ESSAY

THE PERVERSE EFFECTS OF EFFICIENCY IN CRIMINAL PROCESS

Darryl K. Brown*

INTRODUCTION

The need for greater efficiency in legal process is an undisputed premise of modern policy, practice, and scholarship, and efficiency’s virtues hardly merit debate. A central part of the story of modern adjudication is achieving greater efficiency in processing and resolving cases. This is a key explanation for the “vanishing trial” and the dominance of practices that replaced it—settlement, pre-trial judgments, and alternative forums on the civil side, and plea bargaining and pretrial diversion on the criminal. The U.S. Supreme Court cites the need for efficiency in its explanations for a range of decisions, and the desirability of efficiency is a pervasive, core premise of much legal scholarship on adjudication. Yet I argue here that there are unrecognized risks inherent to greater process efficiency; it may do criminal justice more harm than good.

For litigation process in particular, explanations for the unavoidable need for efficiency are familiar. I limit this account to criminal adjudication, but the basic story is the same on both sides of the docket: The number of cases that courts must resolve has grown relentlessly for decades, and the public infrastructure has not kept pace. Courts lack the staff and resources to adjudicate all cases by trial. On the criminal side, prosecutors’ offices likewise lack the capacity to try every case they initiate. Hence the necessity for settlement, alternative dispute resolution,

* O.M. Vicars Professor of Law and E. James Kelly, Jr.-Class of 1965 Research Professor of Law, University of Virginia School of Law. I would like to thank participants in a faculty workshop at Fordham Law School for helpful feedback on an earlier draft. I also thank Mark Patterson, John Pfaff, and Jenia Iontcheva Turner for reading earlier versions and providing extensive, perceptive comments.
and summary judgment on the civil docket, and plea bargaining on the criminal.¹

Why this strain on adjudication resources exists, and how much it has changed over time, are more disputed questions. The most common explanation looks to rising caseloads or, more precisely, to the fact that the ratio of cases to court capacity has grown increasingly strained over time. The number of civil and criminal cases increases over time for many reasons; funding for courts (and prosecutors and defenders) commonly does not keep pace. Adjudication resources must be spread more thinly over the growing number of cases, and the justice system increasingly resolves cases by means other than traditional trials. Without practices to clear cases more quickly from court dockets, delays and backlogs grow.

There are many possible reasons for this increasing mismatch of cases and adjudication resources. Criminal offenses and civil causes of action expanded in the twentieth century; modern life generated more crime, more civil injuries, and more contract disputes; trial processes grew more professionalized, formal, and elaborate as contemporary notions of fairness and due process evolved.² Some of these explanations, such as

¹ Explicit concerns about criminal courts’ efficiency are longstanding. See State v. Worden, 1 Crim. L. Mag. 178, 195 (Conn. 1880) (concluding that a requirement for “cumbersome detail and heavy machinery of trial by jury” would “effectually paralyze[]” enforcement of minor regulatory offenses); Jed Handelsman Shugerman, The People’s Courts: Pursuing Judicial Independence in America 177–207 (2012) (describing 1930s reformers, including California prosecutor Earl Warren, arguing for a need for greater efficiency in criminal courts, notably in the wake of a rise in organized crime, and especially in light of the 1930s organized crime wave).

caseload pressure as the cause of plea bargaining, are hotly disputed by scholars. The reasons that judges, policymakers, scholars, and others experience pressure for greater adjudicative efficiency is not the focus here. The starting point is that the perceived need for greater efficiency is widespread and longstanding; as the primary diagnosis of adjudication’s modern predicament, it has driven the relentless trend to resolve each case more quickly and cheaply. The pressure for efficiency has deeply reshaped adjudication practice, driving innovation of non-trial practices in order to match caseloads to court capacity.

My focus in this Essay is on largely overlooked complications regarding the nature and consequences of efficiency—what we understand that term to mean and what its effects are in the context of caseloads and adjudication resources. The problem is different in the civil law context, where the state’s control over caseloads is largely less direct. In criminal law, the state controls both adjudication resources and caseloads. Legislatures define all offenses, and prosecutors initiate every criminal charge. In that sense, prosecutors determine the caseload burdens that courts (and prosecution staffs) face even if charging is, as a matter of policy, also partially a function of crime rates.

Public officials determine the supply of adjudication capacity (judicial process leading to criminal judgments). They likewise determine the demand for that service, if we understand demand here as the demand for adjudication of criminal charges. When demand for criminal adjudication grows for any number of reasons—rising crime rates, better investigative efforts, changes in enforcement priorities—it puts pressure
on the supply of adjudication services. In response, the state has the option not only of supplying more of the same service—by funding more courts and judges—but of substitution. It can switch from a costlier process for reaching judgments, such as trials, to a cheaper one, such as settlement. That kind of substitution is commonly described in criminal adjudication as improved efficiency. In blunt terms, it describes the production of more criminal court judgments at a lower per-unit cost, and therein lies the complication. Hold aside for the moment that this picture focuses only on how trials and plea bargains are substitutes for each other (as means to achieve convictions) and ignores how they are not (for example, as modes of public process and democratic supervision of prosecution policy). This sort of efficiency gain—a decrease in the cost of the service—can be expected, on the premise of an ordinary demand function, to trigger more demand. Cheaper adjudication can lead to even more cases entering the criminal court system.

That response is a widely recognized and routine effect of efficiency improvements in all sorts of contexts, but it is little discussed in criminal adjudication. In many settings this effect is not only unproblematic but welcome; it may be the goal of improving efficiency. In other contexts, however, increased demand as a result of increased efficiency is undesirable, even yielding perverse results. In what follows, I lay out reasons to suspect that the consequences of criminal adjudication’s greater efficiency take the latter form at least as much as the former. Plea bargaining and a range of related, doctrinally authorized practices for making criminal process more efficient can perversely increase demand for criminal prosecutions, rather than serving as a means to meet demand for enforcement that is driven by crime rates. Put differently, the effects of improved efficiency in criminal adjudication are—at a minimum—ambiguous; demand for criminal adjudication is likely in some part endogenous to the cost of adjudication, even though courts and policymakers typically imply that criminal caseloads are exogenous and are solely a function of non-price determinants such as crime rates and enforcement priorities.

That is the argument of Part III. To understand this possibility, Part I briefly reviews some specifics of adjudication law motivated by efficiency. Part II looks more closely at precisely what efficiency means, at the diverse consequences it can produce even for a single specific activity like adjudication, and thereby how improved efficiency can generate perverse outcomes. Part IV considers some of efficiency’s possible dis-
tributational effects within the criminal justice system, such as enabling a shift in spending from adjudication to incarceration, and it highlights data that suggest this effect has occurred. Part V explores some of the normative implications of these efficiency effects on the norms and purposes of adjudication and for criminal justice policy more broadly. The final Part concludes with observations on efficiency’s implications for criminal law enforcement policy.

I. THE LAW OF EFFICIENCY IN CRIMINAL ADJUDICATION

Because much of this story is familiar, this Part provides only a brief overview of the changes in criminal adjudication rules that have increased criminal process efficiency, while also redefining its purposes. The biggest part of the story is the much-studied triumph of plea bargaining over trials. Whether plea bargaining was in fact a product of necessity—meaning a product of caseloads rising faster than resources for courts and prosecutors—is a matter of debate among plea bargaining scholars. There is no question, however, that the Supreme Court, following others, believes that it was.

