Pleas' Progress

Stephanos Bibas

*University of Iowa College of Law*

Follow this and additional works at: [https://repository.law.umich.edu/mlr](https://repository.law.umich.edu/mlr)

Part of the Criminal Law Commons, Criminal Procedure Commons, Legal History Commons, and the Legal Writing and Research Commons

**Recommended Citation**


Available at: [https://repository.law.umich.edu/mlr/vol102/iss6/5](https://repository.law.umich.edu/mlr/vol102/iss6/5)
PLEAS' PROGRESS

Stephanos Bibas*


George Fisher's new book, Plea Bargaining's Triumph,¹ is really three books in one. The first part is a careful, detailed explanation of how and why plea bargaining exploded in Middlesex County, Massachusetts in the nineteenth century. This part is the fruit of an impressive amount of original research in Massachusetts court records and newspaper archives. The second part of the book looks more broadly at other academic histories of plea bargaining in England, California, and New York. It explains how the forces that produced plea bargaining in Middlesex County likewise contributed to plea bargaining's rise elsewhere. The final part applies the lessons of history to critique current criminal procedure. In particular, Fisher criticizes the U.S. Sentencing Guidelines for tilting the balance of power toward prosecutors.

Academics have already written a number of histories of plea bargaining in Massachusetts and elsewhere, but this one is different. Fisher, a former Middlesex County prosecutor and now a professor at Stanford Law School, brings his prosecutorial perspective to bear in explaining the rise of plea bargaining. I will review Fisher's book from this same perspective, as both of us are plea-bargaining scholars and former prosecutors rather than professional historians.

My thesis is that Fisher adds an important dimension to the history of plea bargaining precisely because he looks at it with a prosecutor's eye. Instead of resting on broader social explanations of plea bargaining, which have become fashionable of late, Fisher emphasizes the caseloads, incentives, and powers of judges and prosecutors. His prosecutor's eye sees the actors' powers and incentives from a rational-actor perspective that purely academic historians often miss.

---

(Though the rational-actor approach looms large in theoretical and normative scholarship about plea bargaining today, it is largely absent from historical accounts.)

In particular, Fisher notes that most histories have focused on the incentive and desire to bargain while assuming that prosecutors had the power to do so. Fisher, however, notes that prosecutors started out with the incentive but not the power to bargain, and judges started with the power but not the incentive. Over the course of the nineteenth century, prosecutors developed powers to bargain unilaterally in some cases. Later, judges developed incentives to cooperate with prosecutors and spread bargaining more broadly, to offset their crushing civil dockets.

Unfortunately, in discussing prosecutors and judges, Fisher pays much less attention to the role of defense counsel. Indeed, his sources limit what he can say, as prosecutors drafted the indictments and motions in the court files on which he relies. Nonetheless, a fuller rational-actor account of plea bargaining must explore the roles of defense counsel as well as prosecutors and judges. Defense counsel develop bonds of trust with prosecutors and judges and establish going rates or prices for particular crimes. In other words, defense counsel serve as repeat players. Plea bargaining is possible without defense counsel, but repeat players grease the wheels, influence defendants' decisions to plead, and balance other actors' powers.

Part I of this Review considers Fisher's in-depth exploration of Middlesex County courts in the nineteenth century. Fisher's thorough review of thousands of court records and newspaper archives explains plea bargaining's rise in two main stages: First, prosecutors used their limited powers to bargain without the cooperation of judges. Fixed penalties for liquor-law violations and murder gave prosecutors the power to charge bargain unilaterally. Prosecutors later developed the technique of placing cases on file (in suspension), which eventually developed into modern probation. Second, judges, who previously had resisted ceding their sentencing power to prosecutorial bargaining, acceded to offset the mounting burden of their civil dockets. This second step allowed bargaining to expand to embrace all kinds and degrees of crime. Fisher is the first scholar to emphasize the growth in civil dockets. His account of the forces at work in Massachusetts is far superior to those of Theodore Ferdinand and especially Mary Vogel, whose work Fisher demolishes.²

Part II considers how the lessons of Middlesex County explain the rise of plea bargaining elsewhere. Fisher is perhaps too dismissive of other scholars, such as John Langbein and Milton Heumann, whose accounts of plea bargaining in many ways complement his own. Regardless, Fisher's rational-actor account convincingly explains how and why plea bargaining grew to dominate in New York, California, and England, as well as Massachusetts. Here, as in his account of Middlesex County, the only thing missing is a fuller appreciation of the role of defense counsel.

Part III concludes with the lessons of the past for the present. In recent years, the academic debate over plea bargaining has all too often revolved around whether or not to abolish plea bargaining. But as Fisher's account shows, plea bargaining will not disappear any time soon, especially because caseloads are so large. Even an unusually large drop in caseloads or a few local efforts to ban bargains probably would not kill plea bargaining. Judges, lawyers, and defendants have come to like the swiftness and certainty of bargaining and would keep liking it even if they had time to try every case. These insiders' preferences carry much more weight than the public's grave suspicion of the process. This is true because the insiders have the power, personal stakes, and better information about the low-visibility bargaining process. Thus, bargaining is here to stay. Plea bargaining has triumphed, in Fisher's phrase, because it has endeared itself to the actors with real power.

We cannot turn back the clock two centuries, and it is pointless to keep writing articles that treat jury trials as the norm. More fruitful and practical reform would give judges and defense counsel the power to check prosecutors. Fisher's critique of the United States Sentencing Guidelines rightly emphasizes this point. Prosecutorial control over charging and sentencing led to plea bargaining's rise. Just so, plea bargaining's moderation depends on checking and balancing prosecutors' unilateral charging and sentencing power. The lessons of the past contain the seeds of reform for the present.

