in a way that dramatically increases bargaining freedom even within a court setting. True, parties may be unable to link a formal

cerns about an improper political motive). Flynn appealed to the U.S. Court of Appeals for the D.C. Circuit seeking to compel the district court to immediately grant the motion to dismiss. Id. at 77 (majority opinion). The panel issued a writ to so compel the district court, and Judge Sullivan filed a petition for hearing en banc. Id. The circuit court granted the petition and vacated the original panel’s ruling, remanding the case to the district court for consideration. Id. at 85. Unfortunately, in the end, little progress was made on resolving this quandary because on November 20, 2020, President Trump pardoned Flynn before the district court could rule on the issue. But in his memorandum opinion dismissing the case, Judge Sullivan stressed that “courts are tasked with making their own determinations on whether dismissal would be in the ‘public interest,’” and he concluded that if not for the pardon he would likely not have granted the government’s motion. See United States v. Flynn, C.A. No. 17-232, 2020 WL 7230702, at *9 (D.D.C. Dec. 8, 2020) (quoting Rinaldi, 434 U.S. at 29 n.15). Accordingly, it remains unclear whether judicial review in this context is derived from Rule 48(a), implicit in Article III authority, or both. Likewise, there is no clear consensus on the proper scope of the court’s discretion. See Rinaldi, 434 U.S. at 29 n.15 (acknowledging that “the circumstances in which [Rule 48(a)] discretion may properly be exercised have not yet been delineated by [the] Court”). For our purposes here, the point is simply that a prosecutor and a defendant cannot unilaterally agree to settle their dispute by dropping charges that have already been filed without judicial review—regardless of the source and extent of that judicial authority. The judge remains a central player in criminal settlement dynamics.


14 This claim is true in that a charged murder cannot be so resolved (with “murder” remaining the crime of conviction); it does not mean that the parties cannot modify the charge to something other than murder that involves the defendant paying a fine (reckless driving, for instance). For our discussion on such issue-modification agreements, see discussion infra Part II.B.

15 Prosecutors have the constitutional right to decline to prosecute altogether as part of an agreement, but they must operate in a more regulated space once a criminal charge of any sort becomes part of the arrangement. See FED. R. CRIM. P. 11.

16 This already occurs to some degree. See discussion infra Part II.B (discussing settlements involving fictitious acts and even fictitious crimes).
first-degree murder conviction with a sentence requiring three months of probation. But if judges simply rubber-stamped criminal settlements—and if we assume that some set of facts will always exist to support all potential punitive outcomes—civil and criminal settlement would be virtually identical. The parties in both contexts would fine-tune their bargaining to identify precise substantive landing points in line with their various interests and without concern over third-party intervention. Any rubber-stamped legal label the court applied to the substantive criminal outcome might be a legal fiction (i.e., the defendant would be “convicted” of a purely fictional crime), but this would not distinguish the two domains in a significant way in this alternative world. After all, a civil settlement also cannot replicate a true jury verdict delivered by a court of law (i.e., a purely private civil settlement does not carry a court’s imprimatur), for example, nor any derivative effects of an accurate court-acknowledged outcome.

In the end, the fundamental distinction between the arenas of criminal and civil settlement revolves around the sharp difference in authority of the judge in the criminal domain to set the stage for the litigation, manage any adjudication, and influence the final outcome, particularly at sentencing. Put simply, judges have more power in the criminal justice system through the legal necessity of their sign-off for particular punitive outcomes, such as incarceration. The Constitution and/or legislation grant judges this authority, and parties cannot entirely counter it. What is more, a judge cannot lawfully be cut into the parties’ bargain in any significant way (i.e., the parties cannot directly share the surplus of any theoretical three-way bargain with a judge), so the parties must anticipate and contract around unresolved judicial discretion. Although there is no policy consensus on the degree to which judges should exercise this discretion—some authors assert that too little oversight unfairly empowers prosecutors, while others warn that too much leads to biased sentencing variation—the fact that judges act as gatekeepers to the state’s power to punish is a defining characteristic of the American criminal justice system.

Because of this judicial custodian, a prosecutor and a defendant cannot unilaterally “settle” their dispute (as that term is traditionally used to mean a complete and private resolution). Instead, they can at best agree to limit or

---


19 See, e.g., Nancy J. King, *Judicial Oversight of Negotiated Sentences in a World of Bargained Punishment*, 58 STAN. L. REV. 293, 294 (2005) (describing judicial discretion as “an essential part of any sentencing policy that simultaneously values both efficiency through negotiated dispositions and consistent application of systemwide sentencing norms”).
reform their dispute as it goes before a judge in some mutually beneficial way. In this sense, all criminal settlement agreements are best understood as partial or incomplete settlements. As an analytical frame, the concept of a partial settlement discards the traditional binary notion of “pure trial,” in which parties rely entirely on background default rules, and “pure settlement,” in which they privately agree on the disposition of their dispute with no fact-finding procedure or decision-maker. Within this framework, agreements between parties fall along a wide spectrum of potential settlements reflecting “some agreement or set of agreements . . . that will reduce litigation costs, minimize trial risk, or improve each party’s expected outcome.” Thus, although a prosecutor and a defendant may not completely bypass judicial involvement in resolving their dispute, they nevertheless can construct and enter into partial-settlement agreements that alter it in ways that better accord with their respective preferences relative to the set of default rules that would otherwise govern.

The existing plea-bargaining literature overlooks these core settlement underpinnings and their implications for criminal law and procedure. Scholars have largely written themselves into a corner by addressing only a small subset of feasible bargains, specifically those that involve defendants agreeing to plead guilty to one or more criminal charges, generally in exchange for prosecutorial concessions with respect to charges or the consequences of convictions. Generally speaking, charge bargaining refers to a prosecutor agreeing to drop certain charges in return for a defendant’s waiver of certain procedural rights—principally, the right to a trial. Sentence bargaining involves a defendant agreeing to relinquish procedural rights in return for a prosecutor’s promise to seek a punishment discount at sentencing. Much of the plea bargaining scholarship builds on this typology either by adding complications, such as the interests of relevant third parties, or by stressing

20 J.J. Prescott and Kathryn E. Spier describe and develop in detail the idea of incomplete or partial settlements in the civil context. See Prescott & Spier, supra note 4, at 71.
21 Id. at 62.
22 Id.
25 Id. at 186.
that these types of bargains routinely fail to return expected or socially desirable results. Recent commentators have further supplemented the mix with the concept of fact bargaining—in which the parties stipulate to a version of facts that is consistent with the charges to be filed with the court—but even this advance has been presented with the defendant’s guilty plea as a fixed starting assumption.

In our view, by maintaining this plea-focused approach to bargaining in the criminal law, scholars have analyzed just one (admittedly important) side of the Rubik’s Cube. They have overlooked other forms of criminal settlement, both actual and theoretical. Consequently, they have missed the fact that the diverse terms of these multifarious arrangements can serve as substitutes and complements during bargaining. Parties can assemble these terms into settlement agreements that together mitigate the costs and risks of trial and improve ex ante expected outcomes, at times in ways that improve on traditional plea bargaining. More problematically, the traditional framework has led many scholars to disregard the critical role that judicial discretion—and the resulting uncertainty parties experience—plays in shaping criminal settlement agreements. The fixation on guilty pleas may have prevented the development of a more comprehensive theory of criminal settlement that accounts for the variety of agreements into which parties may, and regularly do, enter. True, some thinkers have recently begun to push a broader understanding of plea bargaining that stretches beyond its traditional conceptual boundaries. But none has yet offered a comprehensive tax-

27 See, e.g., Russel D. Covey, Signaling and Plea Bargaining’s Innocence Problem, 66 Wash. & Lee L. Rev. 73, 91 (2009).
28 Thea B. Johnson, for instance, describes what she calls “fictional plea[s]” as “a plea bargain agreement in which the defendant pleads guilty to a crime he did not commit, with the consent and knowledge of multiple actors in the criminal justice system.” Thea Johnson, Fictional Pleas, 94 Ind. L.J. 855, 857 (2019) (emphasis added). And, in condemning fact bargaining, Judge William G. Young portrays “it [as] involv[ing] a fraud on the court as the government’s recital of material facts during the plea colloquy and at sentencing necessarily must omit or at minimum gloss over facts material to sentencing.” Berthoff v. United States, 140 F. Supp. 2d 50, 62 (D. Mass. 2001). Although it is true that fact bargaining regularly occurs at sentencing “because certain material ‘facts,’ . . . now mathematically drive every sentencing decision,” as we will discuss below, there is no reason that the benefits of such bargaining cannot be realized by the parties without the defendant pleading guilty. See id.
29 See, e.g., Daniel S. McConkie, Judges as Framers of Plea Bargaining, 26 Stan. L. & Pol’y Rev. 61, 66 (2015) (“Judges ensure that the plea is ‘knowing’ and ‘voluntary’ but essentially rubberstamp most plea agreements . . . .”).
30 For a discussion of this burgeoning scholarship, see Talia Fisher, The Boundaries of Plea Bargaining: Negotiating the Standard of Proof, 97 J. Crim. L. & Criminology 943, 945 (2007) (exploring the border between alienable and inalienable procedural rights); Gregory M. Gilchrist, Trial Bargaining, 101 Iowa L. Rev. 609, 614 (2016) (arguing in support of allowing defendants and prosecutors to bargain over streamlined trial procedures); Russell M. Gold et al., Civilizing Criminal Settlements, 97 B.U. L. Rev. 1607, 1610 (2017) (using the term “settlement” to describe
onomy of the criminal settlement agreements that are likely to surface in the shadow of pervasive judicial discretion or a unifying theory that accounts for the motivations of parties in reaching such settlements. Criminal settlement as an idea and as an art is therefore ripe for academic reassessment, theoretical development, and practical innovation.

This Article looks beyond plea bargaining and offers a framework for conceptualizing all criminal settlements—or, agreements between prosecutors and defendants that alter the default rules of criminal adjudication. We focus our discussion mostly on bargaining in federal courts for the sake of exposition, but our arguments and descriptions hold for criminal settlement agreements generally. In essence, we claim that prosecutors and defendants “settle” by entering into agreements that reformulate their dispute to maximize their respective interests. These interests can be roughly characterized as cost minimization, risk mitigation, and ex ante value maximization. Parties realize these interests by crafting agreements that can be sorted into three non-exclusive categories: procedure-modification, issue-modification, and outcome-modification agreements. But because judges necessarily sit as

plea agreements and contending that criminal proceedings should incorporate certain procedures from the civil context to help make agreements more informed and voluntary); Nancy Jean King, *Priceless Process: Nonnegotiable Features of Criminal Litigation*, 47 UCLA L. REV. 113, 118 (1999) (noting that “[s]hort of foregoing trial altogether, defendants can make trials more economical or less risky for prosecutors . . . by waiving particular rules”); John Rappaport, *Unbundling Criminal Trial Rights*, 82 U. CHI. L. REV. 181, 182 (2015) (striving “to dispel the illusion that plea bargaining occurs within an all-or-nothing framework” and imagining the effects of making all procedural rights negotiable). Additionally, some scholars have referred to non-prosecution agreements (NPAs) and deferred prosecution agreements (DPAs) as forms of “criminal settlement.” See, e.g., Cindy R. Alexander & Mark A. Cohen, *The Evolution of Corporate Criminal Settlements: An Empirical Perspective on Non-prosecution, Deferred Prosecution, and Plea Agreements*, 52 AM. CRIM. L. REV. 537, 544 (2015). These types of agreements do indeed look much like what could be considered “settlement” in the traditional sense. They usually involve a corporate defendant providing an admission of fact, a commitment to cooperate, a specified duration for the arrangement, and an agreement to monetary and non-monetary sanctions. Id. As such, both NPAs and DPAs are examples of bundles of settlement terms, which complement one another in ways that advance the parties’ interests. Id. Like other criminal settlements, judges must review DPA filings to ensure that there is no undue exercise of discretion by prosecutors and that there is compliance with laws. Id. at 547–48. NPAs are distinct, however, in that they do not require the filing of criminal charges with the court so there is no judicial review. Id. at 545. NPAs are thus akin to civil litigants agreeing not to involve the court at all and to resolve their dispute privately. Although NPAs and DPAs are of great interest, this Article focuses instead on criminal settlements involving individuals to illustrate the broad applicability of our criminal settlement framework. Importantly, the model we advance here applies with equal force to criminal charges brought against corporate entities and can be used in that domain to illustrate the parties’ various interests in entering into such agreements.

31 This typology of criminal settlement agreements aligns with the basic arrangement seen in civil litigation, but the mechanisms by which the agreements operate are distinct in the two contexts. See Prescott & Spier, *supra* note 4, at 80 (advancing a theory of partial-settlement agreements in civil litigation).
final arbiters in criminal cases, parties must fashion arrangements that account for ample judicial discretion and uncertainty—such as by constraining judicial discretion to channel possible punitive outcomes in ways that conform with parties’ expectations and preferences.  

Our analysis has real-world payoffs. Concrete examples of criminal settlements that align with our framework abound—and extend far beyond traditional plea bargains. For example, defendants and prosecutors may agree to modify the procedure for determining guilt, trading a constitutionally guaranteed jury trial for a bench trial, with each party potentially believing that relying on a legally trained and experienced decision-maker will produce a better or more expeditious resolution. Or they may agree to modify the substantive issues at play in the dispute, regardless of the procedural scheme employed (such as the defendant stipulating to a specific quantity of drugs), which will in turn limit the issues that the factfinder needs to address as well as the range and types of outcomes available to the judge. Parties may also agree to modify the available punishment options by presenting a mutually agreed-upon sentence as a nonbinding recommendation to the judge rather than as a binding requirement of the plea. This approach has the effect of amplifying uncertainty because the judge may ignore the parties’ recommendation. But a mutually optimistic prosecutor and defendant may wager on this uncertainty, with each expecting the judge to deviate in their favor. Remarkably, parties may even bargain over terms such as the level of enthusiasm the prosecutor will use in making a sentencing proposal to the judge. Indeed, the scope of the settlement terms in any

---

32 Milton Heumann was perhaps the first to describe plea bargains essentially as agreements over how the parties will present the case to the judge. See Milton Heumann, Plea Bargaining: The Experiences of Prosecutors, Judges, and Defense Attorneys 43 (1978) (“The agreement that emerges from the sentence bargain can take several forms, all of which relate to what will, or will not, be said to the judge on sentencing day.”).  

33 See Bibas, supra note 5, at 2500.  

34 See Harris v. United States, 536 U.S. 545, 571 (2002) (Breyer, J., concurring in part and concurring in judgment) (“The upshot is that in many such cases defendant and prosecutor will enter into a stipulation before trial as to drug amounts to be used at sentencing (if the jury finds the defendant guilty.”), overruled on other grounds by Alleyne v. United States, 570 U.S. 99 (2013).  


37 Id. at 1324–26.  

38 See, e.g., United States v. Benchimol, 471 U.S. 453, 455 (1985) (per curiam) (upholding an agreement in which a prosecutor did not make his recommendation “enthusiastically,” at least in part because he had not promised to do so); United States v. Riggs, 347 F.3d 17, 19 (1st Cir.
case is potentially—and often in reality—vast. And, just as in the civil settlement context, criminal settlement agreements may combine procedure-, issue-, and outcome-modification terms as substitutes and complements to allow parties to tailor their dispute even more profitably.

Our partial settlement framework also allows us to advance a better understanding of the push-pull relationship between the judge, on one side, and the prosecutor and the defendant, on the other. By partially settling in some form, the prosecutor and the defendant manipulate the fundamentals of their dispute—including the range and forms of punishment available to the judge—allowing parties to speculate over potential outcomes in line with their respective risk tolerances. The indeterminacy introduced by judicial discretion can operate through party information asymmetries, overoptimism, loss aversion, and so on to encourage such bargains, activating risk-seeking or risk-averse behavior on behalf of both prosecutors and defendants. But judges are not passive players in this game; rather, they have the implicit power (and can take steps to enhance that power) to shape the types of agreements into which the parties will enter. Judges can play their role by strategically signaling what they are willing to accept and/or the potential price of their going along, both by injecting themselves into the litigation and by purposely building a reputation through their decisions in other cases. Even more so than in civil litigation, anticipated judicial behavior is ever-present in criminal settlement dynamics.

The Article concludes by briefly identifying a few implications of our analysis. By conceptually rooting criminal settlement agreements in the partial settlement frame, new solutions to old problems in research and policy debates may emerge. In thinking about how best to balance power between judges, prosecutors, and defendants, we must understand how and why parties structure their settlement agreements—including traditional plea bargains—the way they do. We are unable to address in this Article the legitimate concerns over the moral and political complexities that criminal settlements of all types entail.

2003) (per curiam) (holding that a prosecutor’s bargained-for joining of the defense’s recommendation need not be enthusiastic, if enthusiasm was not part of the bargain).

39 See Bibas, supra note 5, at 2533 (noting that “indeterminate sentencing plays on overoptimism and loss aversion to encourage bargains”).

40 The literature on these issues is immense and far-reaching. See, e.g., ROBERT A. KAGAN, ADVERSARIAL LEGALISM: THE AMERICAN WAY OF LAW 85 (2001) (“Plea bargaining . . . transforms the act of confession from a ritual of moral and social healing into a cynical game, reinforcing the criminal’s alienated view of society.”); John H. Langbein, Torture and Plea Bargaining, 46 U. CHI. L. REV. 3, 12 (1978) (“In twentieth-century America we have duplicated the central experience of medieval European criminal procedure: we have moved from an adjudicatory to a concessionary system.”).
but first things first. By understanding the dynamics of the criminal settlement process, both holistically and piece by piece, we take the constructive first step of disentangling and modeling what has become the central practice of the American criminal justice system.

I. THE INTERESTS AND CONSTRAINTS OF CRIMINAL SETTLEMENT

Criminal settlement is akin to civil settlement in its essential determinants and dynamics. This should not be surprising. In both contexts, parties bargain whenever possible toward a mutually beneficial agreement that modifies any of the less desirable default arrangements set by law.\(^{41}\) Mutual benefit does not mean that any negotiated outcome is objectively good, fair, or equal. Settlement will at times result in the innocent or faultless being imprisoned or paying to resolve accusations against them.\(^ {42}\) Instead, we use “mutual benefit” to refer to the net improvement both parties may perceive in striking a particular bargain in relation to their subjective valuation of their alternative options. In other words, parties ask: “Will this agreement make me better off than all other current—and potentially terrible—options available to me, including the possibility of not entering into any agreement at all and relying on default rules?” In effect, parties assess potential agreements relative to the position in which they believe they would be absent some form of settlement.\(^{43}\) In this way, “[e]ach party . . . trades the possibility of total victory for the certainty of avoiding total defeat.”\(^ {44}\) Criminal settlements, just like civil settlements, “are simply agreements between parties to a dispute that offer value to both on one or more of the following dimensions: reducing adjudication costs, mitigating losses due to risk, and maximizing ex ante expected returns.”\(^ {45}\)

But relative to the broad freedom to contract that characterizes civil settlement, the ambit of viable criminal settlements is more limited. As we

---

\(^{41}\) “Plea bargaining flows from ‘the mutuality of advantage’ to defendants and prosecutors, each with his own reasons for wanting to avoid trial.” Bordenkircher v. Hayes, 434 U.S. 357, 363 (1978) (quoting Brady v. United States, 397 U.S. 742, 752 (1970)).