The point that gets less attention is that the choice for adjudication is not simply a binary one between trials or plea bargaining. There is much room for variation in the specific rules surrounding plea bargaining, including those that define prosecutor charging authority and the parties’ disclosure obligations. Most criminal justice systems in Europe and elsewhere now authorize and widely practice some form of plea bargain-

---

4 For a sample of the debates, see Malcolm M. Feeley, The Process is the Punishment: Handling Cases in a Lower Criminal Court 244–77 (1992) (arguing against caseload pressure as explanation for bargaining); Fisher, supra note 3 (attributing bargaining to civil and criminal caseload pressures and the self-interests of prosecutors and judges); Lawrence M. Friedman & Robert V. Percival, The Roots of Justice: Crime and Punishment in Alameda County, California 1870–1910, at 192–95 (1981) (attributing bargaining to shifting focus from the crime to the criminal and the increased involvement of police and prosecutors in adjudication); Milton Heumann, Plea Bargaining: The Experiences of Prosecutors, Judges, and Defense Attorneys 156–57 (1978); McConville & Mirsky, supra note 2, at 333–37 (attributing bargaining to changes in macro-political economy and emerging state interests in social control); Vogel, Social Origins of Plea Bargaining, supra note 3, at 162–66 (explaining bargaining through changes in political economy).

5 See, e.g., Santobello v. New York, 404 U.S. 257, 260 (1971) (“[Plea bargaining] is an essential component of the administration of justice. . . . 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.”).
ing. Rules elsewhere vary substantially from U.S. practice, however, including those in other common law nations. England, for example, regulates plea bargain sentences in relation to trial sentences, limiting the plea discount to a maximum of one-third less than a post-trial sentence—a policy that in theory compromises efficiency (by limiting incentives for guilty pleas) in order to guard against undue pressure on defendants to plead guilty. While the federal sentencing guidelines fix a comparable discount for “acceptance of responsibility,” the size of plea discounts in federal as well as state practice is in fact effectively unregulated, because no law meaningfully limits prosecutors’ discretion to add or dismiss charges depending on a defendant’s willingness to plead guilty. Defendants can waive appellate review of convictions based on guilty pleas in American courts (save in the military justice system); not so in Germany. Disclosure rules differ as well. In contrast to the rules in Canada that require pre-trial disclosure (as well as those in civil law jurisdictions where defendants have pre-trial access to the state’s full evidence file), the Supreme Court has also held that prosecutors’ duty to


7 On England’s limitation of discounts for plea bargains, see Criminal Justice Act, 2003, c. 44, §§ 144, 172 (U.K.); Sentencing Guidelines Council, Reduction in Sentence for a Guilty Plea (2007); Attorney General’s Reference Nos. 14 & 15 (Tanya French & Alan Webster), [2006] EWCA (Crim) 1335 (appeal taken from Eng.) (noting that “[g]uidelines do no more than provide guidance” to judges and “[t]here may well be circumstances which justify awarding less than a discount of one third where a plea of guilty has been made at the first opportunity”).


9 Turner, supra note 6, at 95–97 (translating Bundesgerichtshof [BGH] [Federal Court of Justice] Mar. 3, 2005 (Ger.)).

disclose exculpatory evidence to defendants is a trial right that defendants can be encouraged to waive as part of a plea bargain.\textsuperscript{11} The Court has also approved convictions based on guilty pleas even for defendants who refuse to admit their factual guilt, although it has never defined the standard of proof that trial judges should apply when assessing the factual basis for guilty pleas.\textsuperscript{12} Statutory rules add little to trial judges’ duties regarding the factual basis for the criminal judgments they enter based on pleas.\textsuperscript{13}

In a wide range of rules and doctrines such as these, the law of criminal adjudication has opted not only to encourage the efficiency of negotiated criminal judgments over trials. In virtually every instance, the choice is for rules that reduce the process costs of reaching judgments, despite the harm to other interests. The most familiar interests include those served by jury decision-making, and by the defendant’s (or victim’s) participation in the public trial process.\textsuperscript{14} But more subtly, the particular rules that define American plea bargaining also minimize judges’ role in—and responsibility for—the content of courts’ criminal judgments. Likewise they reduce judicial supervision of prosecutors and appellate court supervision of criminal process.

II. EFFICIENCY AND ITS CONSEQUENCES

Discussions of criminal process suffer from an inadequate understanding of what efficiency means in this context, and what consequences can follow from efficiency gains. Definitions of efficiency vary and are tailored to different uses and settings. Pareto efficiency, commonly employed in legal scholarship, defines a standard efficiency in allocation of goods.\textsuperscript{15} Maximum production efficiency, by contrast, is defined vari-

\textsuperscript{11} See United States v. Ruiz, 536 U.S. 622, 631–33 (2002) (describing prosecutor’s due process duty to disclose exculpatory and impeachment evidence as a trial right, and finding nondisclosure of impeachment evidence, based on defendant’s consent in a plea agreement, to be constitutional).


\textsuperscript{14} See Stephanos Bibas, The Machinery of Criminal Justice (2012) (describing and criticizing costs to defendants, victims, and public interests in prevailing plea-based adjudication).

\textsuperscript{15} See Pareto Efficiency, in A Dictionary of Economics 333 (John Black et al. eds., 3d ed. 2009) (“Pareto efficiency [is a] form of efficiency for an economic allocation. An allocation is Pareto efficient if there is no feasible reallocation that can raise the welfare of one economic agent without lowering the welfare of some other economic agent.”).
ously as producing goods at the lowest cost or “allocating the available resources between industries so that it would not be possible to produce more of some goods without producing less of any others.”\textsuperscript{16} Production efficiency is closest to the idea invoked in adjudication. Efficiency in this sense is a description of the ratio between the resources required for a productive activity—such as resolving court cases—and the outputs of that activity.\textsuperscript{17}

Some fields of study pay closer attention to the broad set of effects that can follow from efficiency gains in specific settings. Research on energy efficiency is one example, and I use it as an analogy below. A gain in efficiency results from a comparison of the ratio of resources to output from two different times or settings. An increase in efficiency is a decrease in the amount of resources required to produce a unit of a good or service, from one time (or mode of production) compared to another. Thus a court that produces final judgments in five cases per day can be deemed more productive than a court that produces two judgments daily with the same staff. In the same sense, we say that a car’s fuel efficiency improves if its capacity changes from 30 miles-per-gallon (“mpg”) to 35 mpg. Further, this sort of efficiency gain can be described in at least two ways, depending on which factor (distance or fuel) serves as the baseline. One measure focuses on the reduction in fuel required for a given activity, for example, driving a car 30 miles. For that activity, required fuel drops from one gallon to about 0.85 of a gallon. Alternately, the efficiency gain can be described as an increase in activity (or production) with the amount of fuel held constant. With one gallon of fuel, the car now travels 35 miles instead of 30. In this sense, efficiency is interchangeable with productivity; a system that achieves a productivity gain

\textsuperscript{16} Efficiency, \textit{in} A Dictionary of Economics, supra note 15, at 134; cf. Productive Efficiency, \textit{in} Dictionary of Economics 159 (A & C Black 2006) (defining productive efficiency as “a situation in which the most production is achieved from the resources available to the producer”).

\textsuperscript{17} \textit{A Dictionary of Economics}, supra note 15, at 134 provides the following definition of “efficiency”:

Obtaining the maximum output for given inputs. Efficiency in consumption means allocating goods between consumers so that it would not be possible by any reallocation to make some people better off without making anybody else worse off. Efficiency in production means allocating the available resources between industries so that it would not be possible to produce more of some goods without producing less of any others. Efficiency in the choice of goods to produce means choosing so that it would not be possible to change the set of goods so as to make some consumers better off without others becoming worse off.
can produce more goods (or travel more miles) with the same inputs as before.