I. THE LESSONS OF MIDDLESEX COUNTY

A. Prosecutors' Tricks of the Trade

Fisher's tale is a minuet of prosecutors' powers and judges' incentives. His account begins with the appointment of the first Middlesex County prosecutor in 1807 (p. 19). In early nineteenth-century America, public prosecutors worked part-time for low pay, so

they had to maintain private practices on the side. As a result, they had two obvious incentives to plea bargain: First, they sought to ease their crushing workloads and to make more time for their paying private clients (pp. 40-44). Second, they also liked quick, certain, easy victories, which allowed them to boast about their high conviction rates (pp. 48-49).

Though prosecutors' incentives to plea bargain were clear, their powers to plea bargain were not. Before pleading guilty, most defendants demanded a quid pro quo — an express or at least implicit promise of leniency in return. In most cases, prosecutors could not assure defendants that judges would reward pleas with lower sentences. Judges set sentences within broad sentencing ranges, and prosecutors lacked the power to bind judges' hands. So, without judicial cooperation, prosecutors were powerless (pp. 24-25, 49-51). In the early nineteenth century, judges did not share prosecutors' incentives to plea bargain. As full-time, well-paid officials, they lacked prosecutors' incentives to save time for lucrative private practice. Also, unlike prosecutors, judges did not see convictions as statistics in their win-loss ratios. And, as a matter of principle and pride, they were reluctant to surrender their sentencing power to prosecutorial bargaining (pp. 52-58).

Prosecutors could bargain, therefore, only in two exceptional areas where they could unilaterally control sentences. First, the Massachusetts liquor laws carried fines that were fixed or spanned a narrow range. Prosecutors could *nolle prosequi* (dismiss) counts of an indictment unilaterally and set the costs defendants had to pay upon conviction. Clever prosecutors would charge each violator with four liquor-law counts and *nolle* three counts in exchange for a guilty plea to the fourth. By dropping charges that carried heavy fixed penalties and setting costs lower, prosecutors could guarantee lower sentences for plea bargains and threaten heavier sentences after trial (pp. 21-30). Second, prosecutors had similar power in murder cases, because murder carried a mandatory death penalty. Prosecutors could *nolle* the indictment's allegation of malice aforethought and reduce the charge to manslaughter, which was punishable by zero to twenty years' imprisonment. Once again, determinate penalties made charge bargaining possible and reliable (pp. 33-35).

Here Fisher's prosecutorial eye sees what others have overlooked. Theodore Ferdinand, for example, has noticed the concentration of pleas and charge bargains in Massachusetts liquor-law cases. Ferdinand theorized (1) that constables invented plea bargaining while negotiating with criminals for information, (2) that tavern-keepers were savvy enough to seek bargains, and (3) that there were
no victims to oppose bargains involving victimless crimes. Though his speculation sounds plausible, it rests on no hard facts (pp. 58-61). Fisher, in contrast, sees that the pattern of overcharging and *nolles* resembles the way prosecutors overcharge and charge bargain today. He understands the link between fixed liquor-law penalties and the prosecutor’s unilateral power to *nolle*. And so he sees why in 1852 the legislature, to ban plea bargaining, required court approval to *nolle* in liquor-law cases. Once prosecutors lost their power to *nolle* unilaterally, charge bargaining in liquor cases ground to a halt (pp. 49, 51). Fisher’s prosecutorial insight sees the more elegant and compelling explanation for these facts that broad-brush paintings of social forces miss.

After the state legislature ended charge bargaining in liquor cases, prosecutors used another bargaining trick. In liquor-law and other less-serious cases, defendants pleaded guilty in return for having prosecutors place their cases “on file.” Convicted defendants were not sentenced until the prosecutor chose to move for sentencing. Placing a case on file meant putting it in suspense and not moving for sentencing unless the defendant later broke the law again. The on-file technique is nothing more than an old form of probation (pp. 63-77). Here, as elsewhere, Fisher is careful to point out the inferences and lacunae in the evidence, but his story is nonetheless convincing. The broader lesson he draws is that probation flourished because it furthered prosecutors’, and later judges’, interests in plea bargaining. Plea bargaining was so popular and powerful that legislative efforts to stamp it out caused it to spring up in new procedural forms (pp. 89-90).

**B. Defense Counsel**

If Fisher’s account is right, one would expect to see plea rates increase steadily throughout the nineteenth century. The biggest problem for his account is that guilty-plea rates look more like a “V” than a steady upward line. They hovered around sixty percent at the beginning of the century, then dropped to around forty percent for the middle third of the century, and climbed again to eighty percent by the end of the century. Fisher’s explanation is that at the beginning of the century, most defendants were likely

4. Ferdinando, *supra* note 2, at 13, 15, 66, 95, 186 (constabulary explanation); *id.* at 81 (defining Ferdinando’s category of “vice or regulatory” offenses to encompass liquor-law violations); *id.* at 50 tbl.2.2, 59 tbl.2.3, 77, 79 tbl.3.6 (showing pattern of bargaining in these cases).