\(^{42}\) The “innocence problem” in the criminal context and the “nuisance lawsuit” in the civil context are both well-documented in the American justice system. See, e.g., Covey, supra note 27, at 73; Randy J. Kozel & David Rosenberg, Solving the Nuisance-Value Settlement Problem: Mandatory Summary Judgment, 90 VA. L. REV. 1849, 1850 (2004).


\(^{44}\) Church, supra note 3, at 518; see also Alschuler, supra note 3, at 683 (“[T]he defendant abandons his chance of obtaining a sentence of zero, and the prosecutor abandons his chance of obtaining a sentence of life.”).

\(^{45}\) Prescott & Spier, supra note 4, at 78 (discussing civil settlement).
note above, civil litigants, outside of a few select circumstances involving class actions or substantial third-party effects, are free to enter into an almost infinite variety of settlement agreements with or without court say-so. Not so for criminal litigants. The legislature, the Constitution, or both empower a third party, the judge, with the ability to derail many plausible bargains—for example, by enforcing statutory minimum and maximum sentences when the parties would prefer another outcome—and to change the relative attractiveness of others—for instance, by exercising the “last move” in sentencing, which introduces significant uncertainty in any bargain involving outcomes. Judges as scene-stealers often limit the bargaining space but also potentially expand it when they are willing to play roles that would be difficult to reproduce privately at low cost in the civil context. Nevertheless, although party autonomy is an essential distinguishing factor between criminal and civil settlement, the underlying interests that motivate bargaining in both contexts remain in tune.

A. The Parties’ Interests in Criminal Settlement

Criminal settlement is the product of optimizing agents engaged in utility-maximizing “trade.” Each party possesses “assets” that the other may value, so they bargain with the goal of achieving mutual gains. In the most familiar scenario, prosecutors “buy” concessions from defendants by promising some benefit (e.g., a lesser charge or a shorter sentencing recommendation), and in turn defendants “sell” some or all of their procedural rights. A defendant’s procedural bargaining chips are many, with the most valuable often being the right to a trial, the right to withhold information, and the right to appellate review. The price that a prosecutor is willing to pay, and that a defendant is willing to accept, for waiving such rights is a function of the expected effects of the new terms on the costs, the risks, and the expected outcomes in the alternative—i.e., proceeding to trial without some

46 See Weisburst, supra note 6, at 55.
48 The Supreme Court has embraced this model: “A defendant can ‘maximize’ what he has to ‘sell’ only if he is permitted to offer what the prosecutor is most interested in buying.” United States v. Mezzanatto, 513 U.S. 196, 208 (1995); see also RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 10 (5th ed. 1998) (“[R]esources tend to gravitate toward their most valuable uses if voluntary exchange . . . is permitted.”).
49 See, e.g., Russell D. Covey, Plea Bargaining and Price Theory, 84 GEO. WASH. L. REV. 920, 924, 940 n.119 (2016).
agreement. If the proposed terms reduce litigation costs and risk-bearing losses and/or improve the anticipated outcome for either party, the prosecutor and the defendant are more likely to settle, at least if they can share the gains in trade. We should expect to observe in practice all agreements that can improve on the status quo with respect to these considerations, absent some legal limitation or other friction, such as negotiation costs.

Consider first the real burden of adjudicating criminal charges through a full-blown jury trial in a traditional courtroom with all of the innumerable constitutional and statutory protections to which a criminal defendant is legally entitled—what Justice Scalia once called “the exorbitant gold standard of American justice.” For the state, administering the criminal justice system is resource intensive. The costs of investigating and prosecuting crimes, securing defendants’ rights, maintaining often overcrowded courts and jails, and ultimately enforcing sentences all fall solely on the state. The government cannot (easily) abandon an accused’s sundry rights without their consent. Accordingly, the state has incentives to control systemic costs through piecemeal bargaining, lowering marginal costs one case at a time. As Chief Justice Burger observed, “If every criminal charge were subjected to a full-scale trial, the States and the Federal Government would need to multiply by many times the number of judges and court facilities.”

The finite nature of criminal justice resources ultimately constrains the prosecutor’s bargaining position. Research demonstrates a strong positive correlation between judicial caseloads and the rate of guilty pleas; that is, as the supply of criminal justice resources decreases, the prosecutorial demand for resource-efficient settlement increases. Relatedly, data show that pros-

---

50 See, e.g., Schulhofer, supra note 26, at 1980 (noting that “the expected costs of trial, the likelihood of success, and the expected trial sentence” determines a case’s value).

51 We are imagining an ex ante Pareto improvement. If an arrangement will only improve the situation for a single party (and make it no worse for the other party), the benefiting party may be able to induce agreement by sharing the gains by offering something—anything—to the other party, although fairness concerns may force a more equitable split in practice.


54 Richard Birke, Reconciling Loss Aversion and Guilty Pleas, 1999 UTAH L. REV. 205, 233 (“[P]rosecutors turn time-consuming and resource-expensive investigations and trials into time-efficient pleas, and in so doing, they can keep the numbers of cases closed and the numbers of new cases that come in at a manageable equilibrium.”).


56 See, e.g., Frank O. Bowman, III & Michael Heise, Quiet Rebellion II: An Empirical Analysis of Declining Federal Drug Sentences Including Data from the District Level, 87 IOWA L. REV. 477, 538 (2002) (noting the trend between guilty plea rates and judge caseloads); Covey, supra note 49, at 934 (graphing the relationship between the guilty plea rate and judicial caseloads from 1946 to 2003); Welsh S. White, A Proposal for Reform of the Plea Bargaining Process, 119 U.
ecutors who bring cases in federal districts with higher rates of judicial vacancies are more likely to negotiate toward guilty pleas (as well as offer more favorable discounts at sentencing) than their counterparts in jurisdictions with relatively fewer judicial vacancies. The study documenting this pattern quotes one prosecutor as reporting that the “resource in shortest supply [is] the number of district judges.”

In the criminal justice marketplace, the supply and demand of “system” resources necessarily influence the going rate for settling criminal cases. Prosecutors are more willing to offer defendants deep sentencing discounts when case congestion causes the costs of not settling to be especially high for the state—and, particularly and more realistically, for their own offices’ budgets.

Beyond the need to operate within and relieve stress on the state’s adjudicative resources, prosecutors acting in a professional capacity also bear significant case-level costs during litigation, particularly when they find it necessary to proceed to trial sans agreement. These are not personal financial costs like those civil plaintiffs must often bear, but rather the opportunity costs that necessarily accompany a fixed agency budget. The government budgets that support prosecutorial activities are not insignificant, but prosecutors can nevertheless bring only so many cases and often struggle to manage their caseloads. Prosecutors are therefore wont to dispose of their cases in ways that will maximally leverage their resources, which may mean settling cases involving lower levels of criminality quickly, allowing them to turn their attention to higher profile defendants. But whatever the precise tradeoff, prosecutors likely seek “the maximum deterrent punch [(or punch of some sort)] out of whatever resources are committed to crime control.”

PA. L. REV. 439, 448 (1971) (comparing Philadelphia and New York City and noting that “because more guilty pleas must be entered in New York [due to limited resources], the concessions offered to defendants are concomitantly increased”).

57 See Crystal S. Yang, Resource Constraints and the Criminal Justice System: Evidence from Judicial Vacancies, 8 AM. ECON. J. 289, 295 (2016) (quoting William Braniff, Local Discretion, Prosecutorial Choices and the Sentencing Guidelines, 5 FED. SENT’G REP. 309, 312 (1993)) (“Judicial vacancies require prosecutors to allocate limited resources among potential cases, such that the number of active judges in a district court is effectively a bottleneck on prosecution.”).


60 Frank H. Easterbrook, Criminal Procedure as a Market System, 12 J. LEGAL STUD. 289, 290 (1983); see also Albert W. Alschuler, The Prosecutor’s Role in Plea Bargaining, 36 U. CHI.
In practical terms, economizing on the time and money devoted to one case frees up resources for other cases; closing cases quickly and inexpensively is valuable to prosecutors because they can do more with less.\textsuperscript{61}

Prosecutors also have preferences that go beyond the traditional purposes of punishment—most notably, they often pursue opportunities to advance their career. Personal aspirations can affect how prosecutors evaluate individual settlement terms.\textsuperscript{62} For this reason, prosecutors may value a conviction itself—i.e., a “win”—more than they do any particular sentence.\textsuperscript{63} At the same time, prosecutors and defendants are not always playing a zero-sum game like civil litigants. Scenarios exist in which both parties will prefer lighter sentences, given their respective preferences. “[E]xtra months in prison are not like marginal dollars in civil cases,” William Stuntz recognized; “[o]nce the defendant’s sentence has reached the level the prosecutor prefers . . . adding more time offers no benefit to the prosecutor.”\textsuperscript{64} Whether this dynamic actually privileges more reliable convictions or instead longer average sentences, prosecutors nonetheless have views (either personal or informed by the macro policies of their respective offices) on the relative

\textsuperscript{61} Importantly, not all of a state’s criminal justice resource constraints operate on prosecutors. Funding and duties are divided between state actors (e.g., police, corrections), and prosecutors typically do not “pay” for many of the consequences of their decisions. See Daniel Richman, \textit{Prosecutors and Their Agents, Agents and Their Prosecutors}, 103 COLUM. L. REV. 749, 774–76 (2003). This failure of prosecutors to internalize the full costs of their bargaining results has been termed a “correctional free lunch.” See FRANKLIN E. ZIMRING & GORDON HAWKINS, THE SCALE OF IMPRISONMENT 140 (1991); see also Adam M. Gershowitz, \textit{An Informational Approach to the Mass Imprisonment Problem}, 40 ARIZ. ST. L.J. 47, 50–51 (2008) (arguing that line prosecutors should regularly receive information on the prison population so that they might consider it in negotiating plea bargains). That said, there may be reputational costs to prosecutors who burden judges and other actors with trying unreasonable cases or proposing settlements that externalize costs to other agencies to an excessive degree. See Yang, \textit{supra} note 57, at 294.

\textsuperscript{62} Of course, as individuals, prosecutors can be motivated by a host of unique and overlapping considerations. See Richman, \textit{supra} note 61, at 758 (“Convictions? Sentence-years? Deterrence? Agency prestige? Lifetime earnings? Leisure?”).


merits of certain settlement terms, and this affects bargaining preferences and behavior.\textsuperscript{65} So, be it a simple agreement to plead guilty to the crime alleged or some more complex arrangement, achieving a settlement helps prosecutors realize their various interests.

Like prosecutors, criminal defendants also aim to minimize the adjudication costs associated with their defense. Such costs are readily evident for defendants who retain private counsel,\textsuperscript{66} but even indigent defendants suffer financially—they are charged for being prosecuted. Since the 1990s, at least forty-three states and the District of Columbia have enacted laws requiring defendants to repay a portion of the costs associated with their public defense, with some charging a flat fee and others requiring repayment of all accrued costs.\textsuperscript{67} Likewise, certain states assess fees associated with the exercise of other constitutional rights, such as compelling a witness’s testimony,\textsuperscript{68} or simply empaneling a jury for trial.\textsuperscript{69} But more important than any direct fee (which realistically may never be collected) are the psychological costs and delay associated with demanding a trial on the merits—which may not occur until after the defendant has spent months or even years in jail. Appearing for arraignment as well as preparing and being present at all proceedings also impose opportunity and reputational costs on defendants and their families. For many, simply enjoying their due process rights can generate costs far in excess of the punitive burdens that are likely to result from conviction.\textsuperscript{70} Indeed, for some defendants “the process is the punishment.”\textsuperscript{71} All criminal defendants have an interest in reducing adjudication costs, broadly construed, and therefore pursue feasible agreements that deliver more bang for the proverbial buck.

In addition to seeking compromises to limit costs, criminal defendants also aim to minimize the expected sanctions associated with any conviction.

\textsuperscript{65} See Richman, supra note 61, at 756 (explaining, in the federal context, the influence of a given administration’s criminal justice agenda).

\textsuperscript{66} See, e.g., Pamela S. Karlan, Contingent Fees and Criminal Cases, 93 COLUM. L. REV. 595, 599 (1993) (noting that most private criminal defense agreements involve “either entirely prepaid fixed fees . . . or the payment up front of a substantial retainer”).

\textsuperscript{67} States vary greatly in how they assess these costs and the extent to which they factor in a defendant’s ability to pay. See John D. King, Privatizing Criminal Procedure, 107 GEO. L.J. 561, 571–72 (2019) (reviewing criminal justice “user fees”).

\textsuperscript{68} Virginia, for instance, charges a $50 fee for a witness’s appearance at a hearing or trial if the defendant is found guilty of the charges for which the witness is summoned. Id. at 578 (citing VA. CODE ANN. § 19.2-187.1(F) (West 2017)).

\textsuperscript{69} Arkansas, for example, charges a nonrefundable $150 jury fee. Id. at 580 (citing ARK. CODE ANN. § 21-6-403(b)(1) (West 2013)).


\textsuperscript{71} Id. at 294.
All else equal, defendants prefer to spend less rather than more time behind bars.72 At the extreme, defendants may seek settlement to avoid the risk of capital punishment.73 But there are many other forms of punishment beyond the deprivation of life or liberty, including monetary fines, property forfeiture, and mandatory restitution. Collateral consequences under the law, such as sex offender registration and notification requirements, deportation, and the loss of the right to bear arms, to name just a few, also flow from many convictions.74 Some defendants may view these collateral effects as more severe than any formal court-ordered fine or incarceration.75 Criminal convictions can also produce losses associated with having a criminal record, including fewer employment and housing opportunities.76 So while defendants bargain in the hopes of mitigating litigation costs, they also seek to improve the likely outcome of their dispute (conditional on there being a dispute) and stand willing to bargain with the currency of the marketplace—their statutory and constitutional rights—to improve their position.

Both prosecutors and defendants assess costs (e.g., time, money, and anxiety) and outcomes (e.g., convictions, sentences, and any collateral consequences) under conditions of uncertainty and asymmetric information. Neither party can be certain that a conviction or an acquittal will occur at

---

72 This does not mean that some defendants are not willing to spend more time imprisoned to avoid other costs. Preferences across punishment types can vary among rational actors. See generally DERYCK BEYLEVELD, THE EFFECTIVENESS OF GENERAL DETERRENTS AGAINST CRIME: AN ANNOTATED BIBLIOGRAPHY OF EVALUATIVE RESEARCH (1978).


75 So weighty are such consequences that the Supreme Court has recognized a constitutional right to be informed of them—or at least some of them—in advance of a guilty plea. See, e.g., Brian M. Murray, Beyond the Right to Counsel: Increasing Notice of Collateral Consequences, 49 U. RICH. L. REV. 1139, 1161 (2015).

76 For this reason, the potential for expungement of a conviction might be of great value to a defendant. Although there is no general federal expungement statute, state courts with such statutes have recognized expungement as a legitimate bargaining chip. See, e.g., People v. Acuna, 92 Cal. Rptr. 2d 224, 228 (Ct. App. 2000) (rejecting an appellant’s request for expungement, noting that there was “no express provision in his plea bargain that mentions expungement,” nor was expungement “clearly part of the parties’ understanding”); Commonwealth v. Lutz, 788 A.2d 993, 1001 (Pa. Super. Ct. 2001) (“In the absence of an agreement as to expungement, Appellant stands to receive more than he bargained for in the plea agreement if the dismissed charges are later expunged.”); see also J.J. Prescott & Sonja B. Starr, Expungement of Criminal Convictions: An Empirical Study, 133 HARV. L. REV. 2460, 2470 (2020) (showing the logistical and legal difficulties eligible individuals face in obtaining expungement).
the outset of a prosecution or even midway through a trial. Nor can parties be certain about what will flow from any verdict, nor what costs will accrete unexpectedly en route to a case’s final disposition. As individuals, defendants and prosecutors are sensitive to risk, although some defendants may be risk-loving and some prosecutors may approach risk neutrally. By and large, however, parties work to mitigate their risks in line with their preferences to maximize their overall utility. A prosecutor may have an interest in maintaining a high conviction rate and, for that reason, will be motivated to minimize any risk of acquittal by offering terms that reduce a defendant’s expected sanction by enough to induce mutual agreement.\(^77\) Likewise, a defendant may wish to minimize the risk of an outsized punishment and will often offer to trade away the chance of a complete acquittal.\(^78\) Parties suffer risk-bearing costs in relation to the perceived strength or weakness of the evidence and the likelihood of a particular outcome.

In sum, the “inputs” to criminal settlement behavior are the likely litigation costs (both psychological and financial), the expected outcomes post-litigation, and the uncertainty associated with any set of terms (including an absence of terms in the case of a failure to settle), all channeled through parties’ respective preferences—just as in civil litigation. Parties optimize over potential settlement packages by taking into account how much they individually value the costs and expected outcomes as well as how much they dislike the riskiness of a particular path.\(^79\) If there is a set of terms by which the parties can mutually benefit relative to the default arrangement—a “naked” trial—one party will begin by making an offer, assuming negotiation costs are not too high. It might be that no agreement can improve the situation for both parties, in which case settlement negotiations will fail. If there is only one set of terms that offers mutual improvement for the parties, the only questions are whether they will be able to “find” those terms, given the real-world challenges of bargaining and trade, and whether the preferred arrangement of terms is legally permissible.

Finally, if there are many permutations of mutually beneficial and legally permissible terms that the parties could assemble from among the various procedural adjustments, issue alterations, and outcome modifications on the table, the parties must simply negotiate over which combination to

\(^77\) See Easterbrook, \textit{supra} note 60, at 289 (explaining this effect).

\(^78\) Stuntz, \textit{supra} note 64, at 2560 (“[P]lea bargains outside the law’s shadow depend on prosecutors’ ability to make credible threats of severe post-trial sentences.”).

\(^79\) See, e.g., Russell D. Covey, \textit{Plea-Bargaining Law After Lafler and Frye}, 51 DUQ. L. REV. 595, 613 (2013) (noting that “the size of the plea discount offered, the probability of conviction at trial, and the costs of contesting the case” are a defendant’s key considerations in determining whether to plead guilty).
accept in order to maximize their individual share of the resulting collective surplus.\textsuperscript{80} For example, if the parties are mutually pessimistic about their prospects at trial, they are relatively more likely to settle, if only to avoid their private costs of adjudication.\textsuperscript{81} Alternatively, if a prosecutor and a defendant are mutually optimistic about their chances, they may fail to reach any agreement, choosing instead to gamble on a favorable adjudicated outcome under the background default rules. Absent an inefficient failure of trade, this scenario implies that no settlement terms could produce a sharable surplus. Nevertheless, defendants and prosecutors do not bargain with complete autonomy. They cannot always throttle their uncertainty and costs to optimize mutual net returns in the way parties to civil litigation often do. Instead, bargaining for criminal settlement is distinctive in that it is circumscribed by a necessary third party—the judge.