These alternate descriptions of the same efficiency gain go to the heart of policy concerns in settings such as energy consumption—and also in adjudication policy. The alternate formulations reveal an important distinction in how gains from efficiency can be employed. Gains can be used to reduce resource consumption while maintaining the same desired production or activity level. That is generally the goal, of course, behind public policies (but not always private policies) to increase energy efficiency: to do the same things with less fuel. Alternately, we can also employ efficiency gains to maintain the same level of resource consumption while also increasing production. Instead of using less gas, we can use the same amount to drive more. To complicate matters, gains also can be taken in other ways: The same rate of fuel consumption can be used to travel the same number of miles as before, but the efficiency gain now allows for a qualitative improvement in the activity—a larger, heavier vehicle can now get 30 mpg when it formerly got 25. Manufacturers that improve their productive capacity commonly choose among such options. With a gain in efficiency, they can produce more units with the same resources as before, produce the same number while using fewer resources, or produce the same number but of better quality using the same resources. The critical point is the difference between greater efficiency and its consequences. An efficiency gain has nothing necessarily to do with how one takes advantage of that gain.

Real-world choices between these options can be complicated, because how one uses efficiency gains depends also on how much control one has over how they are utilized. Consider again a firm that produces a good with fewer resources than before. The firm may have some control over what it does with that efficiency gain, but it is restricted by market conditions. If no competitors have yet made the same improvement, a market might allow the firm to produce the same goods, sell them for the same price, and pocket the efficiency gain as higher profits. Or, if market demand is price-elastic, the firm can lower the price of the products and sell more units, which leads the firm to produce more. Increased

---

18 This follows the assumption of the standard demand function, according to which demand is determined by price and increases as the price of a good or service decreases, holding aside non-price determinants of demand. See generally Paul Krugman & Robin Wells, Microeconomics 62–77, 147–52 (2d ed. 2009) (offering an overview of demand and supply
production returns the firm’s resource consumption to its levels before the efficiency gain, or perhaps even beyond. For a typical firm, that does not matter; the firm’s goals in utilizing efficiency gains are defined in terms of profits. But if a policy goal were instead defined in terms of lowering resource consumption, that is a bad outcome. The firm’s efficiency gain does not necessarily translate into lower total resource consumption. Depending on conditions, resource consumption could remain constant or increase.

The same holds for any setting in which an efficiency gain can translate into a price reduction, including auto fuel efficiency. Automobile manufacturers that achieve fuel efficiency gains do not fully decide how those gains will be put to use. To a large degree, car drivers do (with variations depending on how the markets for cars and fuel are regulated). Drivers experience greater fuel efficiency as a price cut in the cost of fuel, and in the cost of driving a mile. Other things equal, price cuts increase demand, so drivers might choose more driving while spending the same on gas, instead of driving the same amount and allocating the savings elsewhere. But here too, those are not the only options. Automakers might allocate new fuel efficiency not to producing the same car with better mileage but instead to improving the quality of the driving experience in some sense, by producing a bigger or more powerful car that gets the same mileage. Some drivers will opt for bigger cars that require the same gas-per-mile as smaller cars formerly did, rather than sticking with smaller cars and saving money from fuel costs for other uses.

In these examples we see how effects of efficiency gains depend on a number of factors other than simply the gain itself. Further, even if the policy preference for a specific effect is clear—for example, lower resource consumption rather than increased production—achieving that goal is not straightforward in many contexts. Firms, as noted, are often constrained by market conditions—whether demand is price-elastic, whether competitors have matched their efficiency gains, and so on. Policymakers seeking to reduce carbon fuel usage through greater energy efficiency face challenges at least as great, because energy consumption depends on how consumers respond to efficiency gains; drivers may choose to drive more or drive bigger cars rather than spend less on fuel.
That is why raising fuel prices is a more effective public policy strategy for reducing fuel consumption than mandating efficiency gains; response to a price increase is more predictable than it is to an efficiency gain.

For these reasons, energy and environmental economists worry that efficiency gains in energy consumption for all sorts of technologies will sometimes lead to constant or greater resource usage rather than to their policy goal of lower usage. For technology users, efficiency gains amount to price cuts in the benefits that a technology provides. On the model of a standard demand function, price reductions increase demand for a good or service, at least when non-price determinants of demand do not cut the other way. 19 This consequence of efficiency gains—boosting consumption rather than cutting it—is sometimes called a rebound effect, or a Jevons effect. 20 It is what a manufacturer hopes to achieve by improving productivity and cutting prices: an increase in demand. But if one’s goal is instead to translate efficiency gains into lower resource consumption, an efficiency gain that triggers more consumption and greater resource usage is perverse. 21 Nonetheless, when the risk is real—and when the policy goal is defined as reduced resource usage rather than increased production or supply—it makes efficiency gains less attractive. Depending on one’s goal, then, inefficiency—or its equivalent, keeping the price of a good higher—can be beneficial. The better strategy could be to avoid efficiency gains as a means to keep costs higher and thereby discourage greater demand, just as a gas tax raises fuel costs to discourage consumption.

19 See generally Krugman & Wells, supra note 18, at 62–71 (describing factors, including non-price factors, affecting demand).

20 On Jevons or rebound effects, see Blake Alcott, Jevons’ Paradox, 54 Ecological Econ. 9 (2005) (explaining the theoretical background and surveying modern debates about the Jevons effect); Kenneth A. Small & Kurt Van Dender, Fuel Efficiency and Motor Vehicle Travel: The Declining Rebound Effect, 28 Energy J., no. 1, 2007 at 25. The original argument for rebound effects (since rejected in the specific context) is found in W. Stanley Jevons, The Coal Question; An Inquiry Concerning the Progress of the Nation, and the Probable Exhaustion of our Coal-Mines 122–37 (London, MacMillan & Co., 2d ed. 1866).

21 With respect to energy usage, the consensus seems to be that rebound effects are real but in most settings do not fully offset reductions from increased efficiency. For debates about rebound effects in energy contexts, see Lee Schipper & Michael Grubb, On the Rebound? Feedback Between Energy Intensities and Energy Uses in IEA Countries, 28 Energy Pol’y 367 (2000) (special issue on rebound effects); see also John M. Polimeni & Raluca Iorgulescu Polimeni, Jevons’ Paradox and the Myth of Technological Liberation, 3 Ecological Complexity 344 (2006); James Barrett, Rebounds Gone Wild, Nat’l Geographic (Dec. 20, 2010), http://www.greatenergychallengeblog.com/2010/12/rebounds-gone-wild/.
III. EFFICIENCY IN ADJUDICATION

A. The Ambiguous Relationship of Criminal Offending and Caseloads

With this framework, we can see how efficiency gains in criminal adjudication might create the same sorts of complications. Criminal justice officials have many options for efficiency improvements. Most simply and importantly, they achieve more efficient case processing by substituting plea bargains for trials. Understanding adjudication simply as the productive capacity of judges and prosecutors enables one to conclude that officials who process—or resolve—more cases in less time (producing, say, judgments in ten cases a day instead of five) are more efficient. We can now recognize the options created by this more efficient practice. Officials might produce twice as many case resolutions as before, doubling their production while using the same resources—that is, working the same number of days. Alternately, they could keep the production rate steady, at five judgments per day and then use the resource savings (their work time) for other tasks: Judges might try more civil cases; prosecutors might more carefully supervise police and screen evidence. Still another option is to enjoy the efficiency gain outside the court system; policymakers might reduce the number of judges and prosecutors and spend the savings on other public projects. Some mix of these options is also possible.22

The common understanding, especially among courts, practitioners, and policymakers, is that efficiency gains in criminal process are driven by exogenous increases in demand for enforcement. Put more simply, caseloads increase because more offending occurs, or because better en-
forcement catches a larger percentage of offenders. There is some empirical evidence on crime rates in recent decades to support this view, as well as historical evidence from earlier eras, although it is not unambiguous. On this view, caseload pressures force courts and prosecutors to find ways to adjudicate more efficiently and to take those efficiency gains in the first form—by increasing the supply of adjudication, or the production of convictions, to meet increased demand. The goal of judges and prosecutors who find ways to process cases more quickly is not to handle the same number of cases and allocate the savings elsewhere; the goal (and need) is to process more cases.