5. Cf. Bordenkircher v. Hayes, 434 U.S. 357 (1978) (holding that due process permits prosecutors to threaten and carry out threats to file higher charges if defendants refuse to plead guilty; treating overcharging and forgoing charges as part of the “give-and-take” of negotiation).
unrepresented. Instead of trying to defend hopeless cases themselves, they pleaded guilty without bargains and threw themselves on the mercy of the court (pp. 93 tbl.4.1, 95-100). Thus, increasing rates of representation helped to lower plea rates. Fisher forthrightly admits the evidentiary hole in his argument: there is no solid evidence of representation rates before 1844 (p. 96). The account is plausible but invites more questions. One would think that prosecutors were more tempted to buy off represented defendants with plea bargains. After all, defense counsel made cases more time-consuming to try, were more likely to win, and were shrewder about bargaining possibilities. Fisher’s answer is that prosecutors lacked the power to bargain, because they would not put serious cases on file and let defendants go (p. 101).

The other question is why plea rates spiked upwards late in the century. Fisher’s explanation is that Massachusetts granted defendants the right to testify on their own behalf. This inadvertently hurt their trial prospects: if they remained silent, jurors assumed they were guilty. But if they took the stand, denied guilt, and were convicted, judges penalized them more harshly for perjury. Pleas came to seem a safer course (pp. 104-10). Once again, while Fisher’s account is plausible, there is little solid evidence. The problem is that defendants never explain why they are pleading guilty, and defense lawyers rarely speak publicly about their clients’ motivations. Fisher’s account rests on a few press reports, plausible but unproven. To his credit, Fisher is candid about this limitation (p. 111).

Fisher’s account would be more satisfying if it devoted more attention to defense counsel. Defense counsel improved their clients’ prospects at trial and increased the length of trial. They also increased the likelihood that defendants would plead guilty in exchange for clear concessions (pp. 98-100). Fisher suggests that represented defendants saw that their chances were better and so were less likely to plead guilty without bargains, out of desperation. Unfortunately, he does not reflect upon the other ways in which counsel probably facilitated bargains: Defense counsel doubtless served an informational role, just as they do today. They must have developed a sense of the going rate for particular crimes. They must have told clients what was a good deal for a particular case and must have encouraged them to demand appropriate concessions. And, as repeat players, defense counsel must have built bonds of trust with prosecutors, making bargains more reliable and so more desirable for both prosecutors and defendants.6

All of these factors probably facilitated plea bargaining, setting the

---

6. See HEUMANN, supra note 3, at 69-91 (describing how defense counsel learn the value of the case, develop working relationships with prosecutors, and discover what discounts are possible in various types of cases).
stage for bargaining's dramatic rise once judges saw benefits from accepting pleas.

C. Caseloads Caused Judges to Embrace Bargaining

Fisher's account shows that plea bargaining could progress only so far without the assent of judges. Without it, prosecutors could bargain only where they controlled sentencing: liquor-law cases, murder cases, and less-serious cases that they could put on file. The capstone of the Middlesex County story is why judges abandoned their opposition to plea bargaining and went along with prosecutors. The earliest scholars of plea bargaining explained it as a response to rising caseloads, a way for courts to lighten their dockets. More recently, Mary Vogel, Theodore Ferdinand, and Milton Heumann have disputed "the myth of caseload pressure" as a cause of plea bargaining.

Fisher rehabilitates caseload pressure as the explanation for judicial acquiescence in plea bargaining in several stages. First, Fisher exposes flaws in these authors' data correlating caseloads and pleas. Fisher shows that the data are either unreliable or in fact support the caseload-pressure hypothesis (pp. 9, 45-47, 294 n.31).

Second, Fisher shreds Vogel's cultural and political explanation for plea bargaining. Vogel claims that the newly enfranchised masses chafed under the enforcement of social order. Political elites, she claims, resolved the tension by bestowing "episodic leniency" on guilty pleas, thereby maintaining order while pleasing voters. As Fisher shows, however, the pattern of leniency was too small for the public to notice (pp. 141-43). Contemporary newspapers and legislative reports contained almost no mention of plea bargaining, and what little attention it drew was unfavorable (pp. 144-51). Plea bargaining, in short, was not well known and was unpopular when it was known. If outsiders were not pressing for plea bargaining, the impetus must have come from actors within the criminal justice system. This insight underscores the value of a rational-actor analysis of the actors' incentives.


8. Vogel, Courts of Trade, supra note 2, at 163-65; accord Milton Heumann, *A Note on Plea Bargaining and Case Pressure*, 9 Law & Soc'y Rev. 515 (1975) [hereinafter Heumann, *Case Pressure*]; see also FERDINAND, supra note 2, at 93 ("Whether plea bargaining was a response to burgeoning caseloads... is debatable.").


10. See also Carolyn B. Ramsey, *The Discretionary Power of 'Public' Prosecutors in Historical Perspective*, 39 Am. Crim. L. Rev. 1309, 1336-37, 1385-86 (2002) (citing nineteenth-century New York City press clippings to show that when the public did learn of plea bargaining, it was outraged by the practice).
Finally, Fisher has the insight to look at judges' overall workloads, instead of just their criminal dockets. On the criminal side, cases per judge peaked in the mid-nineteenth century and then declined while pleas rose (pp. 116-17). Moreover, criminal trials did not grow much longer until the end of the century, contrary to Langbein's explanation for plea bargaining's rise. But judges' civil dockets were exploding. As the population grew and mechanization spread, people began to suffer many more injuries on railroads and streetcars and in factories and textile mills. Tort suits were more time-consuming than the old civil-contract cases (pp. 121-24). Judges could do little to reduce their civil caseloads, but they could offset them by disposing of their criminal cases quickly. (This argument of course assumes that the same judges had both civil and criminal dockets, which is often but not universally true.) Judges controlled sentences, whereas juries controlled liability and damages. Thus, as Fisher perpectively notes, judges "had far greater power to coerce pleas on the criminal side than to induce settlements on the civil side" (p. 123). Judges could promote pleas by promising rewards, or by habitually following prosecutors' sentence recommendations, or by letting defendants withdraw their pleas if the sentence exceeded the recommended one (pp. 129-36). Though in hindsight the role of civil cases seems to be common sense, Fisher is the first to see it.