\textbf{B. The Judge as a Constraint on Settlement}

Prosecutors and defendants have an inherent interest in pursuing settlement whenever terms exist that will improve their respective lot relative to the default arrangement of a full criminal trial. Yet unlike in the civil context, where parties are remarkably unfettered, there are forces in criminal litigation that constrain parties’ bargaining options. Prosecutors and defendants deal in imprecise commodities, swapping procedural rights for various concessions in a world of statutorily set penalties. Moreover, any agreement the parties reach—no matter how meticulously they construct it to ensure their mutual benefit—cannot entirely eliminate uncertainty because it must still pass judicial muster. Criminal litigants can never \textit{fully} settle their dispute; rather, parties bargain and partially settle in the looming shadow of a judge serving as the gatekeeper to the state’s punitive power.

Federal Rule of Criminal Procedure 11 governs plea agreements at the federal level.\textsuperscript{82} The Rule explicitly regulates deals in which the defendant pleads guilty, but it is far broader in its practical influence and implications.

\begin{footnotesize}


\textsuperscript{82} As we note in the Introduction of this Article, we focus on federal rules and plea bargaining in the federal courts for the sake of exposition. States have their own substantially similar analogs to Federal Rule of Criminal Procedure 11. Some states mirror the federal system of permitting sentencing proposals to be either binding or nonbinding on the judge, \textit{see}, e.g., TENN. CRIM. P. 11, whereas other states leave the matter to judicial discretion, \textit{see}, e.g., N.Y. CODE CRIM. PROC. § 220.60(3) (McKinney 2021). Regardless, the crucial ingredient for our argument is that all states require some form of judicial review and acceptance of criminal plea agreements.
\end{footnotesize}
Rule 11 describes the concessions the government may offer in return for a defendant’s waiving some or all of their procedural rights. A prosecutor can agree to: (A) not bring, or move to dismiss, other charges; (B) recommend, or agree not to oppose, the defendant’s request that a particular sentence or sentencing range is appropriate or that a particular provision of the federal Sentencing Guidelines, a policy statement, or a sentencing factor does or does not apply; or (C) agree that a specific sentence or sentencing range is the appropriate disposition of the case or that a particular provision of the Sentencing Guidelines, a policy statement, or a sentencing factor does or does not apply. This is by no means an exhaustive list of a prosecutor’s bargaining chips, but (B) and (C) do serve to mark the boundaries of what types of outcome concessions are legitimately on the table.

The defendant, in return, is permitted to waive some of their procedural rights—including the right to trial by jury, the right to remain silent, or the right to have the charges proven by the state. Not all rights can be waived, however. For instance, defendants cannot waive the constitutional prohibition on double jeopardy, nor can they accept an alternative standard of review for determining guilt (such as “preponderance of the evidence’’). Further, defendants cannot consent to punishment in excess of what a statute permits, and they cannot agree to a proceeding that considers constitutionally impermissible factors (such as race or gender). Indeed, at some basic level, core notions of procedural justice must be maintained. In the colorful language of the U.S. Court of Appeals for the Seventh Circuit: “[T]here are limits to waiver; if the parties stipulated to trial by 12 orangutans the defendant’s conviction would be invalid notwithstanding his consent, because some minimum of civilized procedure is required by community feeling regardless of what the defendant wants or is willing to accept.”

---

83 FED. R. CRIM. P. 11(c)(1)(A)–(C).
84 See generally King, supra note 30 (reviewing those features of criminal procedure that can and cannot be bartered away).
86 See Fisher, supra note 30, at 945.
87 See, e.g., United States v. Feichtinger, 105 F.3d 1188, 1190 (7th Cir. 1997) (recognizing that a criminal defendant cannot agree to be sentenced to a term of incarceration “in excess of the statutory maximum”).
88 See United States v. Hicks, 129 F.3d 376, 377 (7th Cir. 1997).
89 See United States v. Josefik, 753 F.2d 585, 588 (7th Cir. 1985). Realize, though, that while a judge will not permit parties to empanel a jury of monkeys to resolve their dispute, a judge may allow parties to alter other procedures or reframe issues as a substitute capable of mimicking the perceived benefits or outcome of such a whimsical procedure. When it is in their mutual interest, parties will try to bargain around any restriction enforced by a judge.
allows plenty of play, but the universe of available criminal settlements has at least some (albeit ill-defined) boundaries.90

The judge is responsible for policing these boundaries, which happens in two main ways. First, Rule 11 (or an equivalent state counterpart) mandates that the judge seek a knowing and intelligent waiver of a defendant’s rights and advise the defendant of the maximum penalties and any mandatory minimum penalties, including certain collateral consequences, flowing from a guilty plea.91 The judge must also ensure that the charge has some factual basis, meaning “that the conduct which the defendant admits constitutes the offense charged in the indictment or information or an offense included therein to which the defendant has pleaded guilty.”92 This examination aims to “protect a defendant who is in the position of pleading voluntarily with an understanding of the nature of the charge but without realizing that his conduct does not actually fall within the charge.”93 Critically, a defendant cannot barter away their right to this judicial examination.94

The second way a judge polices criminal settlements is by imposing reasonable sentences on an individual convicted of a crime despite any party agreement to the contrary. If the parties’ recommended sentence departs from the Sentencing Guidelines’ range, for example, a judge may only impose it if there are “justifiable reasons,” which the judge must then “set forth with specificity in the statement of reasons form.”95 In 2005, when the Supreme Court ruled in United States v. Booker that the federal Sentencing Guidelines were not mandatory but advisory, it effectively expanded judicial discretion by allowing judges to weigh almost any factor during sentenc-

90 See Daniel P. Blank, Plea Bargain Waivers Reconsidered: A Legal Pragmatist’s Guide to Loss, Abandonment and Alienation, 68 FORDHAM L. REV. 2011, 2012 (2000) (arguing that courts have been unable to articulate a cogent explanation for why some rights are waivable and others are not).
91 FED. R. CRIM. P. 11.
93 Id. (quoting FED. R. CRIM. P. 11 advisory committee’s note to 1966 amendment).
94 See FED. R. CRIM. P. 11(b)(1) (mandating the plea colloquy and its substance).
95 U.S. SENT’G GUIDELINES MANUAL § 6B1.2(b)(2) (U.S. SENT’G COMM’N 2018). “Justifiable reasons” include: “(1) [the] circumstances of the offense and the history and characteristics of the defendant”; “(2) the need for the sentence imposed . . . (A) to reflect the offense’s seriousness, promote respect for the law,” and “provide just punishment,” “(B) to afford adequate deterrence,” “(C) to protect the public from further crimes of the defendant,” and “(D) to provide the defendant with needed educational or vocational training, medical care, or other correctional treatment”; “(3) the kinds of sentences available”; “(4) the kinds of sentence and the sentencing range” for the offense; “(5) any pertinent policy statement issued by the Sentencing Commission”; “(6) the need to avoid unwarranted sentence disparities”; and “(7) the need to provide restitution.” See 18 U.S.C. § 3553(a)(1)–(7).
Precedent even permits a sentencing judge to consider personal policy disagreements with the Sentencing Guidelines so long as the ultimate sentence imposed is nevertheless “reasonable.” And although a sentence falling within the Guidelines is presumed reasonable, one falling outside the Guidelines is not presumed unreasonable.

Judges wield many powers capable of shaping final outcomes. For instance, judges can look beyond the parties’ presentation and consider charges dropped during the parties’ negotiations. Judges may accept a criminal settlement providing for the dismissal of any charge—or an agreement not to pursue potential charges—only when “the remaining charges adequately reflect the seriousness of the actual offense behavior and . . . accepting the agreement will not undermine the statutory purposes of sentencing or the sentencing guidelines.” Thus, even if the prosecutor agrees to dismiss a charge, the judge may consider the underlying acts as “[r]elevant [c]onduct” and therefore as inputs to any sentence. True, the policy guidelines caution judges to defer to prosecutors, but they also add that “when the dismissal of charges or agreement not to pursue potential charges is contingent on acceptance of a plea agreement, the court’s authority to adjudicate guilt and impose sentence is implicated.” The U.S. Sentencing Commission’s report for 2019 shows that 7.6% of upward Sentencing Guideline departures were due to dismissed and uncharged conduct. That figure was 18.1% in 2016. Judges can and do impose sentencing outcomes beyond the apparent bounds of parties’ settlement terms.

---

99 U.S. SENT’G GUIDELINES MANUAL § 6B1.2(a).
100 Id.
101 Id. cmt.
104 United States v. Fitch provides an extreme, yet illuminating, example of the potential impact of uncharged conduct on sentencing. 659 F.3d 788, 790 (9th Cir. 2011). There, the government charged the defendant with sixteen counts of fraud against his missing wife, who had disappeared under suspicious circumstances. Id. The jury convicted the defendant on those fraud charges. Id. During sentencing, however, the prosecution requested an upward departure based on the uncharged allegation that the defendant had actually murdered his wife in furtherance of his fraudulent acts. Id. at 793. A U.S. District Court Judge for the District of Nevada concluded that the defendant had in fact murdered his wife and sentenced him to 262 months, which was the statutory maximum for the fraud charges. Id. at 794. The U.S. Court of Appeals for the Ninth
Despite this considerable discretion, some commentators stress that judges tend toward “rubberstamping” criminal settlement agreements rather than closely scrutinizing them.\(^\text{105}\) There are several reasons why some judges may take a less-than-active role in policing bargains. One of the most noted is that overburdened judges have an interest in approving settlements to lighten their own heavy caseloads.\(^\text{106}\) Time and effort matter to judges, too. Imposing a sentence recommended by the parties and within the applicable federal Sentencing Guideline range allows a judge to avoid the additional work of justifying the imposed punishment and makes it less likely their decision will be appealed and, if appealed, reversed.\(^\text{107}\) Beyond this potentially self-serving efficiency, a judge may also feel strongly about ensuring that a prosecutor and a defendant get the benefit of their bargain without undue judicial intrusion, essentially ignoring any third-party effects on crime rates or communities and the possibility of untoward but undetected pressure on the defendant.\(^\text{108}\) Parties should have the power, a judge might believe, to bargain over how best to structure the resolution of their dispute in line with their own interests.

But perhaps the most important reason why judges may tend not to actively (or at least successfully) police criminal settlements is that parties are often adept at evading meaningful judicial review.\(^\text{109}\) By manipulating the contours of their dispute, prosecutors and defendants can present scenarios that hide material facts that a judge might otherwise consider in exercising discretion. For instance, parties can bargain over which and how charges are presented to a judge, provide joint sentencing factor recommendations, and offer fabricated accounts of the defendant’s criminal history so as to con-

\(\text{Circuit affirmed the sentence under the due deference standard. \textit{Id.} at 799. For a broad discussion of } \textit{Fitch}, \text{see generally Robert Alan Semones, Note, } \textit{A Parade of Horribles: Uncharged Relevant Conduct, the FederalProsecutorial Loophole, Tails Wagging Dogs in Federal Sentencing Law, and United States v. Fitch, 46 U.C. Davis L. Rev. 313, 348 (2012).}\)

\(^{105}\) McConkie, \textit{supra} note 29, at 66; \textit{see also} Frank S. Gilbert, Commentary, \textit{The Probation Officer’s Perception of the Allocation of Discretion}, 4 Fed. Sent’g Rep. 109, 109 (1991) (quoting one jurist who likened his role “to that of a ‘notary public’”).


\(^{108}\) This was not always the case. It was not until the nineteenth century, as assistance of counsel in criminal proceedings became more common, that courts felt “[r]elieved of the duty to protect the defendant from his own bad judgment, [and] . . . could not as easily justify disturbing agreements to forego procedural protections.” King, \textit{supra} note 30, at 121.

\(^{109}\) \textit{See, e.g., HEUMANN, \textit{supra} note 32, at 43.}
strain the range of plausible punitive outcomes at the judge’s disposal.\footnote{See, e.g., King, supra note 19, at 297; Stephen J. Schulhofer & Ilene H. Nagel, Plea Negotiations Under the Federal Sentencing Guidelines: Guideline Circumvention and Its Dynamics in the Post-Mistretta Period, 91 NW. U. L. REV. 1284, 1292 (1997).} Importantly, the presentencing reports of probation officers, which are meant to inform the judge about relevant sentencing factors and to provide an independent check on the parties, are often incomplete or lack sufficient evidence to justify a judge’s reliance on them when they contradict the parties’ factual stipulations.\footnote{King, supra note 19, at 303.} Further inquiry is time consuming and costly. Without a reliable alternate account, a judge may feel significant pressure to accept the parties’ predictably misleading presentation.

Even though prosecutors and defendants can successfully evade judicial review of settlement terms in some instances, attempting to do so is both costly and risky. For this reason, it would be wrong to ignore entirely the constraints that judicial discretion places on criminal settlement behavior. The specter of an inquisitive judge who might reject an agreement between the parties or, worse, partially unwind or alter the agreement and then accept it—thereby defeating the parties’ intentions—informs the way prosecutors and defendants structure and accomplish their bargains. Parties can never be certain of how their settlements will play out before the court, but this in no way implies that settlement cannot continue to be mutually beneficial. Parties simply incorporate the judicial wildcard into their calculations and further adjust their preferred terms according to their interests in cost and risk minimization and value maximization.

II. CRIMINAL SETTLEMENT AS PARTIAL SETTLEMENT

All agreements between prosecutors and defendants—including traditional plea bargains—are best understood as partial settlements. All settlement exists along a spectrum, with “pure trial” (no party agreement at all—or a tacit agreement not to agree) at one end, in which background default rules and endowments govern completely, and “pure settlement” at the other, in which parties wholly supplant defaults with their own rules and outcomes, usually resolving the dispute outright.\footnote{Prescott & Spier, supra note 4, at 62.} Prosecutors and defendants are unable to “purely settle” out of court as parties often do in the civil context, but they are just as able to enter into other settlement agreements, exchanging certain concessions whenever doing so is to their mutual benefit. Conceptually, these concessions can be divided into three nonexclusive categories of modifications depending on which default terms they address.
Agreements can modify (A) the procedures the court and parties will use to reach a decision in the case, (B) the substantive issues to be resolved by any adjudication, and (C) the outcomes potentially available.

Procedure-, issue-, and outcome-modification agreements can be restyled as simply categories of deal terms, which parties can combine to alter costs, adjust uncertainty levels, or shift expected outcomes. Terms can work together as mutually reinforcing complements or offer alternative strategies as substitutes. Although some terms can be characterized as belonging to more than one of these three categories, and parties regularly trade terms from multiple categories simultaneously, we address each category separately. By defining boundaries, we can review more clearly how these modifications operate to generate surplus for parties. It also allows us to flesh out the theoretical implications and practical consequences of criminal settlement in the shadow of significant judicial discretion. While each type of modification operates differently, they all work to sidestep the constraints and costs imposed by background rules so that parties can make the most of their situation given the judge’s role and power.

A. Procedure-Modification Agreements

A procedure-modification agreement is one in which the parties agree to change the rules by which their dispute will be resolved. That is, the parties agree to operate according to procedures that diverge from those that a statute, constitution, common law, or court rule establishes for determining the defendant’s guilt or punishment. Procedure modifications are central to most criminal settlement agreements because often a defendant’s most valuable bargaining assets are procedural in nature (or can be so construed), and in practice prosecutors can most readily conserve resources and reduce risk by circumventing certain default procedures.113 Criminal processes can be modified through agreements that include a bundle of procedural changes, such as a defendant agreeing to forgo adjudication on guilt completely,114 or à la carte procedural changes, such as a defendant agreeing only to waive the right to appeal their conviction or sentence. Again, parties agree to mod-

113 Defendants regularly bargain away their procedural rights in bulk by deciding to decline a trial on the merits and advance with little delay to sentencing by the judge, thereby creating value for the state by bypassing trial and its associated burdens.

114 A defendant forgoing adjudication altogether by entering an open guilty plea amounts to a settlement agreement in our view because prosecutors can object to any guilty plea, though this rarely if ever occurs. Defendants sometimes plead guilty unilaterally in hopes of receiving a lenient sentence from the judge or reducing their litigation costs. Of course, a guilty plea is far more often a term the defendant exchanges with the prosecutor for something of value.
ify default procedures to advance their mutual interests by lowering costs, reducing risk, or improving ex ante returns.

The first-to-mind example of what can be described as a procedure-modification is wholly predictable: the defendant offers to plead guilty to a charge, completely waiving all procedural rights to any adjudication of guilt (but not necessarily other post-conviction procedural rights), and to proceed directly to sentencing by a judge, which, if unaccompanied by any other agreement between the two parties, is simply called an “open plea.”115 Importantly, waiving adjudicatory procedures is essentially outcome determinative, so entering a guilty plea can also be characterized as (and overlaps with) an issue modification. A guilty plea alone is akin to civil litigants agreeing to a bifurcated lawsuit in which the question of liability is resolved but the issue of damages remains; the parties have agreed to resolve only a subset of their dispute by adjudication.116 But regardless of how we label it, just as in civil litigation, an open guilty plea allows both sides to avoid the litigation costs and risks associated with the default processes for adjudicating responsibility and instead frees them to concentrate their resources on influencing the sentence the decision-maker must select.

As an extreme, blunt form of procedure or issue modification, a guilty plea may not be the obvious place to start our discussion of the wide variety of criminal settlement terms. Yet a defendant’s agreement to plead guilty is so common that it is not just assumed to be the most frequent form of criminal settlement; rather, the traditional plea-bargaining framework has completely eclipsed broader notions of criminal settlement in academic and policy discussions.117 Nevertheless, even typical plea deals illustrate the underlying mechanisms of our broader theory of criminal settlement. Agreements by a defendant to skip the default procedures for determining guilt are of-

115 The perceived benefits of open pleas from the defendant’s perspective will turn on, among other considerations, the strength of the state’s evidence as well as the judge’s expected response at sentencing. See Nancy J. King & Michael E. O’Neill, Appeal Waivers and the Future of Sentencing Policy, 55 DUKE L.J. 209, 240 n.100 (2005) (“[I]t depends on the judge, whether you want an open plea. . . . If I went in with an open plea and the government is going to present all of this relevant conduct and the court might be receptive to finding by a preponderance that the relevant conduct occurred, it’s a risk. [It is d]ifferent if there is no relevant conduct. If there is, you’d want to stipulate.” (quoting a phone interview with a public defender) (alterations in original) (citation omitted)); see also id. at 233 n.87 (“[A]bout 25–30% of the open pleas are in meth cases. The penalties are so outrageous, and they have a hard time proving quantity. And even with acceptance they don’t give you much.” (quoting a phone interview with a public defender) (alteration in original)).


117 See supra notes 24–29 and accompanying text.
ten—though not inevitably—exchanged for a prosecutor’s reciprocal promise to the defendant, such as an offer to drop certain charges or to propose to the judge that a less severe sentence is more appropriate. These are issue- or outcome-related terms, and in practice they are fellow travelers with procedure modifications, best thought of as highly complementary to a guilty plea. That is, issue and outcome modifications may enhance the perceived value of a given procedure modification. In fact, additional settlement terms of this sort may be necessary to render the entire package of modifications attractive to both parties (and implicitly the judge).

But prosecutors and defendants agree to modify procedures in many other ways beyond guilty pleas, and examples show how fine-grained such procedural settlements can be. For instance, short of waiving adjudication altogether, a defendant can maintain their innocence while agreeing to submit the issue of guilt to an alternative adjudicator. In the federal courts, options include a jury of fewer than twelve persons (though not less than six), an Article III district court judge, or a federal magistrate judge. The parties may simultaneously prefer one of these factfinders, a preference that may be attributable to each of them entertaining optimistic, and therefore inconsistent, beliefs that the new adjudicator will produce a more favorable outcome for them than the default decision-maker. Like a deal to waive trial, an agreement regarding the type of adjudicator is often (but not always) coupled with additional procedure modifications. For instance, at least in some courts, bench trials are regularly accompanied by the defendant’s waiver of their right to cross-examine witnesses or their right to testify in their own defense. Truncating proceedings in this way can save resources for both parties (and possibly reduce risk) while still permitting them to speculate on their mutually optimistic beliefs about the likely outcome by adjudicating the remainder of their dispute.