The question is whether caseload increases are in fact exogenous to adjudication. With a better understanding of efficiency’s effects, we can see the possibility that they are not. Criminal prosecutions are a variable that may be partially dependent on adjudicative capacity. If so, efficiency gains in some part contribute to the rise in caseloads, rather than the rise in caseloads creating a need for greater efficiency. Rising caseloads can be partly a consequence of more plea bargaining—a rebound effect—rather than a cause. Criminal charging may be a response to the reduced costs of court judgments as much as it is a function of crime rates. Do we plea bargain because we have more cases, or do we have more cases because we plea bargain?

B. Discretionary Policies Affecting Caseloads

The question is no easier to answer here than in the context of energy conservation. But it is easier at least to understand why we overlook the question and assume that caseloads mostly respond to rates of offending coupled with changing enforcement priorities. One reason for the as-

See, e.g., Lafler v. Cooper, 132 S. Ct. 1376, 1397 (2012) (Scalia, J., dissenting) (describing plea bargaining as a “necessary evil”); Blackledge v. Allison, 431 U.S. 63, 71 (1977) (“[T]he fact is that the guilty plea and the . . . plea bargain are important components of this country’s criminal justice system. . . . Judges and prosecutors conserve vital and scarce resources.”); Santobello v. New York, 404 U.S. 257, 260 (1971) (describing plea bargaining as “an essential component of the administration of justice” that prevents the “need to multiply by many times the number of judges and court facilities”); Ronald Wright & Marc Miller, The Screening/Bargaining Tradeoff, 55 Stan. L. Rev. 29, 40 (2002) (describing caseload pressures as the “primary engine behind the shift from trials to plea bargaining”).

sumption arises from the sorts of offenses that get primary attention when we think about crime rates. Attention mostly centers on “core” crimes (or FBI index crimes) such as homicides, robberies, assaults, and property theft, as well as drug-related offenses. Some drug crimes aside, these offenses are at the heart of traditional concerns about social order, public safety, and property rights. Their harm and wrongful nature are unambiguous. Also, the data on core offenses are generally good, especially for homicides. Few homicides go undetected, even if many go unsolved, so we can be confident when data tell us these crimes are increasing. Prosecutions of such core offenses do not seem like a function of government policy. If the government has sufficient evidence of an offender’s guilt, a prosecution is almost certain to follow.

But these assumptions do not hold for many other kinds of offenses that play a large role in criminal enforcement practice. We know, for one, that criminal charges are not certain to follow from the state’s possession of sufficient evidence. American prosecutors, following common law tradition, have vast charging discretion, and sometimes have explicit declination policies. Studies reveal much declination in practice, though patterns vary greatly across offices and types of offenses. Further, whether the state has sufficient evidence depends a lot on whether police have made detection and investigation a priority, and whether they have the resources to do so. Not all crimes are as readily reported or easily detected as homicides, robberies, and car theft. This is especially so for the large number of crimes that cover consensual conduct such as production, distribution, and use of contraband—illicit drugs, alcohol, firearms—as well as offenses ranging from prostitution and intoxicated driving to conspiracy and joint crimes of preparation.

25 For change in rates of crime over the last half century, see 5 Historical Statistics of the United States: Earliest Times to the Present 5-224 tbl.Ec11-20 (2006) (listing estimated rates per 100,000 population of crime known to police from 1960–1997, and showing an increase in all categories of crime until the 1990s, when rates start to decline); see also U.S. Dep’t of Justice, Bureau of Justice Statistics, Sourcebook of Criminal Justice Statistics - 1998, at 260 tbl.3.114 (Kathleen Maguire & Ann L. Pastore eds., 1999).

As to all these offenses, the links between commission and charging are much more contingent. Actual rates of offending are harder to determine, because knowledge of offenses depends heavily on how much effort police put into looking for them. Uncovering and prosecuting many offenses depends on choices by policymakers and enforcement officials about how much to invest in doing so. Examples of these sorts of policy choices are widely known; many are touted by policymakers. The “war on drugs” is a large-scale example, and also one with varying enforcement strategies. The federal government has at times announced policies of not prosecuting juvenile illegal immigrants\(^27\) or those who violate federal marijuana offenses for conduct that is legal under state law. Some city police forces (for example, Seattle) have made marijuana possession a low enforcement priority, while others (for example, New York City) have made it a high one.\(^28\) The rise of drug courts adds more variation in how cases are resolved after charging. Other examples abound. The Justice Department has, to varying degrees, made child pornography, corporate fraud, and terrorism-related offenses high priorities for prosecutors and investigators. State and federal officials have cooperated in many districts to seek harsher punishments for illegal firearm possession. Both, on the other hand, have experimented with strategies such as the High Point Drug Initiative that negotiate with gang


members and other suspects to prevent offending in exchange for job assistance and non-prosecution of past offenses.\(^{29}\)

The discretion that affects caseloads goes deeper as well, to legislative decisions about criminalization and understandings of what constitutes wrongdoing and risk of harm. The history of vice and moral crimes is perhaps the clearest example. Temperance and prohibition movements played a large role in the nineteenth and early twentieth centuries in defining alcohol use and abuse as harmful, thus justifying their criminalization and prosecution. Alcohol-related criminal caseloads rose and fell along with the political success of temperance organizations, which influenced both enforcement efforts and what alcohol-related conduct was defined as criminal. The same is true for prostitution; criminalization varies widely among jurisdictions (though very modestly within the United States), as do enforcement policies (even within the United States). Crimes of publishing and possessing sexually explicit texts and images were once widespread but now rare (child pornography aside). Domestic abuse, especially by husbands against wives and children, was once shielded from criminalization by norms and legal doctrines of privacy, at times was addressed by civil family courts, and now is commonly a high priority for prosecution. Longstanding crimes of usury have been repealed in many states as views shifted from ones that understood high-interest loans as harmful and immoral to ones that see such loans as voluntary and beneficial to borrowers as well as lenders.\(^{30}\) The list goes on.

---


C. Discretion over Criminal Caseloads and the Effects of Adjudicative Efficiency

Once one recognizes this flexibility in criminalization and enforcement policy, it is easy to see that the connection between criminal offending and caseloads is far from straightforward. That recognition makes the link between caseloads and adjudicative efficiency more problematic as well. Legislatures, police, and prosecutors (and ultimately, to some degree, public opinion) exercise a lot of discretion in determining caseloads; the number of prosecutions is not simply a direct function of the rates of criminal offending in the world outside the courtroom. The discretion to change the number of crimes by legislating crime definitions, uncovering more with greater policing investments, or addressing some violations with policies other than criminal prosecution—all of this opens the possibility for policymakers to make these decisions in response to changes in the price of adjudication, which changes with gains in efficiency. Policymakers and enforcement officials, then, are in a position something like car drivers in the wake of new models that incorporate fuel efficiency improvements. They have options for how to respond to a more efficient adjudication regime.