Fisher's analysis is brilliant in its elegant simplicity. Until now, scholars assumed that actors had the power to bargain, rather than pointing to the precise sources of this power. Ordinarily, prosecutors and judges must join forces to generate plea bargains. Before the civil-docket crunch, prosecutors had the incentive but not the power to bargain, whereas judges had the power but not the incentive. Prosecutors found small pockets of power that they could exercise unilaterally in liquor-law and murder cases and then in cases minor enough for on-file treatment. But once civil-caseload pressure changed judges' incentives, they worked together with prosecutors to bargain more broadly. Scholars need not invoke a complex, cumbersome conglomeries of cultural and social forces to explain plea bargaining's rise. The courtroom actors all came to like it as a way to counteract caseload pressure, and so it triumphed despite public ignorance and legislative opposition.

II. THE MARCH OF BARGAINING ELSEWHERE

The Middlesex County tale is the core of Fisher's book and occupies about three-fifths of its text. Middlesex County is interesting not only in its own right, but also because it casts light on the march of

plea bargaining elsewhere. Fisher devotes a chapter to explaining how the forces in Middlesex County produced plea bargaining in three other jurisdictions: England, California, and New York. Some scholars have criticized Fisher for building on one narrow case study instead of starting with a broader statistical net, but this objection is unfair. Fisher's strength is his focused depth in the first part of the book, which lays a solid foundation for his breadth later on. Fisher makes a powerful case that the forces he identified in Middlesex County manifested themselves elsewhere as well.

Fisher begins with England and asks why plea bargaining was rare if not absent in eighteenth-century England. John Langbein and Malcolm Feeley have shown that there was no plea bargaining in the Old Bailey (London's main criminal court) in the eighteenth century. They argued that the reason for the lack of bargaining was the brevity of trials, which took hours or minutes instead of days. This explanation is a variant of the caseload-pressure hypothesis. As trial length grew, so did the total time demanded to try all cases. As judges became busier, so the argument goes, they had to use pleas to lessen their workloads.

Fisher agrees that there was little or no plea bargaining but points to a powerful alternative explanation: the absence of public prosecutors until the late nineteenth century. In England, victims had to prosecute their own cases as private prosecutors because there were no public prosecutors. As the Middlesex case study shows, public prosecutors had the incentives to bargain and had some power, especially the power to nolle. English crime victims lacked this power. They did have some incentive to bargain to save time and money and to avoid the risk of loss (pp. 155-56). (Fisher should have added the obvious point that victims still had less incentive to bargain than public prosecutors, who lacked personal stakes in seeing justice done.) Victims did find a few ways to bargain, for example by not showing up to testify in exchange for cash. Generally, however, private prosecutors lacked the legal tools and expertise to bargain (pp. 156-57). The lesson of Middlesex is that bargaining requires power and...
knowledge of the legal system. With neither the power to *nolle* nor the expertise of repeat players, private prosecutors could not bargain. Fisher also notes that Old Bailey judges had little incentive to bargain. Their caseloads were growing very slowly, and the thorough press coverage of the Old Bailey would have exposed any sentence bargaining to uncomfortable scrutiny (pp. 157-58).

I would add that the absence of defense counsel was another impediment to plea bargains.\(^\text{14}\) Though we often think of defense counsel as adversaries who fight the prosecution at every turn, in fact their role is much more cooperative and facilitates bargains. The absence of defense counsel deprived defendants of repeat players who knew the system, had developed trust with other actors, and knew the value of cases. In addition, there were no defense counsel to combat clients' overconfidence, offering them sobering assessments of their trial prospects and persuading them to plead guilty. Sometimes defense counsel's assessments are clear-eyed; at other times, they may be jaundiced and self-serving, as lawyers seek to dispose of cases and collect their fees quickly. Either way, defense counsels' advice might well have produced more plea bargains and fewer trials.

Fisher treats his explanation as a novel alternative to Feeley and Langbein's, but it would be fairer to understand it as a complement. Though sometimes bargaining has risen without trials becoming longer, at other times longer trials have correlated with rises in plea bargaining. More procedural rights, longer trials, and increasing case filings cause congested dockets and caseload pressure to plea bargain. The more time-consuming each trial is and the more cases there are, the more cases have to be disposed of by other means. Thus, as Fisher acknowledges in his prologue, his caseload explanation supplements Langbein's but in truth does not supplant it (pp. 9-10).

Where English judges faced less press scrutiny, plea bargaining took off. Fisher studied sentencing patterns for petty larceny in Manchester in the late eighteenth century. He found that almost half of defendants who stood trial suffered seven years' transportation abroad, which included a dangerous ocean voyage to Australia or elsewhere. In contrast, every single defendant who pleaded guilty got a short term of imprisonment instead of a dangerous exile (pp. 158-59). Fisher's main explanation is that Manchester's petty larceny caseload quadrupled within a generation (pp. 159-60).\(^\text{15}\) In addition, in

\(^{14}\) Langbein, *Understanding*, supra note 13, at 263. Although Langbein noted the absence of prosecutors and defense counsel, he used this fact only to explain that trial procedures were much simpler before lawyers complicated things.