118 Federal statistics on the rate of guilty pleas entered without any agreement from the prosecutor—an “open plea”—are seemingly unreliable. See King & O’Neill, supra note 115, at 228 n.76. Yet logic and informal observation suggest that open pleas are comparatively rare, as parties will typically benefit more from bargaining across multiple dimensions.
119 FED. R. CRIM. P. 23(a) (allowing waiver of a jury trial with the consent of the government and approval by the court).
123 See King, supra note 30, at 155.
124 See Nancy J. King et al., When Process Affects Punishment: Differences in Sentences After Guilty Plea, Bench Trial, and Jury Trial in Five Guidelines States, 105 COLUM. L. REV. 959, 979 n.77 (2005) (quoting a Washington state public defender as stating that a bench trial “doesn’t have the flavor of a trial [and] is really like the conditional plea” (alteration in original)).
125 As one public defender explained the benefits of a bench trial:
The adjudicative procedure the parties select for determining guilt can have a dramatic effect on any final sanction the court is likely to mete out, which provides easy insight into why settling parties “care” about procedures—procedures influence outcomes. To offer one example to make the point, a sizeable literature explores the so-called “trial penalty,” in which judges supposedly inflict harsher sentences on those defendants who choose to proceed to full adjudication rather than enter a guilty plea; indeed, some studies of federal court data show an average sixty-four percent longer sentence for those who stick with a default jury trial.126 In general, sentences that follow a guilty plea are likely to be the least punitive, sentences that follow jury trials are likely to be the most punitive, and sentences that follow bench trials fall somewhere in between.127 This sentencing disparity is surely a strong motivator for a defendant, even one not financially responsible for their own representation, to enter into a procedure-modification agreement. The fact that parties often harbor divergent opinions about how a given adjudicator is likely to respond to a guilty plea can account for their ability to come to terms: “The defendant may plead guilty expecting that the judge will find him sympathetic and sentence him to probation,” while “[t]he prosecution, in contrast, may hope that the same plea will result in one year’s imprisonment.”128 By swapping out the adjudicator, the parties mutually optimize given their respective beliefs and preferences.

If the jury feels really strongly, it puts pressure on the judge at sentencing. If the defendant has committed a really heinous crime, he might want a bench trial. There is less press coverage, [which] leaves [the] judge with less pressure to respond harshly. So if you have someone with no redeeming character, [you] can spend less time in front of the jury seeking an acquittal that is unlikely anyway, [and] more time arguing about sentence. Bench trials tend to be really horrible crimes. The defense attorney thinks, ok, I can shorten this up by going to bench trial and it may help out at sentence.

Id. at 981 (alterations in original). And, when asked why the prosecutor might agree, the defender responded: “[T]he prosecutor knows he’s not going to lose the conviction, and it is more efficient. Plus prosecutors want to avoid the risk of having one lone juror hang it up. He’s handed an easier conviction and will take it, even if sentencing might be a bit more lenient.” Id. (alteration in original).


128 Bibas, supra note 5, at 2500.
Rather than bypassing adjudication altogether or agreeing to an alternative decision-maker, the parties may prefer to operate largely within the confines of standard trial procedures but with relatively minor *a la carte* modifications concerning discrete aspects of their proceedings. That is, a prosecutor and a defendant might mutually consent to waive limited trial rights (such as by agreeing to specific trial time limits, witness rules, or evidentiary concessions) in exchange for some amount of limited prosecutorial leniency while nevertheless still adjudicating their dispute. An interesting example of this identified by Gregory Gilchrist involved a murder defendant waiving his constitutional right to confront an unavailable medical examiner who had performed an autopsy of the victim in exchange for the prosecutor not entering certain graphic photographs of the victim into evidence. These kinds of seemingly small procedural concessions may have real payoffs by allowing the parties to preserve resources, mitigate risk, and take advantage of their private, subjective information while still adjudicating (mostly) within the traditional procedural system.

An especially common (and useful) *a la carte* procedure-modification term is an appeal waiver. A defendant’s right to appeal any conviction, sentence, or both is valuable to the government because the appeals process can be expensive and time consuming. It also poses a risk that the resources invested in obtaining the conviction will all be for naught. Prosecutors may therefore be apt to make more valuable and substantial concessions in exchange for a waiver of this procedural right. Prosecutors offer various forms of consideration (indeed, modifications of all types) such as “[binding] ‘C pleas,’ downward departures, safety-valve credits, and a variety of stipulations to facts that determine[] sentences under the Guidelines” in exchange for a defendant agreeing to waive the right to appeal. Importantly, prosecutors may offer some of these benefits even when the defendant insists on

---


130 Gilchrist, *supra* note 30, at 622.


132 See Gilchrist, *supra* note 30, at 623 (arguing that “to the extent counsel can identify ways of streamlining the trial, or even making the trial less burdensome on the prosecution or particular witnesses, doing so can be exchanged for leniency”).

proceeding to a full trial in every other sense. 134 That is, the parties engage in a typical criminal trial conducted according to default procedures for adjudicating guilt but also ex ante “settle” the appeal (usually in exchange for an outcome-modification concession from the prosecutor) because the prosecutor values appellate rights more than the defendant does. 135 A guilty plea is not necessary for both parties to advance their interests.

Appeal waivers have other settlement benefits as well. They can enhance—i.e., complement—the usefulness of other procedural terms, making them relatively more likely to occur together as part of a settlement agreement. Consider a type of agreement reached during trial in which “the defendant and the government agree that it is in their interest to dispense with a certain procedure so long as the defendant waives any right to appeal his conviction based on the absence of that procedure.” 136 Nancy King offers an illuminating example of this exchange out of the U.S. Court of Appeals for the Ninth Circuit; there, a prosecutor “[c]orrectly anticipating that [the defendant] would claim that his attorneys were ineffective in allowing him to waive his presence, . . . secured, in open court, [the defendant’s] waiver of any claim that counsel was deficient in allowing him to waive his presence at jury selection.” 137 Such a waiver ensures that the defendant cannot use their appellate rights to unwind an earlier understanding. Just as it can be rational for an army to burn the bridge behind it, an appellate waiver can be a useful commitment device, creating value for a prosecutor, a part of which a defendant can then demand at trial in the form of some other favorable

134 King, supra note 30, at 148 (“A defendant convicted after trial may decide that the prosecution’s recommendation of a lower sentence is worth giving up the right to appeal his conviction, his sentence, or both.”).

135 Or consider that the parties might agree at some early stage that the defendant will not seek to expunge or seal conviction records in exchange for some prosecutorial concession. Cf. Andrea Amulic, Note, Humanizing the Corporation While Dehumanizing the Individual: The Misuse of Deferred-Prosecution Agreements in the United States, 116 MICH. L. REV. 123, 150–51 (2017) (discussing the benefits of expungement and the different approaches states employ for seeking it). Such a hypothetical procedural agreement could complement an issue modification even without a plea, such as the prosecutor agreeing to a lesser charge. Cf. supra note 76.

136 Id. at 161 n.151 (quoting Campbell v. Wood, 18 F.3d 662, 673 (9th Cir. 1994) (en banc)). Appeal waivers can take many forms. A defendant’s agreement to waive their right to challenge a sentence might be used separately as a substitute or in conjunction as a complement with a waiver of an ineffective assistance claim to tailor the agreement in line with the parties’ interests. See King & O’Neill, supra note 115, at 231, 244 (noting that 65% of a random sample of federal plea agreements include some form of waiver of review, with approximately 28% of these agreements including a specific exemption for ineffective assistance claims).
term. In effect, appeal waivers liberate parties to construct arrangements that may not otherwise withstand judicial scrutiny. 138

These previous examples all involve parties continuing to operate within the established set of default procedures or modifying them around the edges. According to our theory, these partial settlements create value for both sides by reducing risk, lowering litigation costs, or improving expected outcomes. Their “around the edges” quality hints that most if not virtually all criminal trials are “partially settled” because even small pockets of potential mutual improvement can be accessed by the parties. Note, though, that litigants need not always design new procedures from the ground up. Instead, jurisdictions around the country have regularly offered cohesive procedural packages—“pre-fab” alternatives to default rules, ones designed in advance by nonparties. These packages can help lower negotiation costs and reduce asymmetric information concerns in procedural bargaining: a defendant might smell a skunk if a prosecutor offers a complicated package of procedural adjustments out of the blue. 139

The most widely studied of these pre-fab procedural packages are known as “slow pleas of guilty” from Pittsburgh and Philadelphia state courts during the 1960s and 1970s. 140 Those cities implemented procedural systems involving waiver of the right to a jury trial but not waiver of the right to a trial before a judge—essentially a special type of bench trial, with judges who were known at the time for being particularly lenient. 141 During the hearing, the defendant was required to stipulate both to the truth and the admissibility of the police report; and though the defendant could not “contradict” the police report, they could “explain” and “supplement” it before the judge. Defense attorneys in Philadelphia reported that defendants usual-

138 The Supreme Court recently stated that “no appeal waiver serves as an absolute bar to all appellate claims” and that “even the broadest appeal waiver does not deprive a defendant of all appellate claims.” See Garza v. Idaho, 139 S. Ct. 738, 744–45, 749–50 (2019). As we argue, the threat of judicial review affects the value of settlement terms. If courts regularly undo negotiated appeal waivers, parties will adjust and find value elsewhere (even though some erstwhile surplus might remain unobtainable).

139 We might think of these arrangements as akin to expedited or summary jury trials in civil cases. Expedited jury trials are public proceedings before a judge but with circumscribed procedures such that a decision is reached rapidly, often within a single day. See generally NAT’L CTR. FOR STATE CTS., SHORT, SUMMARY & EXPEDITED: THE EVOLUTION OF CIVIL JURY TRIALS (2019), https://www.ncsc.org/_data/assets/pdf_file/0020/18083/shortsummaryexpedited-online-rev.pdf [https://perma.cc/HK5M-8VF9] (reviewing expedited programs around the United States).


141 Alschuler, supra note 3, at 725 (explaining that these bench trials provided much faster and lower risk versions of full trials while still allowing the defendant “to present his side of the story to an impartial third party . . . [and] not surrender his chance for acquittal”).
ly agreed to this package of procedure modifications because the presiding judges were more lenient than those assigned in other nonjury trials—i.e., the waivers came with a lower expected sentence and (perhaps) less uncertainty.\textsuperscript{142} Further sweetening the deal, it was purportedly understood among the parties that the district attorney’s office would allow any defendant sentenced following such a proceeding to obtain a new, full trial without procedural restrictions.\textsuperscript{143} This bundle thus served a signaling function to risk-averse defendants while also preserving the state’s resources by not providing each defendant with a “Cadillac” trial.

A more recent example of a procedure-modification package is the Prosecutorial Remedies and Other Tools to end the Exploitation of Children Today Act (“PROTECT Act”) of 2003 at the federal level. The law permits certain, heavily burdened federal judicial districts to offer “fast-track” or “early disposition” procedures to quickly resolve cases involving immigration crimes.\textsuperscript{144} Under its terms, a defendant who promptly (pre-indictment) pleads guilty to a fast-track eligible crime and agrees to waive certain pretrial and post-conviction rights receives various concessions and dramatic Guideline departure recommendations from the prosecutor.\textsuperscript{145} This early disposition program appears to be particularly valuable to defendants. In 2019, for instance, those convicted via these procedures received a sentence that was shorter by a median of 42\% and a mean of 46.5\%.\textsuperscript{146} This bundle is also attractive to the government because of the substantial burden prosecuting immigration crimes entails. In 2019, immigration crimes comprised 38.4\% of all criminal charges brought in the federal courts.\textsuperscript{147} In effect, the

\textsuperscript{142} Id.
\textsuperscript{143} Alschuler, supra note 140, at 1036 n.485.
\textsuperscript{144} Prosecutorial Remedies and Other Tools to end the Exploitation of Children Today Act of 2003 (PROTECT Act), Pub. L. No. 108-21, § 401(m), 117 Stat. 650, 675 (codified as amended at 28 U.S.C. § 994) (authorizing the promulgation of downward sentencing departures as part of an early disposition program). Note that the first fast-track programs for handling immigration crimes were implemented by U.S. Attorneys in the mid-1990s without any congressional warrant, with anecdotal evidence suggesting that as many as half of them developed some form of fast-track program. See Thomas E. Gorman, Comment, Fast-Track Sentencing Disparity: Rereading Congressional Intent to Resolve the Circuit Split, 77 U. CHI. L. REV. 479, 485–86 (2010).
\textsuperscript{145} See United States v. Medrano-Duran, 386 F. Supp. 2d 943, 944 (N.D. Ill. 2005) (“The purpose of these programs was and is to facilitate prompt and easy disposition of cases to reduce the burdens they impose in those districts . . . .”).
\textsuperscript{146} U.S. SENT’G COMM’N, supra note 102, at 99 tbl.38; see also United States v. Marcial-Santiago, 447 F.3d 715, 718–19 (9th Cir. 2006) (“In light of Congress’s explicit authorization of fast-track programs in the PROTECT Act, we cannot say that the disparity between Appellants’ sentences and the sentences imposed on similarly-situated defendants in fast-track districts is ‘unwarranted’ within the meaning of § 3553(a)(6) . . . . It is justified by the benefits gained by the government when defendants plead guilty early in criminal proceedings.” (citations omitted) (quoting 18 U.S.C. § 3553(a)(6))).
\textsuperscript{147} See U.S. SENT’G COMM’N, supra note 102, at 45 fig.2.
PROTECT Act assembled alternative procedures to induce settlement, empowering the government to manage its caseload while offering the accused valuable concessions.

Beyond these existing pre-fab bundles and a la carte swaps, there is a growing scholarly conversation that aims to identify hypothetical procedure-modification agreements through which parties might conceivably realize mutual gains. Talia Fisher suggests that prosecutors and defendants may benefit from the liberty to negotiate alternative standards of proof—changing “beyond a reasonable doubt” to “preponderance of the evidence,” for example—which would free mutually optimistic parties to proceed to trial with their dispute more finely manicured to better align with their expectations about the evidence. Similarly, Saul Levmore and Ariel Porat recommend allowing parties to dicker over the prohibition on double jeopardy, which could lower the costs of trial by reducing the incentives for prosecutors to overinvest in their one bite at the apple. And Russell Gold and his collaborators propose a fee-shifting scheme in which acquitted defendants may obtain fees from the prosecutor’s office in order to alleviate the economic burden of a full-throated defense. In theory, criminal procedures are infinitely malleable, allowing parties to adjust their risks and interests with precision—so long as the law does not prohibit the agreement and the judge approves (or is unaware) of the terms.

More than their legal permissibility, however, the plausibility of these theoretical modifications turns on prosecutors and defendants each finding benefits sufficient to justify their respective concessions. Not every theoretical bargain is likely to offer sufficient gains to justify a real-world counterpart. And criminal settlement can be complicated by the lumpiness of certain settlement terms. For example, the dichotomous nature of double jeopardy (a defendant cannot waive fifty percent of the right) implies that a prosecutor would need to offer a sufficiently attractive term or set of terms to induce complete waiver. But in the Supreme Court’s words, “A defendant can ‘maximize’ what he has to ‘sell’ only if he is permitted to offer what the prosecutor is most interested in buying.” And, as Nancy King reminds us,

---

149 Fisher, supra note 30, at 945.
150 Levmore & Porat, supra note 85, at 282–92.
151 Gold et al., supra note 30, at 1655.
152 See, e.g., King, supra note 30, at 124 (arguing that prosecutors are unlikely to find a standard of review alteration to be a valuable concession).
“[T]he market for some rights may yet emerge . . . .” Unbridling parties’ creative bargaining power may nudge them to experiment with procedural structures involving rights beyond a defendant’s right to a trial.

B. Issue-Modification Agreements

An issue-modification agreement is one in which the parties agree to change the underlying substance of their dispute. Rather than amending the governing procedures or altering the sanctions associated with any particular verdict or plea, a prosecutor and a defendant mutually agree to change the issues that need to be resolved in the belief that doing so will advance their respective interests. The literature has traditionally referred to negotiating over such settlement terms in the criminal context as “charge bargaining” because parties often agree to an exchange in which the prosecutor promises to amend the charges facing the defendant. Yet issue modifications can stretch far beyond modifying the alleged crimes. They can include reconstituting the specific elements of crimes, reframing relevant facts, and agreeing on the factors for the judge to consider in determining and imposing a sentence. Critically, these reconstructions can be entirely detached from reality and even the initial charging document. As Thea Johnson observes, “[E]verything has become a bargaining chip” when parties attempt to resolve criminal disputes, even “truth itself.”

The issues in dispute between a prosecutor and a defendant—and over which they can bargain—may be discrete or divisible. That is, the parties may agree that certain charges are no longer in dispute at all, or they may settle on the defendant stipulating to certain elements of a single offense (or to something wanting in an available affirmative defense). They may also settle on a particular value or range of values (e.g., of stolen goods) or, alternatively, agree that certain values are off the table. Defendants can at

---

154 King, supra note 30, at 124.
155 Prescott & Spier, supra note 4, at 98.
156 See supra note 24 and accompanying text.
157 See Johnson, supra note 28, at 891.
158 Id.
159 See King, supra note 30, at 124 (“[T]he defendant perceives a benefit from agreeing to accept as proven certain elements of a criminal offense, and the prosecutor values not having to prove those elements at trial.”). Or consider a case in which a defendant stipulates to two (but not all) elements of a charge—that the property was not the defendant’s and that it was worth over $1,000—in exchange for the prosecutor agreeing to dismiss twenty-one other charges. See State v. Jaroma, 660 A.2d 1131, 1133 (N.H. 1995).
160 Stephen J. Schulhofer and Ilene H. Nagel note, importantly, “that circumvention is more likely to occur in . . . [contexts] in which the Guidelines prescribe sentences that are anchored to mandatory minimums, than in many offense categories . . . in which the Guidelines’ prescribed
times benefit from unilaterally stipulating to elements or incriminating facts. For instance, doing so can reduce a prosecutor’s incentive to offer evidence that could influence a jury’s perception of a secondary issue in the case, known as the “‘halo’ effect.” In a production-of-child-pornography case, a defendant might agree to stipulate that a depicted individual is under eighteen years of age (reducing the expense, effort, and uncertainty for the prosecutor of attempting to prove the same) while still arguing that it was reasonable to be deceived as to that fact. Or the parties might conclude that litigating certain issues is too costly given the likely outcomes on those questions, even if the case as a whole remains viable. For example, a defendant might stipulate that a firearm moved in interstate commerce (a fact that is often readily established by the location of the gun’s manufacturer) but challenge, say, whether they had possession of the weapon at an otherwise full-blown trial. Indeed, parties can and do settle the parts of a case that make sense for them to settle while continuing to litigate the remaining issues in order to make the most of their respective positions.