Given this variety of options and real changes in offending rates, it is hard to estimate how much charging is a response to cheaper adjudication. Prosecutors could simply charge more offenses and more defendants than before. The choices among possible charges may also change: Cheaper adjudication makes it less costly to charge crimes that are more difficult to prove. Some offenses are harder to prove because they include elements that require costlier proof efforts; it is easier to prove possession than sale of drugs, and easier to prove strict liability offenses than crimes with mental-state requirements for every element. Offenses can also be harder to prosecute because the available evidence is better in some cases than others. Prosecutors can reach deeper into the pool of potential cases generated by police as adjudication becomes cheaper, pursuing a greater portion of marginal or weak cases that they would decline if processing costs were higher. Police, in turn, may generate more of these types of cases. In all these ways, more efficient adjudication leads to more criminal law enforcement, even if rates of offending are steady.

That does not necessarily mean better or more optimal enforcement. Pursuit of weaker or more-difficult-to-prove cases may mean more acquittals, or more wrongful convictions. More generally, too much en-
forcement, even of uncontroversial laws, can be detrimental, as legislatures recognize when they deliberately restrict enforcement officials’ capacity to vigorously enforce some offenses. Congress specifically restricts the executive branch’s ability to go after certain politically salient offenses, such as those defined by certain firearms and tax laws. And enforcement officials themselves recognize, in a wide variety of contexts ranging from environmental compliance to neighborhood-based illicit drug markets, that civil or cooperative strategies rather than strict criminal punishment can better reduce rates of offending.

Nearly all of those responses to efficiency gains are at the level of enforcement practice. But the effects of those gains can also trigger responses by legislatures. Broadly speaking, as greater efficiency reduces the overall cost of criminal law enforcement, it makes it less costly for legislatures to create new offenses, and more tempting to choose criminal enforcement over other public policy strategies to address social problems or regulatory agendas. Adding new offenses to criminal codes is cheap, but funding their enforcement is not. Yet more efficient adjudication reduces the effective level of “per unit,” or per-offense, spending on prosecutors, courts, and defenders. That amounts to an incentive for legislatures to expand the types of conduct or social harms that they


33 Professor Bill Stuntz did the most to develop this insight, and the recognition that adding certain kinds of new offenses can also sometimes lower enforcement costs. See Stuntz, Constitution of Criminal Justice, supra note 2, at 782, 810; William J. Stuntz, The Pathological Politics of Criminal Law, 100 Mich. L. Rev. 505, 529–40 (2001).
This is hard to confirm with any precision. But the story of criminal law’s reach and of criminal adjudication practice over two centuries is, broadly speaking, at least as consistent with this account as it is with a thesis that rising rates of harmful conduct forced heavy caseloads onto courts that they were forced to accommodate. The nineteenth century saw the rise of professional prosecutors’ offices and police forces and also of regulatory (or malum prohibitum) offenses. The growth of population, cities, commerce, and industrialization explains much of the rise of regulatory law, but it does not clearly explain why so much regulation in the United States took the form of criminal rather than civil sanctions. The United States lagged behind European nations in developing a modern bureaucratic and administrative government infrastructure, which may help explain why states chose regulatory forms that could be enforced by prosecutors and criminal courts. But it is also clear that plea bargaining in some U.S. localities goes back two centuries and was closely linked to new regulatory offenses. Plea bargaining in Middlesex County, Massachusetts, for example, first arose in the context of new offenses regulating alcohol production and distribution. Social harms stemming from expanded alcohol production—criminal conduct, once it was defined as a crime—may have put real pressure on local officials and courts. But it could also have been that new efficiencies in

34 Scholarship on American political development debates whether the United States is properly characterized as a “weak state,” especially compared to the Weberian model of European states; recent scholarship challenges that characterization and describes distinctive forms of significant government authority. See generally Brian Balogh, A Government Out of Sight: The Mystery of National Authority in Nineteenth-Century America (2009); Karen Orren & Stephen Skowronek, The Search for American Political Development (2004); Stephen Skowronek, Building a New American State: The Expansion of National Administrative Capacities, 1877–1920 (1982) (arguing that pre-established institutional arrangements frame state development); Elisabeth S. Clemens, Lineages of the Rube Goldberg State: Building and Blurring Public Programs, 1900–1940, in Rethinking Political Institutions: The Art of the State 187, 189 (Ian Shapiro et al. eds., 2006) (“[T]he problem of indirect or delegated governance is addressed from the vantage point of state governments during an era when many political actors favored construction of the ‘monocratic bureaucracies’ analyzed by Weber...”); William J. Novak, The Myth of the “Weak” American State, 113 Am. Hist. Rev. 752 (2008) (critiquing the use of exclusively European models of state development and asserting that the American state structure is not “weak” in comparison to the centralized bureaucracies of Europe).

dispensing punishments, achieved through plea bargaining, made criminal regulation a more appealing policy option for responding to such social and commercial activity that posed new threats to public order.

The same ambiguity of cause-and-effect characterizes the expansion of temperance and anti-vice agendas throughout the nineteenth and early twentieth centuries. As genuine or perceived social problems grew, criminal adjudication’s ability to reduce its per-case costs provided an additional incentive to address such problems by alcohol or obscenity prohibition administered through criminal law. Guilty pleas skyrocketed in the 1920s as alcohol prosecutions flooded Prohibition-era courts. That can be read as caseload pressure forcing changes in adjudicative practice. But it is also true that adjudication’s greater efficiency made criminal law a more feasible policy strategy for Prohibition; efficiency gains enabled legislators to expand crime definitions and criminal punishment without proportionate increases in funding for courts and prosecutors. The same account applies to the broad use of criminal law in the nineteenth and early twentieth centuries for other regulatory, morals-driven, and social-order policies implemented through criminal law: from longstanding crimes of public drunkenness, vagrancy, unemployment, and prostitution, to those on topics such as fire- and building-safety codes, operating dance halls, bans on storefront shops for securities trading (“bucket shops”), and various environmental regulations.

Data for recent decades support the same inference for contemporary policy in the era of mass incarceration. Although guilty pleas resolved more cases than did trials long before the 1970s, U.S. jurisdictions nonetheless managed to increase plea bargaining rates since that decade.

---

36 Fisher, supra note 3, at 8 (noting that in the Prohibition-era, “the nation’s courts sank beneath the sheer weight of liquor cases”); id. at 4–8, 21–27 (discussing alcohol prosecutions in other eras); Daniel Okrent, Last Call: The Rise and Fall of Prohibition 112, 264 (2010) (describing caseload strains on courts and guilty plea increases during Prohibition).