\(^{15}\) This dramatic growth makes sense. Manchester was one of the new industrial cities that burgeoned in the Industrial Revolution, which must have brought many more industrial accidents and tort suits as well as crimes, whereas London had long been a large, bustling city and was not growing as fast. See William Roberts, *A Charge to the Grand Jury of the Court Leet, for the Manor of Manchester, in 9 The Court Leet Records of the Manor*
1790 the magistrates’ new house of corrections opened, which at the
time was a novel penological effort to reform convicts. Reform-
mined magistrates may have offered this new opportunity to
defendants who were the most promising candidates for reform,
namely those who pleaded guilty and seemed contrite (pp. 159-60).
The burgeoning caseload of Manchester fits well with Fisher’s
rehabilitation of the caseload pressure explanation in Middlesex
County.

Fisher then crosses the Atlantic to consider two American
jurisdictions. First, he reviews Lawrence Friedman and Robert
Percival’s history of criminal justice in Alameda County, California,
between 1870 and 1910. During that time, guilty pleas grew more
slowly than in Middlesex County. Guilty pleas to the charges in the
original indictment increased from 12% to 32% of all pleas entered as
guilty pleas to lesser offenses decreased from 10% to 4%. There was
little charge bargaining, Fisher argues, because prosecutors had little
power to do so. In 1851, the California legislature took the power to
nolle away from prosecutors and gave the power to dismiss to judges.
Just as this step killed liquor-law charge bargaining in Middlesex
County, so it greatly limited post-indictment charge bargaining in
Alameda County (pp. 161-62). Also, judges had less incentive to
bargain. There were many more judges and many fewer cases in
Alameda County, which meant much less caseload pressure (pp.
163-64).

Although Alameda County judges had little interest in plea
bargaining, the other actors did. Defense lawyers gained incentives to
bargain in 1872 when the California state legislature enacted
recidivism enhancements with stiff mandatory minimum sentences.

Further legislation in 1880 created the preliminary examination, which

OF MANCHESTER app. I, at 243 (Manchester, Henry Blacklock & Co. 1889) (reporting that
Manchester’s population was growing so quickly that the criminal courts could not hear and
punish all crimes); ROBERT ROBSON, THE ATTORNEY IN EIGHTEENTH-CENTURY
ENGLAND app. IV, at 166-67 (1959) (reporting that from 1790 to 1800 the number of
practicing attorneys in London grew from 1755 to 1800, whereas the number in Manchester
grew from 40 to 61), Compare David Cody, A Brief History of London, The Victorian Web,
population increased from 300,000 in 1700 to 750,000 in 1760 to 900,000 in 1800), with
SPARTACUS EDUC., Manchester, at http://www.spartacus.schoolnet.co.uk/
ITmanchester.htm (last modified July 25, 2002) (reporting that Manchester and neighboring
Salford grew from about 25,000 in 1772 to 95,000 in 1800).

16. LAWRENCE M. FRIEDMAN & ROBERT V. PERCIVAL, THE ROOTS OF JUSTICE:

17. Id. at 174 tbl.5.12 (tabulating all pleas of guilty or not guilty, and so excluding cases
that were continued indefinitely or dismissed before plea, for example).


19. See CAL. PENAL CODE §§ 666, 667 (Bancroft-Whitney Co. 1885); Ex parte
Gutierrez, 45 Cal. 429, 430 (1873).
allowed prosecutors and defense counsel to meet and review evidence before prosecutors filed charges. But, as noted, judges still lacked the incentive to go along. So prosecutors and defense counsel struck charge bargains before the filing of formal charges. By doing so, they averted the harsh mandatory sentences and circumvented judges' power over dismissal of already-filed charges. This explains why pleas to the offense charged rose dramatically as pleas to lesser offenses plummeted (pp. 165-66). Once Fisher explains the point in terms of the powers and incentives of the actors, it seems obvious in hindsight. And yet the point eluded a legal historian as eminent as Lawrence Friedman, who lacked a prosecutor's eyes.

Finally, Fisher considers New York, relying on two studies, one by Raymond Moley, the other by Michael McConville and Chester Mirsky. Beginning around 1830, guilty pleas rose from almost zero to more than three-quarters of all dispositions before the end of the century. Seventy percent (or more) of guilty pleas in New York City were to lesser offenses, which suggests a huge amount of charge bargaining. The caseload of New York City's major criminal court quintupled between 1839 and 1865, which probably created great caseload pressure to plea bargain (pp. 167-68).

Based on the example of Middlesex County, Fisher supposes that prosecutors were the ones doing the charge bargaining to alleviate the increased workloads. His question, then, is where they found the power to do so. The 1829 criminal code, enacted just as guilty pleas began to rise, divided crimes into multiple degrees. For example, first-degree burglary was punishable by a minimum of ten years' imprisonment, second-degree burglary by five to ten years, and third-degree burglary by zero to five years. Prosecutors could charge first-, second-, and third-degree burglary in the same indictment and then nolle the higher-degree counts as part of a charge bargain. The code contained many minimum sentences for degrees of burglary, robbery,