It is worth emphasizing the potentially vital risk-reduction benefits that can accrue from settling even just some issues. Overcriminalization means that creative prosecutors can craft complex indictments, picking and choosing which crimes to bring (and under what conditions) to amplify a defend-

---

161 See Edith Greene & Elizabeth F. Loftus, When Crimes Are Joined at Trial, 9 LAW & HUM. BEHAV. 193, 194 (1985) (explaining a juror’s tendency to view a defendant in light of all charges the prosecutor brings); Andrew D. Leipold & Hossein A. Abbasi, The Impact of Joinder and Severance on Federal Criminal Cases: An Empirical Study, 59 Vand. L. Rev. 349, 355 (2006) (“[T]he more counts in the indictment, the quicker the jury may be to assume that the accused must be guilty of something.”).

162 Anne E. Di Salvo, Note, United States v. Malloy: Unreasonably Denying Criminal Defendants a Reasonable Mistake of Age Defense in the Fourth Circuit, 69 Md. L. Rev. 1020, 1031–32 (2010) (considering the legal viability of this defense and reviewing United States v. U.S. District Court for the Central District of California, 858 F.2d 534, 536 (9th Cir. 1988), in which the defendants “stipulated to the fact that the girl appearing in their film was a minor, [but] argued that they were greatly deceived as to her true age”).

163 See, e.g., Alschuler, supra note 140, at 944 (discussing the “process costs” that a defendant incurs, such as attorney’s fees and other costs, by contesting the charges against them); Note, Costs and the Plea Bargaining Process: Reducing the Price of Justice to the Nonindigent Defendant, 89 Yale L.J. 333, 333 (1979) (“The present system of cost allocation . . . creates a strong incentive for nonindigent defendants to plead guilty.”).

164 Such agreements are not uncommon. See, e.g., United States v. Simplice, 687 F. App’x 850, 851–53 (11th Cir. 2017) (per curiam) (defendant stipulated that the firearm at issue had traveled across state lines but challenged at a jury trial “whether he had actual possession of [it]”); United States v. Mosley, 339 F. App’x 568, 571 (6th Cir. 2009) (defendant stipulated both to the fact that he had a previous felony conviction and that the firearm had traveled in interstate commerce but contested at a bench trial whether he was in possession of the firearm).
ant’s exposure, making defenses more costly and outcomes more unpredictable. Imagine how a simple drug-possession case can be easily combined with charges for possession of drug paraphernalia, conspiracy to distribute, and even child endangerment. Prosecutors can use such “charge stacking” to threaten defendants with extreme sentences to pressure them into making concessions. Thus, in practical terms, there are nearly certain to be multiple issues over which a prosecutor and a defendant can negotiate and several concessions they can trade, thus making one or more exchanges more likely than if there were just a single issue in dispute. A prosecutor is able to offer to drop certain charges—such as swapping a burglary charge (a felony) for a criminal trespass charge (a misdemeanor)—if the defendant in return agrees to extend a benefit to the prosecutor, such as waiving a certain defense or cooperating in some other criminal matter. In so doing, the parties alter the foundation of their dispute, leaving in the agreement’s wake an entirely different dispute to be resolved.

Issue modifications can be so dramatic that they effectively morph an “actual” dispute into a truly fictional scenario. A defendant may be convicted of “an offense that the defendant did not commit, and that all the parties in the case know the defendant did not commit.” Although the judge is required to ensure that there is a factual basis behind any plea, “[i]n practice, [this] . . . requirement rarely hinders an effort to plead guilty.” Such settlements can be powerful tools to shape the range of likely consequences that will follow from any criminal conviction. As Thea Johnson explains, “Fictional pleas—through minor sleights of hand or outright manipulation of facts or law—avoid deportation and other collateral consequences, while allowing the prosecutor to secure a disposition in the criminal case.” Even Justice Stevens has recognized the benefits of such issue-modification

165 See Irby v. United States, 390 F.2d 432, 439 (D.C. Cir. 1967) (en banc) (Bazelon, C.J., dissenting) (“[O]ften it takes nothing more than a fertile imagination to spin several crimes out of a single transaction.”).

166 Cf. Paul J. Hofer, Federal Sentencing for Violent and Drug Trafficking Crimes Involving Firearms: Recent Changes and Prospects for Improvement, 37 AM. CRIM. L. REV. 41, 53–57 (2000) (reviewing a study showing that prosecutors did not bring or ultimately dropped federal gun charges in more than half of the cases in which gun charges that carried consecutive sentences appeared to be available and/or appropriate).


169 Gilchrist, supra note 30, at 629 n.110; see also United States v. Yeje-Cabrera, 430 F.3d 1, 27–29 (1st Cir. 2005) (explaining that although the law prohibits parties from lying or affirmatively misrepresenting evidence, they may agree to proffer certain facts and omit others).

170 Johnson, supra note 28, at 858.
agreements, encouraging parties to “plea bargain creatively” so they might avoid disproportionate punishments. Notwithstanding this blessing from on high, others describe this issue-bargaining practice pejoratively as “systemic lying” and view it as emblematic of courts abandoning their institutional role in uncovering truth.

Examples of this negotiated fictionalizing abound, and many of them establish that criminal settlements need not involve the dropping or swapping of an entire charge. Settlements can be more fine-grained and can even sidestep a guilty plea. One type of issue-modification (and perhaps also outcome-modification) agreement common in the 1980s and early 1990s in the federal courts was known as “[d]ate bargaining,” whereby the parties would agree that an act occurred prior to the enactment of the Sentencing Guidelines in order to render them inapplicable. Similarly, parties may stipulate to possession of a lesser quantity of drugs than the facts actually make plain so as to avoid harsh mandatory sentences. In fact, a defendant’s criminal history—a documented, extant set of facts—is itself often a negotiated story in which the parties agree on how to present the defendant’s criminal past to the judge. According to one supervising probation officer for the Eastern District of California after he closely studied sentencing recommendations, “Stipulations seem[] to be fictional writings, when compared with the known facts of the cases they attempted to address.” And a study of probation officers from 1996 indicates that “approximately forty percent . . . believe[d] that guideline calculations set forth in plea agreements in a majority of cases are not supported by offense facts that accurately and completely reflect all aspects of the case.”

At the most extreme, agreements to modify a dispute can transform a case into a fictional contest entirely unrelated to the original case. Con-

171 Padilla v. Kentucky, 559 U.S. 356, 373 (2010) (“Counsel . . . may be able to plea bargain creatively with the prosecutor in order to craft a conviction and sentence that reduce the likelihood of deportation, as by avoiding a conviction for an offense that automatically triggers the removal consequence.”).


174 See, e.g., id. at 272–73.


178 See, e.g., People v. Keizer, 790 N.E.2d 1149, 1152 n.2 (N.Y. 2003) (“[A] defendant may plead guilty to a crime for which there is no factual basis and even plead guilty to a hypothetical
sider the following example of this idea in action: In 2006, a defendant in Washington who was accused of participating in a violent robbery pleaded guilty to creating no less than one thousand illegal music recordings without consent. But as the prosecutor made clear in a later newspaper interview: “There were no allegations of sound recordings or videos. We were just being creative to get to the point we needed to get in sentencing.” And he acknowledged that the robbery charge would have been difficult to prove because witnesses had refused to testify. As such, by modifying the issues in contention, the parties crafted a hypothetical dispute more amenable to their joint perspective on how the original case ought to resolve, creating new surplus to share. Such fabrications are not altogether uncommon. The Supreme Court of California, for instance, has blessed factually baseless agreements, stating: “We reiterate our conviction that the plea bargain plays a vital role in our system of criminal procedure; we would be loath to reduce its usefulness by confining it within the straitjacket of ‘necessarily included offenses.’” If judges cannot (or refuse to) see the underlying truth, facts can be mere speed bumps in parties’ pursuit of their interests.

Beyond these types of factually unmoored partial settlements, some courts have also tolerated convictions resulting from guilty pleas to nonexistent crimes. For example, a New York state court sustained a defendant’s guilty plea to the logically inconsistent crime of “attempted manslaughter in the second degree” specifically on the basis that the plea “was sought by [the] defendant and freely taken as part of a bargain which was struck for the defendant’s benefit.” The U.S. Court of Appeals for the Fifth Circuit has also recognized the plea of “attempted recklessness.” Thus, with judi-

---

180 Id.
181 West, 477 P.2d at 421.
183 Dale v. Holder, 610 F.3d 294, 302–03 (5th Cir. 2010) (rejecting a Board of Immigration Appeals decision that a non-citizen could not plead guilty to “attempted recklessness”).
cial consent, parties are able to construct the law itself, making it easier to tailor settlements to their respective preferences. Once prosecutors and defendants are released from factual or legal predicates, they can craft entirely new criminal disputes that change which statutory requirements apply (if they exist at all) and unlock punitive resolutions that were previously unavailable. By concocting their own truth, parties can contractually innovate to create value they can split.

Partial settlements in criminal cases—like all settlements—are driven by party preferences and resources, and issue-modification agreements reflect the same three motivating interests as procedure-modification agreements: reworking the substantive issues in dispute can cut overall adjudication costs, reduce risk burdens by cabining uncertainty, and shift the landscape of likely outcomes, even opening up previously unavailable sentencing scenarios. Such agreements can be observed in conjunction with various complementary procedural and outcome modifications, but they may also occur on their own. Issue modification alone can be a powerful partial-settlement tool, freeing mutually optimistic parties to realize the value of agreement without a more comprehensive settlement. But again, this is only so long as the parties can effectively evade the judicial gaze, or the judge otherwise agrees to play along with what may be a charade.

C. Outcome-Modification Agreements

An outcome-modification agreement is usually one in which the parties agree to restrict or alter the range of potential remedies that a court may impose.185 In the civil context, litigants can agree by contract to transform an initial litigation outcome (e.g., a damage award) into an entirely new outcome after a jury or judge announces a decision (using a function that relates raw verdicts to final awards).186 In the criminal context, however, prosecutors and defendants cannot easily modify final sentencing outcomes, at least not in the same way, because the judge’s sentencing determination is

---

185 See Prescott & Spier, supra note 4, at 85. These agreements often restrict or compress the range of final outcomes to reduce uncertainty, which can benefit risk-averse parties. Outcome agreements, however, as a general matter simply “translate” an initial outcome (the “default” or raw outcome) into a final outcome by which the parties agree to abide. Risk-loving and mutually optimistic parties might in theory enter into agreements in which both “bet” they will win, thus increasing outcome variation. Consider these terms: “If the judge does not dismiss charge X, the defendant agrees to accept the maximum sentence on all remaining charges. If the judge dismiss charge X, the prosecutor agrees to drop all other charges as well.”

186 One classic example is a “high-low agreement,” in which a “defendant agrees to pay the plaintiff a minimum recovery in return for the plaintiff’s agreement to accept a maximum amount regardless of the outcome of the trial.” See High-Low Agreement, BLACK’S LAW DICTIONARY (11th ed. 2019).
more or less final. Accordingly, criminal outcome modifications often see a prosecutor and a defendant agreeing to bind or nudge the judge toward a negotiated outcome by recommending a maximum and minimum sentence of incarceration. These agreements are often, though again not necessarily, accompanied by procedure- and issue-modification agreements as well, either as compensating concessions for the sentencing recommendation or because the promise of an outcome modification increases, or at least does not eliminate, the value of these other exchanges.

Again, Federal Rule of Criminal Procedure 11 outlines the most obvious outcome-modification options available in federal court. “Type B” agreements involve the prosecutor recommending a particular sentence or sentencing range, while “Type C” agreements involve the prosecutor forcing a judge to impose (or reject) a specific sentence or one within a specific sentencing range. A judge must accept or reject Type C agreements in their entirety. If the judge does not agree with the parties’ proposal, the judge is not permitted to enter a different sentence, and the defendant may withdraw the guilty plea. If the judge rejects the proposal, the prosecutor can abandon the charges, the parties can propose an alternative recommendation, or they can proceed to trial. Type B agreements, on the other hand, do not bind the judge to an outcome but act as a nudge. Judges must consider such recommendations, but they are free to impose any reasonable sentence, within certain statutory and constitutional constraints. Defendants may not withdraw a guilty plea if the imposed sentence departs from the nonbinding recommendation. The defendant received the benefit of the bargain; after all, the prosecutor did make the recommendation.

The parties who come to agreement must choose whether a binding or nonbinding proposal is more or less valuable to them given what might follow. We can expect that if more than one sentence or an entire sentencing

187 We can envision agreements that are somewhat close, however. For example, imagine an agreement between the parties in which the prosecutor promises not to contest an appeal of the sentence (and perhaps is willing to grant other procedural claims) if the judge imposes a sentence above a set limit. Such an agreement might reduce the risk associated with judicial discretion at the final sentencing stage, and it might even cause the judge to avoid imposing a sentence that the court knows the prosecutor would not defend. See discussion infra Part III.

188 For instance, a defendant’s agreement to modify the procedure for adjudication on the merits (by waiving their right to a jury trial) may be accompanied by an outcome modification in which the prosecutor agrees to recommend a lesser sentence. By combining these modifications, the value of the entire package increases for both the prosecutor and the defendant, making it more likely that parties will reach a mutually beneficial accord.

189 FED. R. CRIM. P. 11.

190 Id. R. 11(c)(1)(A)–(C).

191 Id. R. 11(c)(1)(C).

192 Id. R. 11(c)(1)(B).
range (without a trial) will improve matters for the parties relative to a trial with a less predictable outcome or where certain unattractive outcomes are relatively more likely with no agreement, they will negotiate over whether to select a sentencing recommendation or instead a binding sentence or sentencing range.\textsuperscript{193} Thus, if a defendant believes that the judge is likely to impose a sentence in line with the prosecutor’s proposal, whatever it is, it would not be worthwhile to offer large concessions to the prosecutor to obtain a Type C agreement (more binding) over a Type B agreement (less binding). Defendants may also value the opportunity to argue for leniency at sentencing, particularly if they are confident that their circumstances will make them appear more deserving.\textsuperscript{194} Alternatively, if defendants consider the risk of a less predictable sentence to be too high, they may be more inclined to pursue the comfortable certainty of a Type C bargain even though such a choice may result in a longer expected sentence because, for example, they will no longer have an opportunity to convince the judge to choose a sentence below the prosecutor’s recommendation.

The decision of whether to agree to a binding or nonbinding proposal (or to agree at all), then, is a function of the judge’s anticipated reaction to any specific sentencing recommendation—and to such agreements generally—as well as the parties’ level and dislike of uncertainty surrounding those reactions. Consequently, parties will attempt to anticipate how their judge will react to the underlying facts and legal aspects of their case when deciding whether to enter into an outcome-modification agreement at all. Some judges, for instance, are adverse toward Type C agreements because they see these bargains as usurping judicial authority as well as being disruptive, at least in those instances in which the judge must reject the proposal and reschedule the hearing.\textsuperscript{195} Parties can account for these expectations and uncertainties when shaping settlement terms. Moreover, parties have strong incentives to identify and incorporate terms that transmit information supporting the proposal and signal reasonableness to the judge.\textsuperscript{196}

\textsuperscript{193} Sigman, supra note 36, at 1332.


\textsuperscript{195} See, e.g., United States v. Seidman, 483 F. Supp. 156, 158 (E.D. Wis.) (refusing Type C agreements because “[i]t is this Court’s prerogative to determine the type of sentence that should be imposed upon a defendant for the offense of which he or she has been adjudged guilty”), aff’d, 636 F.2d 1222 (7th Cir. 1980); Stephen J. Schulhofer, \textit{Due Process of Sentencing}, 128 U. PA. L. REV. 733, 745 n.55 (1980) (suggesting that judges disfavor Type C agreements because if rejected the case goes back onto the trial calendar, disrupting the docket).

\textsuperscript{196} An apt comparison is final-offer arbitration, which is commonly used in Major League Baseball. Final-offer arbitration introduces uncertainty into the negotiations in that the arbiter is constrained to choose one of the disputant’s final offers as the binding settlement. The parties’
Whether as part of a Type B or Type C agreement, parties can decide to recommend to the judge a specific sentence, a sentencing range, or a sentencing cap.\textsuperscript{197} The fact that prosecutors and defendants can propose a sentencing range or cap (versus an exact sentence) renders it more likely that parties will agree on terms—more wiggle room allows both parties to indulge their divergent beliefs and effectively expands the bargaining zone.\textsuperscript{198} Partial settlements of this sort cannot eliminate all uncertainty, and therefore risk-averse parties continue to endure risk-bearing costs. But attempting to achieve greater certainty by continuing negotiations may be so pricey (in both time and money as well as ongoing uncertainty) that both parties may prefer to just agree to outsource the determination of the final sentence to the judge (constrained, of course, to be within the agreement’s limiting parameters).\textsuperscript{199} In fact, there are sure to be many scenarios in which an agreement between the parties is only feasible in terms of a sentencing range.\textsuperscript{200} That is, if the parties have access to different information or hold particularly divergent assessments of the defendant’s chance of success during or after trial, settlement on an exact sentence may be impossible even as the parties are able to reach agreement on a sentencing range. Indeed, optimistic, risk-neutral parties may even prefer to have the minimum and maximum sentences in the range as far apart as possible.\textsuperscript{201} Regardless, given party preferences, beliefs, and resources, agreement to a sentencing range or cap may be an optimal partial settlement.\textsuperscript{202}

Parties can also enter agreements that have the effect of expanding the array of potential outcomes to include sentences that were previously una-

\textsuperscript{197} See FED. R. CRIM. P. 11.

\textsuperscript{198} The situation is highly comparable to outcome modifications in the civil context. For example, a high-low agreement (in which the parties pick a maximum and minimum damages award) allows defendants to “believe” that they will probably pay the “low” when deciding to settle and allows plaintiffs to “believe” that they will probably pay the “high.” Both cannot be right, but overoptimism, divergent priors, or asymmetric information is sufficient to persuade both parties to accept the arrangement. See Kathryn E. Spier & J.J. Prescott, Contracting on Litigation, 50 RAND J. ECON. 391, 410 (2019) (discussing these factors in driving contingent settlement contracts in the civil context).

\textsuperscript{199} For a discussion on how transaction costs impact settlements, see generally George L. Priest & Benjamin Klein, The Selection of Disputes for Litigation, 13 J. LEGAL STUD. 1 (1984).


\textsuperscript{201} See id. at 705 n.19 (noting this in the context of civil high-low agreements).

\textsuperscript{202} Again, consider high-low agreements. There, the plaintiff agrees to a cap on potential damages in exchange for a guaranteed minimum return. High-low agreements can be valuable to parties too mutually optimistic about their chances at trial to find fully settling attractive. See Prescott & Spier, supra note 4, at 63.
Available—for instance, by agreeing to trigger certain “safety valves.” In the federal system, “safety valves” remove a judge’s obligation to impose a sentence that complies with statutory or Sentencing Guidelines’ minimum requirements. There are two main types of safety valves when federal judges sentence defendants. The first involves the prosecutor consenting to stipulate that the defendant has agreed to cooperate with the government and has provided “substantial assistance.” The other involves a defendant taking responsibility or expressing remorse for their crimes. The judge ultimately determines whether a defendant has provided sufficient assistance or taken sufficient responsibility, but the prosecutor can agree to offer corroborating input to improve the defendant’s chance of success. Some judges claim that taking responsibility for one’s transgressions provides valuable information about the likelihood of the defendant’s rehabilitation and thus future dangerousness. For prosecutors and defendants centered on practicalities, however, these safety valves function chiefly to loosen and expand the boundaries of available outcome options, which the parties then navigate to advance their respective interests.