37 For discussions and examples of such regulatory statutes in the nineteenth and twentieth centuries, see Booth v. Illinois, 184 U.S. 425 (1902) (affirming conviction under a state law criminalizing options contracts on commodities); Davis v. Beason, 133 U.S. 333, 341 (1890) (“Bigamy and polygamy are crimes by the laws of all civilized and Christian countries. . . . They tend to destroy the purity of the marriage relation, to disturb the peace of families, to degrade woman and to debase man. Few crimes are more pernicious to the best interests of society and receive . . . more deserved punishment.”); Novak, supra note 35, 1–18 (describing pervasiveness of nineteenth-century regulatory crimes); Donna I. Dennis, Obscenity Law and its Consequences in Mid-Nineteenth-Century America, 16 Colum. J. Gender & L. 43 (2007); Louis Henkin, Morals and the Constitution: The Sin of Obscenity, 63 Colum. L. Rev. 391, 411–14 (1963).
Federal court guilty pleas as a portion of all convictions rose from 86% in 1970 to 97% in 2009. In the 1990s, the federal government instituted “fast-track” plea bargaining policies designed to further reduce the time required to resolve cases by guilty pleas. For the category of cases in which they did so first and most consistently—immigration-related crimes—plea rates rose to 99.4% by 2010. Notably, from 2006 to 2010, prosecutions for illegal entry/re-entry offenses increased, and immigration-related crimes grew from 25% to 36% of federal cases. At the same time, rates of unauthorized migration—that is, the exogenous crime rate—decreased. Discretionary enforcement policies, rather than raw numbers of offenses, drove up caseloads, and quicker guilty-plea procedures made those policy priorities easier and cheaper to implement.

State court data is less detailed but consistent, and more plea bargaining is the most likely explanation for prosecutors’ ability to take on growing caseloads over the last forty years—or, put differently, for prosecutors’ greater efficiency and productivity. From 1974 to 2005, the estimated number of local prosecutors in the United States grew from about 17,000 to approximately 27,000, a 59% increase. Over the same

---


41 On the decline in unauthorized migration up to 2010, see Jeffrey S. Passel & D’Vera Cohn, U.S. Unauthorized Immigration Flows Are Down Sharply Since Mid-Decade, at i (2010). For an excellent overview discussing these developments, see Fan, supra note 39.
period—during which the incarceration rate quadrupled—felony prose-
cutions tripled from roughly 300,000 a year to more than one million.42
(Violent crime rates rose during the first half of this period and declined
for the second.) As the unit cost of prosecutions declined, the number of
prosecutions increased. Without that efficiency gain, criminal law would
have been more difficult—because more costly—to deploy so expa-
sively.

Whether that impediment to expansion of criminal law enforcement43
would have led to better or worse policy is a contentious debate, though
the burden would seem now to be on defenders of contemporary incar-
ceration policies. The correlation of this efficiency gain with caseload
increases does not tell us in which direction causation runs. But to rule
out any role for process efficiency in causing higher caseload is to deny
the operation of the basic demand function in this setting—the premise
that demand is a function of price. By contrast, if the demand function
holds in this context, it supports the inference that a lower price for con-
victions leads to greater demand for them.

More efficient adjudication can also affect the size and complexity of
offenses in the criminal justice system. Long-term growth in corporate
and white-collar crime is primarily a product of expansions of the corpo-
rates and in the breadth and sophistication of commercial activity
over the last century. But harmful conduct in these settings can be ad-
ressed through criminal rather than through civil regulatory law in part
because criminal adjudication has been able to adapt to the challenges of
large-scale cases governed by complex statutes with vast evidentiary

(providing these figures and sources for them, primarily data from the U.S. Bureau of Justice
Statistics and from the National Center for State Courts). The U.S. incarceration rate rose
from 104 to 492 per 100,000 between 1974 and 2005. See Margaret Werner Cahalan, Histori-
6 tbl.6 (2012) (reporting incarceration rates for 2000–2011). For data on rising crime rates
until the early 1990s and declining rates thereafter, see Steven D. Levitt, Understanding Why
Crime Fell in the 1990s: Four Factors that Explain the Decline and Six that Do Not, 18 J.
Econ. Persp. 163, 165 fig.1 & 166 tbl.2 (2004); U.S. Dep’t of Justice, Crime in the United
43 See generally Markus Dirk Dubber, Policing Possession: The War on Crime and the End
of Criminal Law, 91 J. Crim. L. & Criminology 829 (2001) (analyzing the success of the war
on crime and arguing that arrests for possession have replaced vagrancy laws as law en-
forcement’s main tool for social control).
records.\textsuperscript{44} In the traditional trial process, these cases are very costly to litigate, as we see from the few that get that far. But nearly all settle without trial, sometimes before indictment, and often early enough to reduce the burdens of evidence gathering.\textsuperscript{45} This ability of criminal adjudication to adapt to the challenges posed by complex corporate activity may be an added reason why the United States uses criminal sanctions more than other countries, which instead rely on civil or administrative mechanisms.\textsuperscript{46} Criminal adjudication’s efficiency in such contexts incentivizes legislators to adopt regulations backed by criminal rather than civil sanctions. Even if criminal penalties serve only as leverage for civil settlements, as is often the case in federal enforcement practice, cheaper adjudication increases the leverage of the criminal alternative.

Finally, legislatures’ reliance on efficient adjudication can also show up in decisions to fund police and other investigative-enforcement agencies. As cheaper adjudication creates incentives for more prosecutions, prosecutors can either pressure police for more arrests and evidence files or initiate charges in a larger share of the cases police already send to them.\textsuperscript{47} That, in turn, might lead legislatures to increase funding for po-

\textsuperscript{44} This is so even though federal enforcement officials frequently resolve wrongdoing that could be prosecuted criminally through settlements for civil fines and regulatory remedies. See Darryl K. Brown, Street Crime, Corporate Crime, and the Contingency of Criminal Liability, 149 U. Pa. L. Rev. 1295, 1323–44 (2001) (describing some effects of greater availability of civil remedies for criminal wrongdoing in corporate and white-collar settings); Brandon L. Garrett, Globalized Corporate Prosecutions, 97 Va. L. Rev. 1775, 1800–14 (2011) (describing the mix of guilty pleas and civil settlements used to settle corporate crime cases).

\textsuperscript{45} See generally Brandon L. Garrett, Aggregation in Criminal Law, 95 Calif. L. Rev. 383, 385 (2007) (“[T]he appropriate use of aggregation can potentially transform criminal adjudication[,] by providing an avenue to vindicate criminal procedure rights, and by encouraging efforts to create a more efficient, accurate, and fair criminal justice system.”); Garrett, supra note 44 (describing settlements in corporate prosecutions); Brandon L. Garrett, Structural Reform Prosecution, 93 Va. L. Rev. 853, 886–902 (2007) (describing corporate reform and compliance monitoring components common in U.S. Justice Department settlements with firms for criminal wrongdoing).


\textsuperscript{47} Interestingly, Professor John Pfaff’s research suggests that the latter explains most of the rise in U.S. incarceration rates in the last two decades—that is, that prosecutors are charging a greater percentage of arrestees, rather than the police making more arrests. See John F. Pfaff, The Micro and Macro Causes of Prison Growth, 28 Ga. St. U. L. Rev. 1239, 1242–55 (2012); John F. Pfaff, The Myths and Realities of Correctional Severity: Evidence from the National Corrections Reporting Program on Sentencing Practices, 13 Am. L. & Econ. Rev. 491, 518–19 (2011) (arguing that admission practices rather than longer sentences are driv-
lice and related enforcement activity. Alternately (or in addition), it might prompt investigative agencies to change tactics. Police can generate more arrests without more resources by increasing their own efficiency; enforcement officials can respond to the greater adjudicative capacity just as prosecutors and legislatures do. At one end of the spectrum, federal agencies can pursue more cases of complex corporate wrongdoing, knowing that investigative and adjudicative costs are lower with settlements achieved in the shadow of formal adjudication. At the other end—local policing and routine state criminal law—police often can reallocate existing resources to increase their “production” of certain kinds of arrests that have lower investigative costs—marijuana possession, disorderly conduct, alcohol-use offenses, even street-based drug sales—in response to prosecutors’ and courts’ greater adjudicative capacity. New York City police have demonstrated the ease with which they can increase stop-and-frisk searches of pedestrians and thereby ratchet up drug- and weapon-possession charges.48 Sting operations against street drug sellers are cheaper to perform than those targeting buyers, and focusing on street markets for drug offenses is cheaper than investigating the same activity in homes or offices.49 All other things being equal, demand for more enforcement encourages a familiar form of investigative efficiency—a focus on low-cost arrests rather than on higher-cost investigations or on prevention-oriented policies less likely to lead to arrests. Police can have multiple reasons to favor easier rather than harder targets, but adjudicative efficiency is one factor that encourages that focus. In doing so, it encourages a policing strategy that often aggravates racial and class disparities.50
None of this means that faster and cheaper adjudication is the primary cause of new offenses and enforcement agendas. Various policy concerns and political decisions contribute to criminal law enforcement agendas. Social conditions change and new social problems arise. Policymakers and public pressure respond to myriad influences and developments: immigration-driven population growth and ethnic dissension; alcohol customs that vary among groups and that had to adapt to the rise of industrial work settings and automobile culture; the advent of powerful illicit drugs; the expanded capacity of firms to cause widespread financial or environmental harms. The degree to which and the form in which criminal law plays a role in managing these challenges to social order are deeply political.