20. FRIEDMAN & PERCIVAL, supra note 16, at 43-44.
22. Mcconville & Mirsky, supra note 21, at 466 fig.3; Moley, supra note 7, at 108.
23. McConville & Mirsky, supra note 21, at 466 (reporting 70% in an 1865 sample); Moley, supra note 7, at 111 (reporting 85% in 1926, though rates were much lower outside of New York City).
24. McConville & Mirsky, supra note 21, at 463.
26. 2 N.Y. REV. STAT. pt. 4, ch. 2, tit. 4, § 51 (1829); People v. Porter, 4 Park. 524, 526 (N.Y. Oyer & Terminer 1860). Although some statutory language required court approval for nolles, prosecutors appear to have continued to nolle individual counts freely, either because the nolle restriction was not read to limit nolles of individual counts or because judges, used to the prior nolle practice, routinely assented. Fp. 171-74.
forgery, manslaughter, and arson. These mandatory minima allowed prosecutors to tie judges’ hands and to credibly threaten higher sentences after trial, which induced defendants to plead guilty. Defendants could also plead guilty to attempts instead of completed crimes, which lowered their sentencing ranges. Thus, in McConville and Mirsky’s 1865 sample of guilty pleas to lesser offenses, forty-seven percent were reductions from completed crimes to attempt. Finally, the 1829 New York Code enacted stiff minimum sentences for repeat offenders, which prosecutors may have been able to bargain away. Fisher’s analysis of statutory power explains why plea bargaining in New York took the form of charge bargaining and why it started its rise earlier than elsewhere, around 1830. Once the law gave prosecutors the tools to bargain, they used them as broadly as the law allowed, just as they did in Middlesex County.

III. LESSONS OF THE PAST FOR THE PRESENT

In his last two chapters, Fisher draws together the strands of his history to understand the principles underlying plea bargaining. As earlier in the book, his focus is on the courtroom actors and their powers and incentives to bargain. He begins by looking at how plea bargaining serves courtroom actors. Bargaining is efficient, assures victory, and guards against reversal (especially once appellate review expanded in the nineteenth century). So far, all of what Fisher says is consistent with what Milton Heumann and others have long pointed out. There is only one flaw: Fisher’s discussion of prosecutors’ and judges’ incentives is once again marred by his omission of defense counsel’s incentives from the picture. (As we shall see, however, he does touch on the issue somewhat later.)

Next, Fisher turns to how plea bargaining serves the criminal-justice system. By delivering verdicts that appear truthful in clear cases, it keeps juries from getting easy cases wrong. When a jury gets the Rodney King or O.J. Simpson case wrong, the public loses faith in the system as a whole. But these cases are anomalies that would ordinarily result in plea bargains. By skimming off almost all easy cases, plea bargaining leaves only hard ones for trial. And when only hard cases go into the black box, there are no clear cases left for juries to get wrong. The result is more faith in the system. In short, plea

27. 2 N.Y. REV. STAT. pt. 4, ch. 1, tit. 1, § 1; id. at pt. 4, ch. 1, tit. 3, §§ 9, 20, 21, 57 (1829).
28. See, e.g., 2 N.Y. REV. STAT. pt. 4, ch. 1, tit. 3, § 21; id. at pt. 4, ch. 1, tit. 7, §§ 3(2), 27 (1829) (implying that attempts were lesser-included offenses of completed crimes).
29. McConville & Mirsky, supra note 21, at 466-67.
31. Pp. 175-77; see, e.g., HEUMANN, supra note 3, at 110-14, 144.
bargaining protects the public perception of legitimacy and justice, by averting blatantly wrong outcomes.\textsuperscript{32} The irony is that plea bargaining, having swallowed up most trials, now protects the jury ideal and the few actual trials that remain.

Plea bargaining has proven to be so powerful that it fostered procedures that are compatible with it and stunted those that could stand in its way. Fisher's previous discussion of how on-file plea bargaining led to probation's rise in Middlesex County is an example of a symbiosis of these two procedures. In contrast, he argues that plea bargaining thwarted the growth of truly indeterminate sentences, which were premised on individualized assessments of rehabilitation while in prison. Because true indeterminate sentencing would have left all sentencing to the discretion of prison and parole authorities, it would have made the benefits of guilty pleas speculative and thus undesirable. To avert this problem, judges rarely exercised their option to impose indeterminate sentences when they had a choice. Parole boards granted near-automatic parole after the minimum term, and they were more likely to parole defendants who had pleaded guilty. Finally, judges sentenced defendants to narrow ranges (say, eight to nine years instead of five to fifteen), leaving much less discretion to parole boards. In short, defendants who pleaded guilty could effectively predict their sentences, because judges and parole boards behaved in predictable ways. The effect was to stunt the ideal of indeterminacy as the price for plea bargaining (pp. 181-93). I have only one quibble: Fisher personifies plea bargaining and often speaks of how it swept away all obstacles in its path, as if it had life and power of its own. This anthropomorphic locution is not only jarring, but also obscures the human actors who had the incentives and found the powers to make it happen. Nonetheless, his account is a compelling tale of how the actors who benefited from plea bargaining could sabotage reforms that got in their way.