The above examples all consist of modifications concerning the potential length of a defendant’s sentence, but outcome modifications can stretch beyond simply the length of incarceration. As part of a settlement, parties may construct and agree to outcomes regarding probation, community service obligations, facility placement, drug rehabilitation opportunities, restitution, asset forfeiture, and the like. A large difference in party valuation means more gains from trade, and many of these outcomes matter a great deal to some defendants but are often of little value to prosecutors, at least those who are not solely interested in maximizing the harm sentences im-

204 U.S. Sent’g Guidelines Manual § 3E1.1.
205 See Brady v. United States, 397 U.S. 742, 753 (1970) (claiming that a defendant’s admission of guilt demonstrates “hope for success in rehabilitation over a shorter period of time than might otherwise be necessary”).
206 One interesting set of examples are “release-dismissal” agreements, which usually involve a prosecutor agreeing to dismiss criminal charges against a defendant if the defendant agrees to release any civil claims they might have against the state. See, e.g., Town of Newton v. Rumery, 480 U.S. 386, 393–94 (1987). Such agreements are particularly common in cases involving resisting arrest and disorderly conduct scenarios in which the accused claims to have suffered harm; they can provide gains on all sides by lowering the exposure experienced by both the defendant and the state. See Andrew B. Coan, The Legal Ethics of Release-Dismissal Agreements: Theory and Practice, 1 Stan. J. C.R. & C.L. 371, 387 (2005).
pose on defendants. Most prosecutors, we suspect, do not chalk up preventing a defendant’s incarceration in a facility nearer to the defendant’s family as a win.207 Likewise, in some circumstances, offering a defendant the prospect of serving their term of incarceration at a “cushy” minimum security prison—derisively referred to as “Club Fed” in the federal prison system—may be sufficient to seal a deal.208 A defendant may also be willing to spend more time incarcerated so long as, say, conjugal visits are permitted.209 Despite the risk that a judge may fail to implement these terms (or may not be made aware of them), such partial settlements create value by shifting the likely outcome and reducing uncertainty surrounding it.

Moreover, outcome-modification agreements can go beyond a defendant and directly affect third parties, which occurs only because the defendant has an interest in these parties. For instance, hardball prosecutors may threaten to investigate and indict a defendant’s friends, family members, or co-conspirators unless the defendant relinquishes some or all of their procedural rights.210 Under such circumstances, defendants might feel compelled to settle to protect their network from the costs of such attention or to save themselves from any reputational effects.211 As another example, consider

---

207 In significant part, this is because these incidental aspects of punishment are unlikely to matter to the career concerns of a prosecutor. See supra note 63 and accompanying text.


209 See Joseph A. Colquitt, Ad Hoc Plea Bargaining, 75 TUL. L. REV. 695, 712 & n.100 (2001) (noting this as an “unauthorized benefit” and citing State v. Horning, 761 P.2d 728, 732–33 (Ariz. Ct. App. 1988), as “reviewing a plea agreement that included a stipulation by the State that it would take no position on the issue of whether the defendant would be allowed to have conjugal visits with his wife in the county jail”).

210 See, e.g., United States v. Yong, 926 F.3d 582, 591 (9th Cir. 2019) (discussing a settlement in which a defendant pleaded guilty as part of a package deal in which his son would not be prosecuted if the defendant and all other co-defendants pleaded guilty); United States v. McElhaney, 469 F.3d 382, 385–86 (5th Cir. 2006) (permitting a defendant to plead guilty in exchange for immunity for him and his family against tax-related prosecution); United States v. Abbott, 241 F.3d 29, 32 (1st Cir. 2001) (allowing a defendant to plead guilty in exchange for the prosecutor recommending no prison time for the defendant’s mother).

211 See United States v. Lopez, 944 F.2d 33, 36 (1st Cir. 1991) (“A plea agreement entailing lenity to a third party ‘imposes a special responsibility on the district court to ascertain [the] plea’s voluntariness due to its coercive potential.’” (alteration in original) (citations omitted) (quoting United States v. Buckley, 847 F.2d 991, 1000 n.6 (1st Cir. 1988))). In addition, it is important to note that it may not be possible to enforce these types of agreements. See id. at 37 (observing without deciding that “even if third party beneficiary principles [from contract law] were applicable to a plea agreement in a criminal case, . . . we are unaware of authority to that effect”). But see United States v. Tursi, 576 F.2d 396, 396–97 (1st Cir. 1978) (rejecting an appeal from an incarcerated father who had pleaded guilty on the condition that his son be recommended for probation—but who was ultimately sentenced to jail in spite of the recommendation—on the grounds that the prosecutor had fulfilled the agreement by in fact recommending probation).
that a prosecutor and a defendant can bargain over the type or amount of restitution that must be paid to any victims. In this instance, the prosecutor may agree to drop certain charges if the defendant agrees to pay for the victim’s physical or psychological treatment.\footnote{See People v. Johns, No. F047499, 2006 WL 798454, at *3 (Cal. Ct. App. Mar. 29, 2006) (outlining the parameters of such a bargain).} Or the parties may agree that the defendant will surrender all profits from a book or movie deal describing the offense, on the understanding that these are ill-gotten gains.\footnote{These restrictions were originally legislative in nature, known as “Son-of-Sam” laws, until the Supreme Court ruled them unconstitutional. Nevertheless, parties are free to construct such agreements privately. See Colquitt, supra note 211, at 735.} Outcome modifications implicating third parties can be powerful negotiation tools to encourage defendants and prosecutors to reach agreement.

There are also peculiar outcome modifications that involve defendants agreeing to take certain actions entirely outside of the criminal justice system. During World War II and the Vietnam War (and probably all previous armed conflicts of any duration), for instance, defendants would sometimes agree to join the military rather than serve jail or prison time.\footnote{See Carissa Byrne Hessick, Why Are Only Bad Acts Good Sentencing Factors?, 88 B.U. L. REV. 1109, 1154 (2008). Such agreements appear to be less common today, but parties still do enter into settlements involving military service. See State v. Hamrick, 595 N.W.2d 492, 493 (Iowa 1999) (en banc) (discussing a settlement in which the defendant agreed to join the army within five months of the state’s dismissal of his charges).} And when the military put restrictions on its willingness to accept “jailbirds” into service, parties crafted (and judges approved) deals meant to circumvent those restrictions, such as issue-modification agreements that reframed the specific charges facing the defendant, so that military service would remain a viable possibility down the road.\footnote{Hessick, supra note 214, at 1154 n.235 (citing Rod Powers, Can a Judge Order Someone to Join the Military or Go to Jail?, BALANCE CAREERS, https://www.thebalancecareers.com/join-the-military-or-go-to-jail-3354033 [https://perma.cc/G69G-WPQY] (Nov. 25, 2019)).}

Troublingly, defendants may trade even aspects of their bodily integrity for prosecutorial benefits. For instance, in 2005 in Fulton County, Georgia, a mother suffering post-partum depression allegedly shook her five-week-old daughter to death. The prosecutor originally charged the defendant with murder, but he later agreed to reduce the charge to manslaughter if the mother would agree to plead guilty and undergo a sterilization procedure. The prosecutor later stated that he assented to the settlement in part because of evidentiary weaknesses in the government’s case.\footnote{Cf. Beth Warren, Mother Chooses Sterilization Over Murder Trial, ATLANTA J.-CONST., Feb. 10, 2005, at A1. For an example involving a similar settlement that was ultimately rejected by a court, see Bruno v. State, 837 So. 2d 521, 522 (Fla. Dist. Ct. App. 2003) (rejecting a settlement involving a defendant charged with lewd and lascivious assault agreeing to undergo surgical castration because no statute authorized punitive castration for a conviction on those charges).} Importantly, we do
not offer these examples of unusual bargains to suggest that they are socially
valuable or morally acceptable but instead to underscore the notion that par-
ties may at least contemplate exchanging whatever they have available that
is valuable to the other side if such a trade can improve their position,
broadly construed.\footnote{This is a discussion of the United States’ horrendous history of sterilization, see generally Vanessa Volz, Note, A Matter of Choice: Women with Disabilities, Sterilization, and Reproductive Autonomy in the Twenty-First Century, 27 WOMEN’S RTS. L. REP. 203 (2006).}

Aside from the actual outcome-modification agreements that we docu-
ment above, hypothetical agreements are easy to imagine. We can deduce,
for instance, that prosecutors and defendants may structure agreements in
ways that directly incorporate the content of a judge’s decision—i.e., con-
tingent arrangements. The parties might agree that if the judge imposes a
sentence at the high end of a guideline or agreed-upon sentencing range,
then the prosecutor will make a recommendation (or not oppose a request)
regarding some of the defendant’s other interests—perhaps regarding charg-
es against friends or family members, treatment by immigration authorities,
or placement in a nearby facility.\footnote{Criminal settlement agreements contingent on third-party actions are not uncommon. Consider for instance agreements in which prosecutors agree to drop charges against a defendant if they offer testimony leading to a conviction of a co-conspirator. Parties could easily modify such an agreement to take into account not merely the fact of conviction but also the specific sentence that the judge imposes. Parties are not necessarily powerless to respond to judicial decisions, and they may find it valuable (when possible) to build backstops into their agreements.}

Or parties might structure settlements
that “operate” during the sentencing colloquy, hydraulically linking party
actions to incremental choices made by a judge. To illustrate this idea, and
assuming that the judge in question addresses one count or sentencing issue
at a time (or can be convinced to do so), the parties might construct an
agreement with the logic of a flowchart such that “if the judge finds A and
sentences to X” then “the prosecutor does B and recommends Y.” The par-
ties could fine-tune their outcome modifications in real time, reacting on the
fly to judicial decisions as they are handed down, perhaps to dampen varia-
tion and to reduce the likelihood of extreme outcomes.\footnote{The expense of constructing such a detailed and comprehensive agreement ex ante may render it unrealistic in most circumstances. Yet the parties’ interests in cabining judicial discre-
may encourage or discourage such parsing, depending on their own preferences. Innovative outcome-modification terms, just like creative procedure- and issue-modification terms, may expand the bargaining zone to allow settlements beyond what most view today as feasible.

* * *

Criminal and civil settlement are different. Judicially enforced statutory and constitutional limitations pervade criminal adjudication, often leaving the parties with noticeably less latitude and control when they bargain. But just like civil litigants, prosecutors and defendants can mutually benefit by agreeing to modify procedures, issues, and potential outcomes, exchanging valuable concessions in line with their preferences. By rebuilding their dispute to minimize costs, mitigate risk, and maximize returns, the parties can improve on default arrangements through partial settlement. Yet settlement agreements in the criminal domain are implicitly built around judges and judicial decision-making—unlike in civil cases, where judicial involvement is often optional. Judges always retain some discretion in determining final outcomes in a criminal case. This unavoidable kernel of judicial discretion necessarily shapes the types of agreements parties are likely to reach. Consequently, it is important to explicitly incorporate into our analysis of criminal settlement the relationship between, on the one side, the prosecutor and the defendant who collaborate on a deal, and on the other, the judge who decides the outcome in light of that deal.

III. MODELING JUDICIAL REVIEW OF CRIMINAL SETTLEMENT

The unavoidable discretion of judges and the power of courts as gatekeepers to state-imposed punishment distinguish settlement in the criminal domain from traditional civil settlement. To explore how these distinct features affect the dynamics of partial settlement behavior and outcomes, we analyze how judges can both respond to and explicitly shape the bargains of parties that come before them. Most plea-bargaining scholarship—beyond

---

See Shai Danziger et al., Extraneous Factors in Judicial Decisions, 108 PROC. NAT’L ACAD. SCI. U.S. AM. 6889, 6889 (2011) (presenting findings—subsequently the subject of considerable criticism from other scholars on data, methodological, and interpretation grounds—indicating that the probability of a judge granting a prisoner’s parole request is markedly higher at the beginning of the day and just after scheduled food breaks).
overly stressing guilty-plea agreements to the point of ignoring many other potential criminal settlements—has been quick to dismiss the judge’s role, emphasizing the centrality of the prosecutor instead.220 Our model brings the judge back into frame and places judicial discretion very near the center of criminal settlement dynamics. We argue that prosecutors and defendants use procedure-, issue-, and outcome-modification agreements to recast their disputes in ways that optimally constrain or influence the exercise of judicial discretion to better satisfy their respective preferences.

To make our points concrete, we focus on judicial discretion with respect to determining a defendant’s post-conviction sentence, studying how settlement agreements can influence this type of outcome by shifting as many as five separate constraint points, binding or nudging a judge toward the parties’ desired outcome, often in ways that also work to reduce uncertainty and lower adjudication costs. Our conclusions apply just as well to the use of judicial discretion at other stages of criminal prosecution, but the precise structure of federal sentencing renders it a useful platform on which to illustrate key mechanisms and moving parts.221 As we show in this Part, the various terms parties exchange may operate as complements to or substitutes for one another as defendants and prosecutors seek to reformulate their dispute, always with the aim of improving the efficiency, certainty, and consequences of criminal adjudication in jointly optimal ways.

Placing judges at the center of our analysis also allows us to highlight the potentially active role judges may play in shaping partial criminal settlements. Although judicial review may be minimal with respect to certain procedural- or issue-modification terms, it is hard to gainsay the fact that a judge could reject virtually any agreement (of which the court is aware) that is plausibly at odds with the “interests of justice.”222 Parties have robust incentives to erect agreements that they believe will survive review.223 But just

---

220 See, e.g., Standen, supra note 17, at 1477.
221 Also, because parties are especially interested in this final exercise of judicial discretion (and because it comes very close in time to the conclusion of a criminal adjudication), it is clearly salient and surely preoccupies parties in most partial-settlement negotiations.
222 This sort of active judicial management is presumably also possible in the civil context and may exist in some dark corners, but there can be little doubt that judicial regulation of procedures, issues, and outcomes is more significant in criminal cases on account of the more highly regulated environment in which the agreements take shape.
223 This is a simplifying assumption, but it is ultimately not critical to our argument, which is that parties will consciously design settlement agreements to influence judicial decision-making or, when necessary, anticipate and circumvent it. Conceivably, parties may agree to try to influence a judge by recommending sentences or other adjustments that the judge is sure to reject. One could imagine parties trying to leverage some sort of cognitive bias, such as framing, availability, or compromise, in how the judge rejects the agreement or makes other decisions in the aftermath.
as significant, judges can indirectly make terms that courts traditionally do not review more or less attractive to parties by allowing, refusing, or adjusting other terms that regularly do receive more searching scrutiny. This underappreciated indirect reach of judicial influence follows from the fact that terms from the same or different categories can operate as substitutes or complements. Consequently, parties are wise to keep in mind the potentially far-reaching consequences of later judicial decisions when negotiating over procedural or substantive issues early in litigation.

Judicial discretion (even if exercised only during sentencing) implicitly affords judges broad power to shape the contours of criminal settlements (and thus the entire litigation)—so long as parties have at least some understanding ex ante of how judges are likely to respond to the negotiated terms. Not surprisingly, judges can and do act strategically to further develop and deploy this power by signaling which types of settlements they are more or less likely to tolerate—or by otherwise revealing how they will respond to certain kinds of proposals. Judges can accomplish this by communicating their preferences directly and by signaling through their actions in the case and through their reputations established in their work on other cases. By how they use power over attempts by prosecutors and defendants to optimize via agreement, judges can make certain types of settlements more or less likely. Judges’ gatekeeper status and discretion assure them a central role in shaping criminal settlement dynamics and outcomes.

A. Constraining Judicial Discretion

Criminal settlement has at its foundation the idea that defendants and prosecutors build agreements collaboratively to bypass or channel judicial discretion in ways that align with their preferences. With respect to federal sentencing (our case study), we can model that activity by organizing limits on judicial discretion as five loci of control: (1) the crime charged, (2) the federal Sentencing Guidelines, (3) the applicability of statutory mandatory minimums, (4) the terms of parties’ criminal settlements, and (5) an overarching reasonableness requirement. By agreeing to procedure, issue, and outcome modifications that loosen or tighten these five constraints, parties can make it easier or more difficult for the judge to choose a particular sentence. Sources of constraint can be inflexible (binding a judge to a subset of
available outcomes) or flexible (nudging a judge by making it more difficult or costly to make certain choices). Critically, parties often employ settlement terms in synergistic combinations to create comprehensive settlements that better achieve their various interests.

Before we detail this line of thinking in more general terms, consider first how each of the five loci of constraint affect the scope of judicial discretion. The charges a prosecutor brings have the greatest influence—they restrict what a judge can do by setting inflexible bounds on sentence severity. Criminal charge bargaining has consumed a great deal of attention precisely because of the charge’s central role in defining the relevant ambit of possible punitive outcomes. In criminal law, the criminality of an act and its punishment are tightly linked—acts with higher degrees of criminality are generally met with more severe sentences. Thus, the charge filed effectively defines the universe of punishments available and shapes the statutory maximum and minimum sentences, the range of plausibly relevant Guideline factors, and the judge’s determination of what would constitute a reasonable sentence in light of those factors.

Second, sentencing guidelines in general function (or are intended to function) to constrain judicial discretion, at least relative to a world without guidelines. Even though the federal Sentencing Guidelines today are merely “advisory,” they continue to exert influence—on average more than fifty percent of federal sentences still fall within the relevant guideline range. The Guidelines require that judges “consult” them prior to issuing any sentence, and the figures they contain offer anchoring points for judicial decision-making. Moreover, on appeal, sentences falling within the range sanctioned by the Guidelines typically garner a presumption of reasonableness. To the extent that judges care about a court later overturning a sentence or their need to revisit a case, a sentence that is presumed reasonable is more valuable than one that is not. Because of the enduring influence of the Guidelines, prosecutors and defendants bargain over inputs to Guideline

---


227 See U.S. SENT’G COMM’N, supra note 102, at 11, 86 fig.9 (showing that, on average in years 2010–2019, 50.65% of all federal criminal sentences fell within the range recommended by the Sentencing Guidelines).

228 Booker, 543 U.S. at 264.


calculations. Facts, such as a defendant’s “role in the offense” can heavily influence sentencing determinations.\textsuperscript{231} So, too, can a finding that a defendant “clearly demonstrates acceptance of responsibility for his offense.”\textsuperscript{232} By reaching agreement on these factors, parties can make it easier for a judge to issue a mutually agreeable sentence.

Third, a judge’s tools to overcome mandatory statutory limits are few and far between. Unlike the federal Sentencing Guidelines and their advisory sentencing recommendations, the law requires judges to issue sentences within the statutory bounds set by Congress. Issue modifications can shift which limits apply, but judges enjoy no authority to issue a sentence falling outside of the relevant statutory constraints. There are, however, two ways for parties to lever open these constraints. The first allows first-time offenders who meet the criteria set forth in 18 U.S.C. § 3553(f)—many of which are susceptible to negotiation between the parties—to receive a sentence below the statutory mandatory minimum.\textsuperscript{233} The second operates through a defendant’s “substantial assistance” under 18 U.S.C. § 3553(e).\textsuperscript{234} If an agreement between the parties can unlock sentences beyond the applicable statutory minimum, the parties can expand their bargaining range and the punishment options at the judge’s disposal. Safety valves increase the parties’ uncertainty regarding possible sentences (judges have more potential landing points available to them), but the change can be desirable if it also improves the parties’ expected outcomes or if it reduces their costs, especially if they can control the additional risk in some manner.