Nonetheless, costs play a role in legislative choices about criminalization as well as in police and prosecutor priorities, and efficiency gains lower costs. One way to put this is that plea bargaining subsidizes criminal law enforcement, compared to full trials as the baseline for achieving convictions. The efficiency subsidy to criminal law enforcement cannot be separated, even conceptually, from other reasons that affect decisions about what to criminalize, investigate, and charge. Costs, efficacy, moral or social appropriateness—all are intertwined concerns in these discretionary choices. By lowering enforcement costs, more efficient adjudication tilts the policymaking scales in favor of more enforcement and punishment. In this way, efficient criminal process expands the state’s enforcement capacity. In the context of American criminal justice policy over the last generation, this suggests that adjudicative efficiency contributed to the nation’s unprecedented incarceration rates. How police and prisons reap gains from adjudicative efficiency is the focus of the next Part.

IV. DISTRIBUTIVE IMPLICATIONS OF ADJUDICATION EFFICIENCY

Thus far we have seen how improving criminal process efficiency can affect the number and type of criminal cases that are sent to courts by
making case resolutions cheaper. But the public costs of criminal cases include much more than adjudication; most obviously, they include policing and investigation costs that precede charging, and the punishment expenditures that follow conviction. Yet adjudicative efficiency gains can alter the distribution of public expenditures across these various sectors. Lowering the per-unit adjudication cost of convictions reduces the proportional cost of adjudication—judges, prosecutors, and defense lawyers—relative to the remaining, non-adjudication parts of the criminal justice system. Imagine a simple pie chart of criminal justice spending divided into three slices—one each for policing, adjudication, and punishment. The adjudication slice shrinks, relative to the other two, with the innovation of quicker processes for reaching judgments. This relative shift occurs so long as police and prisons do not match adjudication’s gains with their own efficiency improvements. The prospect of gains is more feasible for policing, which can benefit from new technologies. For incarceration, the opportunities for reducing costs of humane detention are inherently more limited.

Gains from efficiency, we have noted, can be used in various ways. The most common explanation for plea bargaining is a story of efficiency gains used to process more cases with the same level of resources. There is no evidence that efficiencies instead allowed for reductions in court and prosecution staffs—hardly a surprise given the nature of bureaucracies. But what about transferring those gains outside of the adjudication system to other parts of the criminal justice system? Efficiency gains leave legislators with a surplus to allocate; they can spend those gains entirely on processing higher caseloads, but the savings can also go elsewhere within the criminal justice system, or even outside of criminal justice altogether.

With this recognition, we can see how adjudicative efficiency could help to subsidize higher spending on incarceration, the sector of criminal justice spending that has grown the most in the last quarter century (faster, in fact, than nearly every other program in state budgets over this pe-

---

52 This is a problem even for large private firms, but it is an especially familiar one for the public sector with no competitive market pressures to force downsizing. See generally James Q. Wilson, Bureaucracy: What Government Agencies Do and Why They Do It (1989) (discussing the behavior and incentives of public and private organizations, for example, to maintain constituencies and budgets).
The broad figures for criminal justice spending over the last generation are consistent with this interpretation that savings from adjudication have been transferred to budgets for prisons. Put differently, efficiency gains from more and faster plea bargaining have enabled legislatures not only to increase prosecutions and convictions, but to redistribute some criminal justice spending from adjudication to incarceration.

Consider the data on criminal justice spending during this period. We noted above that spending on prosecutors per case effectively declined from 1975 to 2005 as felony prosecutions tripled and prosecution staffs did not even double. The U.S. Bureau of Justice Statistics estimates that, between 1982 and 2001, total U.S. government spending on corrections (prisons, jails, probation services) rose by 400%, which enabled the U.S. incarceration rate to rise from 170 to almost 500 individuals per 100,000 members of the population (and over 500 by 2006). Spending on “judicial and legal services” (that is, on court officials and prosecutors) rose less, by 288%—recall that plea bargaining intensified in this period, enabling each prosecutor and judge to produce more judgments. The same pattern holds solely for federal criminal justice spending, and for data available through the year 2007. So, total criminal justice

---

53 Pew Center on the States, One in 31: The Long Reach of American Corrections 1 (2009), http://www.pewstates.org/uploadedFiles/PCS_Assets/2009/PSPPP_fin31_report_FINAL_WEB_3-26-09.pdf (noting that in recent years corrections spending “was the fastest expanding major segment of state budgets, and over the past two decades, its growth as a share of state expenditures has been second only to Medicaid”).

54 See supra note 42 and accompanying text.


57 Bauer & Owens, supra note 55, at 2. Spending on police increased at the slowest rate, by 202% for 1982–2001, but started from a much higher per-capita funding level, more than double the level of spending on either corrections or courts and prosecutors. Id. Regarding federal spending alone in this period, justice personnel spending went up 636%, while corrections spending went up 861%. Id. at 3; see also id. at 8–9 (defining personnel and institutions in each category). For raw data on the same budget trends for these three categories ex-
spending grew; the United States spent much more for criminal enforcement in the 2000s than three decades earlier. But spending also shifted somewhat—adjudication’s relative share of the criminal enforcement budget shrank while punishment’s share grew. Adjudication’s efficiency gains made it cheaper to prosecute more cases, but it also subsidized, in effect, some of the growth in incarceration budgets.

This broad-brush picture overlooks some complications in criminal justice funding structures in the states. Typically, state legislatures fund prisons from state general funds, but judges, prosecutors, and police are funded (with varying state-level contributions) mostly from local budgets defined by city, county, or judicial-district boundaries. Thus, a single decision-maker such as the legislature does not have full control over funding in all sectors with the power to move funds from one to another. Still, the systemic response to adjudicative efficiency can work much as if one did, in part because of a moral hazard problem created by these funding distinctions. When local prosecutors and courts improve their adjudicative efficiency, they can increase convictions without increasing local adjudication budgets or paying for the increased costs of incarceration. Instead, the convictions and sentences they generate put pressure on state legislators to increase prison budgets. For three decades ending perhaps four years ago, this seems to be exactly what happened. Prosecutors charged a larger percentage of offenders in the pool of arrestee case files that police provided to them; local courts produced more convictions; and legislatures increased spending on prisons.