Fisher also does a nice job of showing how the growth of public defenders complemented plea bargaining. Some of the early advocates of public-defender systems argued that public defenders would work not only to acquit the innocent, but also to convict the guilty! Encouraging guilty pleas, they argued, was both more efficient and better for rehabilitation because confession is the first step toward reform (pp. 194-96). Perhaps these statements were sincere, or perhaps they were merely for public consumption. Either way, advocates knew they had to sell the system to the powers that be as an

\textsuperscript{32} Pp. 178-80; cf. Stephanos Bibas, Harmonizing Substantive-Criminal-Law Values and Criminal Procedure: The Case of Alford and Nolo Contendere Pleas, 88 CORNELL L. REV. 1361, 1386-88, 1407 (2003) (arguing that Alford pleas, which are guilty pleas by those whom the public perceives may be innocent, undercut public faith in and the legitimacy of the criminal-justice system).
adjunct to plea bargaining (pp. 196-97). Thus, it is no surprise that public defenders had the highest guilty-plea rates: seventy percent, as compared to sixty-two percent for appointed private counsel and forty-nine percent for privately paid lawyers. Fisher correctly chalks these plea rates up to caseload pressure and the desire to keep good working relations with judges and prosecutors (pp. 198-200).

Closely related to Fisher's points are some other incentives that public defenders have to plea bargain. As repeat players, public defenders have more plea-bargaining expertise, which should make bargains better, more trustworthy, and more predictable for defendants. And the guilty-plea rates cited above suggest that fee structures may have affected bargains. Public defenders, often overworked and paid a flat salary, have the strongest incentives to plea bargain. Court-appointed private counsel typically receive hourly fees subject to a low cap, which makes trials unprofitable. In contrast, private lawyers who are paid by the hour earn more if they go to trial and so have less pressure to bargain. Though Fisher does not express these observations, all of these points complement his account.

By focusing on powers and efficiency, Fisher brings a fresh perspective to the debate over the Warren Court's criminal procedure revolution. Liberals typically praise *Miranda* and other rights as ensuring fair trials, while conservatives fear that they let the guilty walk free on technicalities. Both criticisms, however, ignore the impact of plea bargaining. Because the new rights made trials even more cumbersome (echoing Langbein's point), they increased the efficiency gains from plea bargaining. And by giving the parties yet another dimension along which to bargain, the rights created more opportunities to bargain. The upshot was not fairer trials or scot-free criminals, but discounted sentences in exchange for waivers of these rights (pp. 202-03). New procedural rights, ironically, meant fewer trials and less process.

Fisher's final chapter looks at the balance of power between judges and prosecutors. First, he notes that the advent of bench trials in the twentieth century may have checked prosecutors. The early plea-bargaining theorist Raymond Moley thought that judges would check plea bargaining if they could. Fisher sees, however, that judges have come to like bargaining's efficiency and would use the power of bench trials to strike more bargains. Ordinarily, prosecutors have great plea-bargaining power because they can charge offenses that carry mandatory minimum sentences and then bargain them away. But in those states that allow defendants to choose bench trials without prosecutorial approval, judges can subvert prosecutorial overcharging. The judge can advise the defendant to choose a bench trial, at which

the judge will acquit of the highest charge to reduce the sentence. This counterweight restores some check on prosecutorial power. The result is that judges moderate sentences, resulting in more plea bargains at more lenient levels (pp. 205-09). As far as I know, Fisher is the first to make this point. He draws on his personal experience as a prosecutor, and his point accords with my own experience as well.34

Conversely, if prosecutors can veto bench trials and charge many crimes with mandatory minima, they can thwart judicial leniency. Today, many states give prosecutors this veto, and many statutes contain mandatory minima. As Fisher perceptively notes, rigid sentencing guidelines are the functional equivalent of mandatory minima, as they tie judges’ hands and their ability to check prosecutors (p. 210).

Fisher critiques the U.S. Sentencing Guidelines for further skewing the balance of plea-bargaining power. The Guidelines have arguably succeeded in their stated goal of cabining judicial discretion. They have not, however, counteracted prosecutors’ substantial charging discretion. The Guidelines try to offset prosecutorial discretion by considering the defendant’s actual conduct and not just the offense charged. Prosecutors can trump the Guidelines, however, by charging offenses with low statutory maxima or high statutory minima. The result is a substantial imbalance of power in favor of prosecutors (pp. 210-13).

The Sentencing Commission thought that prosecutorial leniency was a source of sentencing disparity and that the solution was judicial power to increase sentences. Thus, the Guidelines empower judges to reject overly lenient plea bargains and to include uncharged conduct in sentences. Most judges, however, like plea bargains and dislike the newly increased penalties. So, few judges reject bargains or demand harsher terms at the risk of triggering a trial. The bigger problem is not checking prosecutorial leniency but checking prosecutorial harshness. Prosecutors can overcharge to gain leverage for harsh sentences, and judges have little power to check prosecutorial harshness. If judges try to cut sweet deals unilaterally, say by departing from the Guidelines, they face appellate reversal (pp. 212-21).

In theory, the Guidelines are merely guidelines and leave judges flexibility to depart upward or downward in atypical cases.35 Indeed, the Supreme Court reaffirmed district courts’ discretion to do so in Koon v. United States.36 Nonetheless, appellate courts have read Koon


narrowly, district courts remain afraid of reversal if they depart, and Congress has recently abrogated Koon by legislation.37 Thus, fewer than one percent of cases receive upward departures, and no more than about seven percent of cases receive downward departures over a prosecutor’s objection (pp. 217-18). Appellate review deters deviation from the sentencing range set by the prosecutor’s charges and the parties’ factual stipulations. The result is an imbalance of bargaining power.

The Guidelines have had the unintended consequence of driving up plea bargaining and reducing the number of trials still further. By making sentences more predictable, they have reduced uncertainty and the corresponding chance that defendants will go to trial out of optimistic hope for judicial leniency (pp. 223, 225). And with their intricate calculations, they create many more dimensions along which parties can strike deals. The parties can now bargain over charges, the true facts, the applicability of sentencing factors, and cooperation with law enforcement (pp. 227-29). Far from thwarting plea bargaining, the Guidelines have simply channeled it into new avenues.