Fourth, any explicit agreement between parties regarding punishment also constrains judges. Recall that under federal law, parties’ sentencing recommendations can be either binding or nonbinding.\textsuperscript{235} Binding recommendations, of course, impose a greater constraint on judicial discretion than those that are nonbinding. But because judges, like all people, are subject to anchoring, scaling, and other cognitive biases, any recommendation can sway a judge to move in predictable ways.\textsuperscript{236} Knowing this, the parties can craft recommendations that communicate to the judge in a manner likely to produce a better outcome. Moreover, judges often have little incentive to

---

\textsuperscript{231} See Schulhofer & Nagel, supra note 110, at 1293 (quoting U.S. SENT’G GUIDELINES MANUAL §§ 3B.1–2 (U.S. SENT’G COMM’N 1995)) (noting the “role in the offense” provision of the Sentencing Guidelines is often central to parties’ negotiations).

\textsuperscript{232} See U.S. SENT’G GUIDELINES MANUAL § 3E1.1 cmt. background (U.S. SENT’G COMM’N 2018).

\textsuperscript{233} 18 U.S.C. § 3553(f).

\textsuperscript{234} Id. § 3553(e).

\textsuperscript{235} FED. R. CRIM. P. 11(c)(1)(B)–(C).

impose a sentence contrary to a recommendation (or to the preferred outcome implied by a recommendation) because doing so may result in an appeal and potential reversal. Complying with parties’ sentencing requests—i.e., allowing themselves to be nudged toward the parties’ preferences—permits overburdened judges to dispense with the work of fully justifying their decisions. Correspondingly, by offering this carrot, parties can prod a judge toward a desired outcome.

Finally, laid atop these constraints is the need for any criminal sentence to be “reasonable” and for the judge’s exercise of discretion in determining a sentence to involve some individualized consideration of the defendant’s case, characteristics, and experiences. In federal proceedings, judges must consider the statutory factors set forth in 18 U.S.C. § 3553(a) when devising a sentence. These sentencing factors are broad and include the circumstances of the offense and the defendant’s history, the need for the sentence to reflect the motivating purposes of punishment, the available sentences, including the Sentencing Guidelines range, and the “need to avoid unwarranted sentencing disparities.” A court need not give one factor greater weight than another, which leaves judges with significant power and flexibility to pursue their own particular ends. Parties, however, can coopt this power and influence judicial sentencing outcomes by representing (or misrepresenting) these factors, painting the defendant as more or less culpable. As Judge Robert J. Conrad and Katy L. Clements explain, “Instead of airing out dirty laundry at trial, defendants . . . [opt] for plea resolution where their lawyers have the best opportunity to divert the court’s gaze from the offense conduct to arguments for variance.”

Although a judge may impose any reasonable sentence, including one inconsistent with the parties’ presentations, departing from the recommendation creates more work for them (as they must justify their decision in writ-

---

239 Id.
241 See Gall v. United States, 552 U.S. 38, 51 (2007) (stating that appellate courts “must give due deference to the district court’s decision that the § 3553(a) factors, on a whole, justify the extent of the variance”).
242 Hon. Robert J. Conrad, Jr. & Katy L. Clements, The Vanishing Criminal Jury Trial: From Trial Judges to Sentencing Judges, 86 GEO. WASH. L. REV. 99, 135 (2018); see also King, supra note 19, at 295 (“[P]arties can minimize the impact of the presentence report by stipulating in their plea agreement to facts or to applications of factors, hoping the judge will accept their stipulations rather than take the time to adjudicate the accuracy of those facts or issues.”).
ing) and also adds risk (by opening the door to reversal on appeal). The power of any particular recommendation to force a judge’s hand is at its zenith when both parties are in agreement, offering them not only the benefit of a unified front on the appropriate sentence under the circumstances but also the ability to coordinate (or collude) on the presentation of relevant facts and arguments to enhance their persuasive power. For this reason, settlement terms that specify how the parties will make a joint recommendation (as well as other factual stipulations) are unsurprising. Where there is bargaining surplus to be had, working together at sentencing can increase certainty and shift the expected outcome, generating gains for parties to share through the exchange of other terms.

Figure 1 below illustrates the scope of judicial sentencing discretion in a partial criminal settlement framework. The X-axis measures the length of sentences that a judge may consider, and the Y-axis indicates the intensity of the constraint (i.e., difficulty or cost) a judge faces in imposing the sentence in question (given the circumstances of the case, which we hold fixed in this illustration). As a judge considers whether to issue a sentence that is in-

---

243 See United States v. Simpson, 430 F.3d 1177, 1187 (D.C. Cir. 2005) (“When a defendant has not asserted the import of a particular § 3553(a) factor, nothing in the statute requires the court to explain sua sponte why it did not find that factor relevant to its discretionary decision. And nothing in Booker added such a requirement.”); id. at 1187 n.10 (“Something more is required if a district court imposes a sentence outside the Guidelines range. Section 3553(c)(2) provides that, if a sentence ‘is not of the kind, or is outside the range’ described by the Guidelines, the court must state ‘the specific reason for the imposition of a sentence different from that described, which reasons must also be stated with specificity in the written order of judgment and commitment.’” (quoting 18 U.S.C. § 3553(c)(2) (2000 & Supp. 2004))); United States v. Mares, 402 F.3d 511, 519 (5th Cir. 2005) (“When the judge exercises her discretion to impose a sentence within the Guidelines range and states for the record that she is doing so, little explanation is required. However, when the judge elects to give a non-Guideline sentence, she should carefully articulate the reasons she concludes that the sentence she has selected is appropriate for that defendant.” (footnote omitted)).

244 See, e.g., United States v. Lopez-Flores, 444 F.3d 1218, 1222 (10th Cir. 2006) (“When the defendant has not raised any substantial contentions concerning non-Guidelines § 3553(a) factors and the district court imposes a sentence within the Guidelines range, our post-Booker precedents do not require the court to explain on the record how the § 3553(a) factors justify the sentence.”).

245 The shape of this bowl-like function makes plain that both the law and the parties—through their recommendations, which leverage law, and any issue- or outcome-modification agreements—can constrain judicial discretion. The vertical lines represent sentencing “thresholds” at which a decision becomes easier or more difficult, depending on the direction of the change, and they capture the addition of a new source of constraint or the subtraction of a newly irrelevant source of constraint. Broken vertical lines indicate that a judge is permitted—even though the costs to the judge differ—to impose a punishment on either side of the line, whereas solid vertical lines at the statutory limits impose strict mandatory limits on judicial discretion. Dashed vertical lines demarcate the parties’ recommendation, which may or may not bind the court depending on whether it is a Type C or Type B agreement. The slopes of the function need not be same across different regions nor symmetrical on either side of the parties’ joint recommendation. It may also be more difficult for a judge to move the same distance in one direction
creasingly further away from a joint party recommendation, they experience greater limitations on their prerogative—in terms of an increase in workload, greater risk of being overturned on appeal, and any reputational sanctions from colleagues or the Bar for too often frustrating parties’ wishes. Some legal constraints on judicial discretion are more robust than others. Consequently, a judge can legitimately make the costly decision to impose a sentence that does not conform to parties’ Type B nonbinding recommendation or even one that departs from a Guidelines-recommended sentence, but the law does not permit a judge to issue a sentence outside of the statutory maximum and minimum set by Congress.246

To be clear, Figure 1 is best thought of as an example of what a particular judge might face at the sentencing stage after the parties have modified the dispute through partial settlement—i.e., anticipating how the judge is likely to exercise discretion, the prosecutor and the defendant assemble the playing field depicted in Figure 1 to maximize the likelihood of achieving their desired outcome, which is itself a function of the costs and risks associated with achieving that outcome.247 The parties have already agreed (or than in another. But the point is that these lines represent sharp shifts in the extent of judicial discretion.

246 That is, unless there is a “safety valve” or other basis for the judge to depart. See 18 U.S.C. § 3553(e); U.S. Sent’g Guidelines Manual § 5K1.1 (U.S. Sent’g Comm’n 2018).

247 Figure 1 does not depict “limitations” on judicial discretion that emerge indirectly and organically from procedure- or issue-modification agreements that alter a judge’s effective information set. For instance, a procedural change could lead to different evidence being presented or
decided not to agree) on the applicable charge, Guidelines facts, and so on, and they must now leave it to the judge to make the final move (in a game the parties have tried to rig) by imposing a specific sentence.

Implicit in the figure is the landscape over which parties can influence the exercise of judicial discretion by shaping the costs and constraints the judge must navigate when selecting a sentence. By strategically entering into partial-settlement agreements that have the effect of narrowing, broadening, or removing the vertical lines or altering the levels and slopes of the constraints, a prosecutor and a defendant can corner or liberate a judge in line with their preferences. Also implicit in the figure is the remaining power of the judge to pursue their own ends through the exercise of discretion. Judges will seek to maximize their utility in the face of these constraints.\(^\text{248}\)

Figure 1 does not mean to imply that parties cannot constrain judges any further, nor that parties always seek to maximally bind judges when settling. In any particular case, they might, in light of the judge’s signals or reputation—or lack thereof. But parties may also benefit by speculating on their disparate predictions of the judge’s future actions. For instance, if both parties believe that the judge is likely to favor the other party, maximally constraining the judge would be attractive to both sides, likely leading to a more restricted judicial environment. Or, if the prosecutor and the defendant were convinced that the judge would exercise discretion in their favor—i.e., both were “mutually optimistic”—the parties might “agree to disagree” and free the judge to have a more significant role in sentenced. They would, in all likelihood, also exchange other terms that would maximize the judge’s influence on the outcome later as well, such as agreeing to procedural changes that provide the judge with more ammunition or opportunities to shift the case in their direction.\(^\text{249}\)

recorded in a presentencing report, which could in turn make it more difficult for a judge to justify a particular sentence. Presumably, these partial settlements can also influence a judge’s preferences over particular sentences, which we might also think of as self-imposed constraints in the sense that choosing a sentence that does not “feel” right to a judge is psychologically costly.\(^\text{248}\)


\(^{249}\) We see a similar phenomenon resulting from mutual optimism in the civil context. When each party is confident it will win, full settlement is less attractive. But this does not mean that there are not pockets of value to be uncovered. To be sure, “there are good reasons to think that because full settlement is a big step . . . partial settlements have a unique potential to offer both parties significant net benefits even when full settlement does not.” Prescott & Spier, supra note 4, at 78 (footnote omitted). So too in the criminal context. By entering into agreements that leverage judicial discretion, prosecutors and defendants can craft partial settlements that make the most of their optimistic positions, at least ex ante.
Accordingly, when parties weigh the value of a particular criminal settlement against another agreement or against a “naked” trial, they incorporate expectations about judicial behavior—designing their partial settlement both with a clear-eyed recognition of a judge’s power and with the audacious aim of working together to control the scope of this power. Especially risk-averse parties may enter partial settlements that include multiple terms that have the collective effect of reducing the level of uncertainty dramatically, or they may take advantage of options that explicitly reduce risk, like agreeing to a specific sentence.\(^{250}\) The choice will depend on party preferences, any collateral effects on other relevant considerations (e.g., litigation costs), and the likely reaction of the judge. Alternatively, a risk-seeking or very optimistic party may enter agreements, or refuse to enter agreements, in ways that expand judicial discretion in the hope of besting the alternative. The point is that, by partially settling, parties can push a judge toward their preferred outcome even in the face of significant judicial autonomy. Critically, however, judges can push back.

**B. Judicial Shaping of Criminal Settlement**

At first blush, one might treat a judge’s sentencing discretion as a non-strategic source of uncertainty for parties negotiating a partial settlement.\(^{251}\) Just as civil litigants could, in theory, write a settlement contract in which damages turned on whether the Dow Jones went up or down in two weeks, one might think parties to a criminal dispute would be free to craft their partial settlements to incorporate judicial discretion and a judge’s preferences with no fear of the judge reacting in a counterproductive way (from the parties’ perspective) to the existence or terms of any settlement.\(^{252}\) But judges need not be, and most likely are not, passive audiences to parties’ agreements. Judges may have goals of their own in presiding over a dispute, and their ability to achieve those goals may depend on the settlement behavior.

\(^{250}\) Consider a Type C agreement with a specific sentence rather than a sentencing range. Hypothetically, the parties risk only that the judge will reject their plea. *But see* discussion *infra* Part III.B (examining how judges can exert influence to constrain parties from entering agreements that would fully bind their discretion).

\(^{251}\) *See, e.g.*, Albert W. Alschuler, *The Trial Judge’s Role in Plea Bargaining* (pt. 1), 76 COLUM. L. REV. 1059, 1081 (1976) (quoting one public defender as describing judicial discretion over sentencing as “a plunge from an unknown height”).

\(^{252}\) *See* James D. Miller, *Using Lotteries to Expand the Range of Litigation Settlements*, 26 J. LEGAL STUD. 69, 89 (1997) (discussing the benefits of using randomness to resolve outstanding issues and uncertainty, and noting that “[p]rosecutors could use lotteries and other devices in plea bargaining negotiations to achieve more efficient sentencing”).
of the parties. Although prosecutors and defense attorneys are very often repeat players, privileged to know judges individually and in general and to understand their decision-making, judges can deploy uncertainty strategically, and can do so case by case, to enhance or undermine parties’ negotiations and otherwise encourage or discourage certain types of settlements. In the lyrical words of Albert W. Alschuler, “[I]n the minuet of plea negotiation, truly delicate movements by a trial judge may have a meaning of their own.” In the discussion that follows, we consider how judges may exert counter-pressure on prosecutors and defendants in order to avoid or resist any unwanted attempt to cabin the court’s discretion.

One reason judicial power in criminal settlements is often overlooked is because, at least at the federal level, judges are prohibited from directly participating in party settlement discussions. This prohibition blossomed in 1974 out of a concern that judicial involvement might increase a defendant’s feelings of coercion, and as a result, could have the effect of undermining a judge’s image as a neutral arbiter. The rule has been strictly enforced such that it treats even slight intrusions as violations, and prosecutors and defendants cannot even jointly consent to judicial participation in their settlement negotiations. Formally speaking, at least, federal judges can play no direct role in facilitating or discouraging party negotiations regarding settlement, and their oversight role as presiding court officials is limited to

253 Judges, like other rational actors, are motivated by a variety of interests that they wish to advance, including success in their careers, the respect of their colleagues, ideological commitments, or simply leisure. See, e.g., Posner, supra note 248, at 39.

254 See Alschuler, supra note 251, at 1094; see also Shay Lavie & Avraham Tabbach, Litigation Signals, 58 SANTA CLARA L. REV. 1, 55 (2018) (discussing strategic signaling under conditions of asymmetric information).


257 See McConkie, supra note 29, at 65 (collecting and reviewing cases on plea bargaining). Notably, the prohibition applies most clearly to bargains involving guilty pleas. As we conceive of criminal settlement, it can include agreements that alter the nature of the dispute in many ways, including procedure modifications and the like. Of course, we doubt that Rule 11 could be interpreted to apply to a judge trying to facilitate or discourage an agreement between parties over a minor procedural adjustment. Even so, a hands-off mentality induced by Rule 11 may discourage active involvement in other areas.
the review of any agreements—specifically, plea deals—that parties strike and decide to present to them.258

But this formal prohibition does not halt considerable informal judicial influence on parties and their conversations. How could it? At a minimum, parties will always attempt to anticipate a judge’s future decisions, and those expectations may have significant effects on the strategic choices of parties.259 By uttering offhand remarks during pretrial proceedings as well as through their previous decisions in this or other cases, a judge can basically shout or whisper the degree to which they are willing to go along with a given settlement term.260 Prosecutors and defenders are often well aware of a given judge’s proclivities.261 And the parties may be able to estimate with some confidence how a particular judge will react to particular settlement practices, diminishing their uncertainty and, often, shifting the relative value of particular terms.262

Not surprisingly, judges can disrupt these expectations whenever doing so advances their own goals. They can behave strategically to enhance or

258 In those states without a Rule 11 analog prohibiting judicial involvement, judges’ ability to inject themselves strategically into negotiations is more apparent. Yet even where permitted, some judges tend to be more active and others more passive in their involvement. See King & Wright, supra note 255, at 388–92 (citing interviewees reporting that older judges, those in rural districts, and those facing reelection were more likely to participate in plea negotiations than their younger, urban, and safe counterparts).


260 For instance, “a judge might intimate that a favorable sentence would follow a guilty plea by suggesting that they can resolve the case quickly or by discussing the equities of the case in a sympathetic manner.” Alschuler, supra note 251, at 1092–94 (footnote omitted) (describing how judges influence plea bargains through “[h]ints, [i]ndirection and [c]ajolery” (emphasis omitted)). Furthermore, Federal Rule of Criminal Procedure 32(h) requires that “[b]efore the court may depart from the applicable sentencing range on a ground not identified for departure either in the presentence report or in a party’s prehearing submission, the court must give the parties reasonable notice that it is contemplating such a departure.” F ED. R. CRIM. P. 32(h). This provides yet another opportunity for a judge to strategically signal to the parties their assessment of the case.

261 Even before the Guidelines era, when sentences were “random, ill-reasoned, or disparate,” sentencing was done with “sufficient predictability and uniformity to permit widespread plea settlements.” Standen, supra note 17, at 1503; see also Jennifer L. Mnookin, Uncertain Bargains: The Rise of Plea Bargaining in America, 57 STAN. L. REV. 1721, 1741 (2005) (reviewing GEORGE FISHER, PLEA BARGAINING’S TRIUMPH: A HISTORY OF PLEA BARGAINING IN AMERICA (2003)) (“Defense attorneys and prosecutors, repeat players all, will probably be able to make reasonably informed guesses about expected sentences. Thus, the additional discretion provided under Booker’s approach to sentencing may affect the ‘price’ of the plea bargain, but it is not likely to prevent the parties from agreeing on a deal.”).

262 See Offit, supra note 259, at 1114 (drawing on ethnographic research to show that judicial proclivities influence bargaining between prosecutors and defendants); cf. Norman Lefstein, Plea Bargaining and the Trial Judge, the New ABA Standards, and the Need to Control Judicial Discretion, 59 N.C. L. REV. 477, 490 (1981) (noting this effect).
limit the degree of randomness or unpredictability that parties perceive, which can impact both the likelihood of settlement and the types of terms any settlement between the parties is likely to contain. By upping uncertainty through mixed signaling or inconsistent rulings—in all cases or just in the instant one—judges can alter settlement dynamics, perhaps reducing the likelihood of settlement by playing on the parties’ loss aversion or optimism when they have divergent priors about the probable outcome of any decision (i.e., each party interprets greater uncertainty as increasing their chances of success) or instead stimulating the parties’ desire to settle by triggering risk aversion.  

Alternatively, judges who behave predictably—by consistently sentencing within the Sentencing Guidelines’ recommended range, for instance—encourage parties to settle by reducing the expected payoff of a sentencing hearing, by making delay and further investment of resources unattractive, or by nudging parties toward settlement terms that are relatively more valuable in a world with less uncertainty.