The broader lesson here is that plea bargaining is a force that is here to stay. Legislatures have tried to stamp it out, academics have argued against it, and theorists have proposed elaborate alternatives to it. To give two examples, the Guidelines’ real-offense sentencing system was supposed to peg sentences to defendants’ actual conduct and thus eliminate bargaining over what crimes to charge. And the fixed reduction for acceptance of responsibility was supposed to replace negotiable, bargained discounts for pleas.38 But today, no set of rules can abolish bargaining or put the genie back in the bottle. The font of bargaining, as Fisher’s history shows, is the desire for efficiency and the power to manipulate charges and sentences. Prosecutors, judges, and defense counsel have long since become addicted to the efficiency of bargaining. Perhaps in isolated places a powerful figure can stand against the tide for a time, by appealing to the public’s understandable distrust of bargaining.39 Over time, however, the desire for speedy, certain, efficient convictions causes even these isolated


38. See U.S. SENTENCING GUIDELINES MANUAL § 3E1.1 (2002).

bargaining bans to crumble. What matters ultimately is not the public's suspicions, but the courtroom actors' powers. As overlapping statutes and mandatory minima have spawned more charge bargaining, and sentencing laws have grown more complex, the actors have gained more dimensions along which to bargain. The Guidelines, which were supposed to shackle prosecutors, have instead become their tool.

If plea bargaining has triumphed, it is pointless to write yet more articles lamenting the demise of the jury trial. The better solution is to view the problem in Fisher's terms, as a balance of power. Efforts to destroy prosecutors' powers by drafting paper rules, such as the Guidelines, have failed. The more promising possibility is to create a balance of power, by giving other actors more power to check line prosecutors.

For example, mandatory minimum sentences are not truly mandatory, because prosecutors can usually choose not to charge them. Thus, mandatory minima skew the balance of power by binding judges but not prosecutors. Some states have experimented with forbidding prosecutorial bargaining over certain crimes. Prosecutors may, however, circumvent these laws by bargaining before indictment. The way to check prosecutorial abuse of mandatory minima is not by enacting paper restrictions on bargaining, but rather by creating a genuine balance. To balance this power, legislatures should restrict mandatory minima or allow judges to go below them in more situations. In an ideal world, legislatures would also simplify their criminal codes, to reduce overlapping statutes and to lessen the opportunities for charge bargaining. (If complex laws such as the Guidelines are giving prosecutors more tools, simplified laws would take some of these tools away.) Legislatures could also make recidivist enhancements turn not on prosecutorial charging decisions, but on whether the enhancement applies on the facts. These reforms, however, are unlikely to occur. Legislatures like giving more bargaining chips to prosecutors, so that legislatures both look tough on crime and dispose of more cases via pleas.

If legislatures will not lead reform, perhaps head prosecutors will. If line prosecutors have the power to overcharge and penalize


41. See, e.g., CAL. PENAL CODE § 1192.7 (West 2004).

42. See CANDACE MCCOY, POLITICS AND PLEA BARGAINING: VICTIMS' RIGHTS IN CALIFORNIA 89-128 (1993) (noting that prosecutors evaded bargaining restrictions by bargaining at earlier stages, before the preliminary examination or much discovery).

disfavored clients or lawyers, we may want to limit charging. There are various ways to do this. For instance, a separate charging unit within a district attorney’s office would have less incentive to overcharge than an assistant who is trying to hoard plea-bargaining chips. More supervision of line prosecutors’ charging and nolle decisions by higher-ranking prosecutors would restrict their freedom to overcharge and bargain away. For example, the Reagan and first Bush Justice Departments adopted policies that restricted line prosecutors’ charge bargaining. Prosecutors had to file the most serious readily provable charges, had to refrain from bargaining them away, had to put plea bargains in writing, and had to get supervisory approval for bargains. The effect was to ensure more consistency, counteracting line prosecutors’ incentives to bargain to ease their own workloads. In the same vein, Attorney General John Ashcroft has recently forbidden federal prosecutors to conceal or misrepresent sentencing facts or acquiesce in illegal departures.

There are other ways to adjust prosecutors’ incentives to bargain. Because most head prosecutors are elected, greater transparency and publicity of large charge reductions could deter overcharging in the first place. A few may even successfully campaign on a platform of restricting bargaining, as the New Orleans district attorney did for a time. (The difficulty is explaining to the public why prosecutors are declining to prosecute many marginal cases rather than bargain them away.) The point is that restrictions on bargaining will come not from paper bargaining bans, but from creating counterweights and contrary incentives.

Finally, one might reform the presentence report process, giving probation officers a larger role. If prosecutorial manipulation of the facts skews judges’ sentencing decisions, probation officers may need


47. See Wright & Miller, supra note 39.
more independent investigative resources to ferret out what actually happened. This focus on powers, resources, and incentives is much more realistic than simply announcing a ban and pretending that prosecutors will not circumvent it.

Fisher's thoughtful history, in short, is much more than history. The past is prologue, and the present follows the same rules of power and incentive that Middlesex County did two centuries ago. Thus, his account sparks fruitful reflection on how to reform the system we have today. Fisher's prosecutorial eye adds a welcome dimension to the historical literature, correcting the recent emphasis on broader social forces. The key is to implement these lessons by creating concrete checks and balances.