To understand how a judge’s decision to increase or reduce uncertainty can affect the settlement behavior of parties, consider once more the five loci of control we outline above. Begin again with the settlement on a particular charge. Although parties generally enjoy considerable discretion in dismissing and amending charges, the prospect of judicial review matters to these decisions. A judge can scrutinize a prosecutor’s charging decisions and consider at sentencing those facts “found by a jury or admitted by the defendant, . . . conduct that was not charged, as well as . . . conduct underlying charges of which the defendant was acquitted.” Moreover, in cases involving multiple criminal charges, the parties know that the judge will eventually determine whether sentences will be served concurrently or consecutively, which can dramatically affect the amount of time a defendant spends incarcerated. Accordingly, a judge’s proclivities impact the value of certain settlement types. To illustrate, a defendant will attribute less value to an issue-modification term in which the prosecutor agrees to drop charges if the judge is known to uniformly issue concurrent sentences. Likewise, judges who have reputations for scrutinizing uncharged conduct will likely be met with settlements containing highly edited fact patterns or with what

---

263 See, e.g., Bibas, supra note 5, at 2533 (“Indeterminacy leaves more room for each side to be overly optimistic, to take risks, to anchor on irrelevant benchmarks, or to otherwise misestimate the likely sentence.”).

264 See supra note 13 and accompanying text.


266 See Crespo, supra note 63, at 1335.
appear to be fact patterns involving significantly less uncharged criminal conduct than we may otherwise expect to see.267

Consider next the role of the Sentencing Guidelines, and the changes in judicial behavior that appear to have followed the transition from mandatory to permissive guidance. Research shows that, after the Supreme Court decided United States v. Booker in 2005, judges follow the recommendations of the Sentencing Guidelines less than half of the time.268 Furthermore, the extent to which judges follow these Guidelines differs by type of crime.269 Moreover, some judges appear to be more amenable to guidance than others. Surveys suggest that opposition to the Guidelines is higher among federal judges with pre-Guidelines sentencing experience and lower among those without that experience.270 Thus, the level to which the Guidelines will matter in negotiations will turn on whether the prosecutor and the defendant believe their judge will find them controlling under the circumstances. A judge who always issues Guideline-compliant sentences sends a signal to parties about how best to construct their settlements. Parties will build their settlement agreements on those expectations—unless and until such a judge decides to chart a new path.

A judge’s view of the Sentencing Guidelines can also shape their position on the trial/settlement differential, or the “trial penalty.” This affects the value of certain settlement terms—particularly procedure modifications, like jury waiver, but also issue modifications that are isomorphic in their effects. The Sentencing Guidelines instruct judges to take into account the defendant’s “acceptance of responsibility” in sentencing,271 and data reveal that judges grant the responsibility-taking reduction in over ninety-five percent of federal convictions.272 Data are not available on how often a judge grants the reduction when a defendant does not plead guilty outright, but the

---

267 For an example, in 1979, the Supreme Court of California determined in People v. Harvey that dismissed charges could not be considered at sentencing as “[f]actors relating to the crime.” 602 P.2d 396, 398 (Cal. 1979) (quoting CAL. R. CT. 4.421(a) (1977) (amended 2018)). Parties responded by entering “Harvey waivers” that effectively undid the court’s decision and allowed judges to consider uncharged conduct for specific purposes, most often for restitution. See CAL. PENAL CODE § 1192.3(b) (West 2020). Under these arrangements, the defendant gets the benefit of the dropped charge without the state losing all of the benefits of bringing the charge by subsequently agreeing to dismiss it.

268 U.S. SENT’G COMM’N, supra note 102, at 78 tbl.24.

269 In 2019, for instance, 25.4% of cases involving tax offenses carried Guideline-compliant sentences. Id. at 90 tbl.31. That figure is only 18.5% for bribery/corruption offenses, and it is a mere 10% for antitrust offenses. Id. Compare those figures to the 90.8% of Guideline-compliant sentences for drug possession offenses. Id.


271 U.S. SENT’G GUIDELINES MANUAL § 3E1.1(a) (U.S. SENT’G COMM’N 2018).

272 U.S. SENT’G COMM’N, supra note 102, at 72–73 tbl.21.
Beyond Plea Bargaining

Guidelines instruct that a reduction in such a scenario should only be granted in “rare situations.”273 Yet the same Guidelines also instruct that “[t]he sentencing judge is in a unique position to evaluate a defendant’s acceptance of responsibility” and that a judicial determination is thus “entitled to great deference on review.”274 Some courts, including the U.S. Court of Appeals for the Ninth Circuit, have held that a defendant who exercises their right to trial is not per se barred from the responsibility-taking discount.275 Judges can dramatically upset the apple cart by granting sentencing discounts to defendants who accept responsibility only after enjoying an unadulterated trial. In the end, judges are free to change the “price” of jury trials as parties negotiate criminal settlements.

Judges can also influence criminal settlements by how they react to mandatory statutory sentencing scenarios. Again, judges may signal to parties by always (or usually) implementing the maximum sentence in certain types of cases or, alternatively, always concluding that the minimum is appropriate.276 Even in those instances in which prosecutors petition judges to deviate from statutory constraints—by stating that the defendant offered the government substantial assistance, for instance—uncertainty remains. This is because 18 U.S.C. § 3553(b) requires judges to determine whether the assistance is in fact substantial enough to warrant a departure.277 And there is much debate among jurists over what should be considered an “appropriate reduction.”278 A judge who endeavors to scrutinize closely the government’s recommendation or who finds few instances in which a reduction is appropriate will undercut the value of safety valves in negotiation. In response, parties will re-optimize in the face of this anticipated judicial behavior—they are free to respond by focusing on other terms and by looking for

273 The only example provided is “where a defendant goes to trial to assert and preserve issues that do not relate to factual guilt (e.g., to make a constitutional challenge to a statute or a challenge to the applicability of a statute to his conduct).” U.S. SENT’G GUIDELINES MANUAL § 3E1.1 cmt. 2.

274 Id. § 3E1.1 cmt. 5.

275 See, e.g., United States v. Dhingra, 371 F.3d 557, 568 (9th Cir. 2004) (“[A] defendant who exercises his constitutional right to a trial does not automatically forfeit the benefit of the adjustment for acceptance of responsibility.”).

276 Some judges are known as being particular about certain types of crime. Judge James S. Gwin of the U.S. District Court for the Northern District of Ohio, for instance, rails against mandatory punishments for child pornography possession and recently conducted in one trial a jury poll on sentencing in order to show “how off the mark the Federal Sentencing Guidelines are.” United States v. Collins, 828 F.3d 386, 388 (6th Cir. 2016). The mean juror recommendation was 14.5 months, while the statutory range was 262 to 237 months. Id. Judge Gwin sentenced the defendant to the statutory minimum. Id.


278 See id. at 825–26.
the most effective substitutes. Still, judicial discretion can limit the value of specific types of bargains that the judge finds unappetizing, producing a different set of criminal settlements than we may otherwise expect.

It is also relatively easy for judges to influence whether the sentencing recommendations parties present to them are binding or nonbinding. The value of one over the other turns significantly on how the parties suspect the judge will behave in light of their proposal. Some judges are known to loath binding agreements. Other judges may be more willing to accept any type of negotiated agreement so long as the proposal offers a sufficiently wide range for them to exert meaningful discretion. Imagine the frustration of a judge presented with, say, only a three-month difference between the high and the low sentences in a proposal. Alternatively, some judges may make plain that they welcome narrow binding agreements, either to reduce their workload or, more interestingly, perhaps as an invitation to engage in more meaningful back-and-forth communication. In theory, a precise recommendation would allow a judge to review the parties’ proposal and reject it, so as to signal what the court is willing to accept without running afoul of prohibitions on judicial participation in negotiations. A judge’s reputation will shape the parameters of the parties’ proposal.

Finally, judges likely have the most power to influence the dynamics and terms of criminal settlements when assessing whether a recommendation satisfies the mandatory reasonableness requirement. Judges have considerable discretion and appear to regularly consider—either consciously or unconsciously—legally irrelevant factors in fashioning sentences. For example, some data show that Black and white judges alike sentence Black defendants more harshly when compared to white defendants. Likewise, some evidence suggests that female judges sentence defendants to longer

279 See supra note 195 and accompanying text.
280 Cf. supra note 200 and accompanying text.
281 See, e.g., United States v. Haack, 403 F.3d 997, 1004 (8th Cir. 2005) (likening appellate review for reasonableness to the review for abuse of discretion).
282 See, e.g., United States v. Green, 436 F.3d 449, 456–57 (4th Cir. 2006) (“[T]he overarching standard of review for unreasonableness will not depend on whether we agree with the particular sentence selected but whether the sentence was selected pursuant to a reasoned process in accordance with law, in which the court did not give excessive weight to any relevant factor, and which effected a fair and just result in light of the relevant facts and law.” (citation omitted)); see also United States v. Haack, 403 F.3d 997, 1004 (8th Cir. 2005) (likening appellate review for reasonableness to the review for abuse of discretion).
283 See, e.g., THOMAS M. UHLMAN, RACIAL JUSTICE: BLACK JUDGES AND DEFENDANTS IN AN URBAN TRIAL COURT 78 (1979) (arguing that Black defendants are sentenced to longer sentences than white defendants by both Black and white judges).
terms and are less likely to incarcerate women. Judges appointed by Democrats appear to impose shorter sentences than Republican-appointed judges for crimes involving violence, theft, and drugs.

And so on. The main takeaway is that judges have broad sentencing discretion and appear to use their discretion in ways that are individually distinctive—systematically so. When this happens, or when a judge signals that it will happen going forward, we should expect parties to negotiate settlement agreements in light of that information. Obviously, defendants cannot alter their demographic characteristics or pick specific judges likely to be lenient toward someone in the defendant’s group, but parties can take information about judicial tendencies and biases into account in fashioning their settlements. They can use this information to constrain or leverage a judge’s decision-making in accord with their preferences. A binding, narrow recommendation, say, may be more valuable to a defendant who anticipates a biased or otherwise inequitable sentence.

In Figure 2, we illustrate our basic claim about the role of judicial discretion in the dynamics of criminal settlement. We add a second function—a concave parabola—over the top of Figure 1 in order to demonstrate how a judge’s willingness to impose a particular sentence interacts with the various

---


constraints the parties place on the judge’s decision-making freedom. The difference between the two figures is now that the Y-axis also measures the judge’s baseline preference to impose any given sentence (i.e., not taking into account the constraints imposed by the law and party settlement). In this particular example, we draw the parabola in a way that indicates that the judge’s and the parties’ preferences over various sentences are roughly aligned—i.e., the judge happens to prefer to impose the sentence recommended by the parties. But imagine the parabola shifting to the right or the left or becoming more concave (a higher peak with narrower base). If the curve shifts far to the left, we could find ourselves in a situation in which the judge is entirely unwilling to impose the statutory maximum—or even the Guideline minimum. In that scenario, there would be little for the parties to gain by dickering over terms that solely affect those values.

Just as parties manipulate the stringency of constraints on judges through how they recast their dispute, judges can behave in ways that signal how they will exercise discretion in the future and thereby change the settlements we are likely to encounter. These revelations may be truthful or strategic. Either way, judges have the power to affect the relative usefulness of various constraints and, thus, the “prices” for concessions in settlement negotiations. Imagine these two functions sliding atop one another, as both the judge and the parties manipulate the contours according to the choices and signaling of the other player. This push-pull relationship sets the values of partial-settlement terms and agreements as a whole. Modifications—whether procedural, issue-oriented, or outcome-related—are largely fungible, at least when prosecutors and defendants care only about the final sentence. Thus, as one potential manipulation of the scope of judicial discretion becomes less valuable in light of judicial preferences or signaling, optimizing prosecutors and defendants will shift their focus toward terms

---

286 As in Figure 1, the X-axis represents the length of the sentence. Solid, dashed, and dotted vertical lines continue to represent constraint points on judicial discretion that are either flexible or inflexible. See supra note 245 and accompanying text.

287 See supra note 269 and accompanying text; see also Ilya Beylin, Comment, Booker’s Unnoticed Victim: The Importance of Providing Notice Prior to Sua Sponte Non-Guideline Sentences, 74 U. CHI. L. REV. 961, 964 (2007) (“When a judge rejects a plea bargain and imposes, sua sponte, a non-Guideline sentence without providing parties an opportunity to dispute her reasoning [under Federal Rule of Criminal Procedure 32(h)], she places the bargaining parties in a precarious position: they must craft stipulated facts in a manner that provides no conceivable basis for overturning their bargain or risk the imposition of a sentence substantially different from that on which the parties agreed.”).

288 We can imagine defendants who, in addition to a lesser sentence, also value exercising their due process rights. Notions of procedural justice can be a strong and rational motivator. Cf. Prescott & Spier, supra note 4, at 79 (“Opposing litigants may agree to settlements for entirely different reasons.”).
“in the money.” Judicial decision-making influences the settlement terms we should expect to observe in practice, and judges can use this power strategically if they are so inclined.

***

Our analysis has shortcomings. First, although we claim throughout that the universe of potential criminal settlement agreements is large and diverse, we are unable to comprehensively categorize this universe. We also do not explicitly treat the role judges have in policing procedure modifications and the like during proceedings themselves. Second, judges are not necessarily privy to all settlement-related information, and this information asymmetry affects the analysis. Judges may not be aware, for instance, of agreements over collateral effects or agreements not to prosecute a defendant’s co-conspirators. These types of agreements, more so than others, occur offstage and beyond the reach of the judge—or the judge is at least likely to be in the dark about these terms. While the judge has no authority to compel a prosecutor to bring charges against a co-conspirator or recommend removal proceedings, this is not dispositive. Judicial power is fungible in the sense that although a judge may not have the right to formally review certain decisions, a judge does have indirect, informal power to sway them. As lawyers are well aware, it is often a bad idea to anger a judge even when you have the absolute right to do so.289

Our analysis may also suggest that parties bargain over each settlement term individually as they try to constrain or leverage judicial discretion and maximize their position. This often may not be the case. Instead, many crimes have a going price for “off the shelf” pleas and other concessions—options with more of a take-it-or-leave-it feel.290 Establishing standard prices and useful settlement terms for regularly occurring scenarios allows prosecutors and defenders to bargain efficiently without unrealistically burden-some transaction costs. The analysis we offer above can be reframed to represent these going rates and the logic of our arguments remains true, but it would be an error to suggest that parties always engage in careful consideration of all terms or means of guiding judicial discretion. A corollary is that, in many instances, the cost of negotiating a settlement or a particular term is too high, and so the parties stick with default rules. Complex criminal set-


290 See, e.g., Stephanos Bibas, The Myth of the Fully Informed Rational Actor, 31 ST. LOUIS U. PUB. L. REV. 79, 82 (2011) (analogizing the plea bargain market to car dealerships where most consumers get the going rate (negotiated outcomes) and “only a few suckers pay the full sticker price” (non-negotiated outcomes)).
tlements likely make the most sense in high-stakes criminal disputes in which the parties have a lot of resources or a lot to lose.

Another limitation of our settlement analysis is that it may give the erroneous impression that judges and parties inevitably stand on equal footing in affirmatively shaping settlement agreements. Often, they do not. Parties will tend to care much more about using agreements to generate surplus to share in the dispute than the judge, whose preferences may mean there is often little to gain from strategically influencing criminal settlement terms. Moreover, parties can be adept at hiding inconvenient facts from a judge, and judges are likely to be at a considerable disadvantage in terms of access to relevant settlement information. Because judges are so often overburdened with cases, they also have a strong incentive to honor party agreements and hurry the disposition of cases. Even if a judge plays their role strategically, our model suggests that nimble parties with time and resources and a lot on the line will maneuver in response so as to re-optimize. Judges in theory may have power but in practice may be outnumbered. Still, our model helps to demonstrate how partial-settlement agreements in the criminal context operate to help parties capitalize on their shared interests despite the inability to fully “settle” their dispute.

CONCLUSION

Basic economics and psychology tell us that prosecutors and defendants are likely to enter into agreements with each other when they are mutually beneficial. Our analysis shows that they can and do benefit by partially settling criminal cases. The process and prevalence of criminal settlement, however, implicates moral and political considerations, that, although beyond the reach of this Article, we should not leave completely unaddressed. The scholarship and popular discourse are replete with criticism of, and resignation to, plea bargaining as the central component of our criminal justice system. It is normatively wrong, some say, to allow the state to buy a defendant’s constitutional rights by offering coercive discounts. This wrong is particularly pronounced given the race and gender disparities that permeate

291 See supra notes 109–111 and accompanying text.
292 See supra notes 105–108 and accompanying text.
293 Albert Alschuler, for instance, fiercely criticizes criminal settlement because it “place[s] a price in dollars . . . on things that we should be reluctant to sell: human liberty, the legitimate objectives of the criminal sanction, and the right to a hearing.” Alschuler, supra note 3, at 678–79. But, as Judge Frank H. Easterbrook retorts, regulation is likely to drive bargains to the black market, and judges “serve best by preventing fraud and ensuring that bargains reflect voluntary decisions.” Hon. Frank H. Easterbrook, Plea Bargaining Is a Shadow Market, 51 DUQ. L. REV. 551, 551 (2013).
the plea-bargaining process. Yet the desire among actors to improve their lot is strong, and parties to a dispute will seek to do so even within draconian constraints. So while we can model the motivations of the prosecutor and the defendant in “buying” and “selling” to hopefully profit before a judge, there remains much tension over how and the degree to which we as a society believe this marketplace should be regulated.

The partial or incomplete settlement framework we advance here does not resolve this tension, though it may offer a fresh lens through which to view it. By recognizing that prosecutors and defendants bargain to maximize their interests along multiple dimensions—and that they will always bargain, no matter where the starting points, as long as each party possesses something the other party values more—decision-makers may choose to regulate in ways that alter the relative costs and risks associated with specific types of agreements. For instance, if our concern with criminal settlement is the ease with which the state can convince a defendant to proceed directly to sentencing, restricting bargaining per se is unnecessary when we can instead change party endowments. Altering each party’s initial stake could limit the extent to which the guilty plea is the premier concession most defendants have to offer. Or, if we worry over prosecutors’ outsized power to shape sentences, we could authorize judges to more forcefully inject themselves into negotiations to regulate the amount of risk the parties face in the bargains they consider. By deliberately releveling the playing field through changing the values of various settlement terms, policy-makers can channel parties toward more socially palatable criminal settlements.

Accordingly, shifting focus away from “plea bargaining” and toward “criminal settlement” is not an exercise in semantics. It is a comprehensive shift in thinking about prosecutors, defendants, and judges, and their varied (and at times opposing) interests and resources. Further, it allows us to see the underlying connections between what motivates all settlement behavior across contexts—civil and criminal—which is, to mitigate risks, minimize costs, and maximize ex ante returns. Because parties to criminal litigation cannot completely settle their disputes, their agreements are shaped in light of formal and informal constraints. Prosecutors and defendants modify procedures, issues, and outcomes in order to arrive at bargains most congruent with their preferences and in line with the anticipated judicial response. By giving definition to these partial settlements and demonstrating how they help parties realize their interests, our analysis underscores how conceptual-

ly limiting the idea of “plea bargaining” can be, and we hope it paves the way for a more complete and cohesive understanding of how our system resolves the vast majority of criminal disputes.