60 See sources cited supra note 47 (analyzing arrest and charging data to find that prosecutors increased their rate of felony-charge-filing-per-arrest in recent decades, and this increased charging rate, rather than an increase in sentences, seems to account for most of the rise in U.S. incarceration rates).

61 Carson & Sabol, supra note 42 (reporting that total U.S. prison populations decreased in 2010 and 2011 for the first time in decades); see also State Expenditure Report, supra note 58, at 52 (reporting that total state spending on corrections grew in FYs 2011 and 2012 but at
V. NORMATIVE EFFECTS OF EFFICIENCY’S RATIONALITY

All of the foregoing is an instrumental analysis of efficiency’s effects. But making efficiency a central ambition of adjudication has effects in non-instrumental dimensions as well. As new practices gain dominance, they usually gain acceptance, which requires rationales to explain and justify that acceptance. The pervasive adoption of adjudication strategies in service of efficiency, especially plea bargaining, helps to redefine the norms that inform both ideas of adjudication’s purposes and, more broadly, those that influence the state’s policies of criminal law enforcement. Efficiency has become a dominant value and purpose of adjudication, and its utilitarian nature undermines the non-utilitarian public and social norms of criminal process and policy.

A. Non-Instrumental Purposes of Adjudication

Adjudication’s traditional purposes and rationales have been predominantly non-utilitarian. Constitutional rights to introduce evidence and confront state witnesses serve political norms that value individual autonomy and process participation, independent of whether they improve accuracy in trial judgments. Similarly, statutory rights for victims to be present at court proceedings acknowledge a non-instrumental value in adjudicative process. More generally, the mandatory public nature of criminal courts has political value to defendants and citizens, aside from much lower rates than in previous years). For an analysis suggesting that increased state prison spending comes at the cost of other priorities funded by state general funds, see Prema Anand, Winners and Losers: Corrections and Higher Education in California (Sept. 5, 2012), http://www.cacs.org/ca/article/44 (reporting that, following the 2007 recession, California state spending on prisons exceeded spending on higher education for the first time); NAACP, Misplaced Priorities: Over Incarcerate, Under Educate 1 (2d ed. 2011) (reporting that “[o]ver the last two decades, . . . state spending on prisons grew at six times the rate of state spending on higher education”).

One version of this phenomenon is the status quo bias. See Daniel Kahneman et al., Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias, 5 J. Econ. Persp. 193, 197–98 (1991) (describing the status quo bias). A related aspect is the acceptance of new programs (such as Medicare) or practices (such as racial integration or gay marriage) that gain wide acceptance after overcoming initial unpopularity or resistance. An example in legal doctrine is the argument that judicial supremacy in constitutional interpretation is now widely accepted but was not until the Supreme Court asserted that supremacy in Cooper v. Aaron, 358 U.S. 1 (1958). See Larry D. Kramer, The Supreme Court 2000 Term Foreword: We the Court, 115 Harv. L. Rev. 4, 6–7 (2001) (arguing for this account).
the instrumental benefits of transparency. Even the requirement for judges to confirm the accuracy of guilty pleas, minimal though it is in American law, serves not only to enhance accuracy but to respect a commitment to truth for its own sake.

The jury, a critical feature of common law adjudication, has strong non-utilitarian rationales. The jury has long been explained and defended with reference to its constitutive role in democratic governance and the political value of lay participation (including the benefit of jury service to jurors themselves as citizens), in addition to consequentialist functions such as a check on government power. The Supreme Court has been quite explicit about the criminal jury’s non-utilitarian normative roles, including its presumed disposition to rest verdicts as much on moral assessments as on logic, formal proof, and “any linear scheme of reasoning.” The Court endorsed lay jurors’ power to demand evidence of sufficient narrative power and detail “to implicate the law’s moral underpinnings” so that its “guilty verdict would be morally reasonable,” rather than resting merely on rational proof of “the discrete elements of a defendant’s legal fault.” In death penalty cases especially, it guards “the jury’s moral judgment about the defendant’s actions” through rules governing evidence and jury instructions. Decisions defining proper jury selection procedures acknowledge the public value, independent of effects on accuracy, of diverse and representative jury membership. All of this is part of a long tradition of understanding the common law jury as a body especially suited to render judgments in light of community values as well as legal ones. Its verdicts are expected to be different than those of professional judges in ways that common law jurisdictions val-

63 Bibas, supra note 14, at 144–150 (discussing non-utilitarian normative significance of adjudication procedures).
64 See Fed. R. Crim. P. 11(b)(3) (“Before entering judgment on a guilty plea, the court must determine that there is a factual basis for the plea.”).
66 Id. at 188.
ue. (Thus we might find some irony in how common law systems have moved more quickly than civil law ones to displace lay decision-making with judgments negotiated by professionals.)

These long-established accounts of juries and the nature of judgments implicitly embrace a reality that efficiency-dominated accounts of adjudication ignore: Adjudication plays a **constitutive** role in substantive justice. The process of reaching judgments is context-specific and underdetermined by the applicable legal rules. This is especially obvious when the governing standard is one that is defined broadly by terms such as due care, fault, or negligence. Law’s necessary generality, as Professor Kyron Huigens has explained, inevitably leaves, “in the interstices of the offense definition, . . . some discretion over precisely what the law will require in specific circumstances.” Because of this, the choice of adjudication process matters. From the choice of process follows differences in available evidence, legal and moral argument, modes of advocacy or persuasion, and, not least, differences in the character of the decision-maker. Judges bring legal training while jurors are valued for untutored access to lay moral sentiments; popular election produces different sorts of judges than merit selection with life tenure; demographically diverse juries are different from special panels of elite lay persons selected by “blue ribbon” or “key man” processes.

The same is true regarding prosecutors’ character and decision-making, as we can see in the law that defines prosecutors’ broad charging discretion and insulates it from judicial review. Charging decisions are significant, though non-final, judgments about liability. Public pros-

---


70 Huigens, Mechanics of Fault, supra note 69, at 477.

71 Processes designed to yield demographically representative juries are not required under the fair-cross-section doctrine of the Jury Clause and Due Process Clause. See Taylor, 419 U.S. at 538; see also Duren, 439 U.S. at 367–68. On the older model of special blue ribbon juries and key man selection systems, see, for example, Fay v. New York, 332 U.S. 261, 270 (1947) (approving New York’s system of blue ribbon special juries for criminal cases, authorized in state statutes); Jeffrey Abramson, We, the Jury. The Jury System and the Ideal of Democracy 99, 115–17 (1994) (noting the key man method of selecting elite citizens for juries was the dominant model in the United States through the 1960s); Neil Vidmar & Valerie P. Hans, American Juries: The Verdict 67 (2007) (describing key man jury selection).
ecutors are granted that authority not only because they have better knowledge of available evidence and enforcement priorities but also because of the character of the prosecutor’s office. They are more disinterested than victims who could serve as private prosecutors; their role is informed to some degree by a norm that prioritizes disinterested justice and public interests. But in the United States, the prosecutor’s office is also designed to be politically responsive, which intentionally informs the character of the professional role and the substance of prosecutorial judgments.

Whether in charging decisions or final verdicts, the decision-maker’s character is critical because it affects—or helps to constitute—the legal judgment. In some number of cases, a difference in decision-maker means a difference in judgments, without either judgment necessarily deemed an incontrovertible error. In the common law tradition, particularly with regard to the jury, this character difference has long been taken as not only inevitable but desirable. Likewise, in the United States (but only the United States), the prosecutors’ politically attuned motivations and judgments are mostly taken as virtues.

B. Consequentialism and Incommensurable Values

Prioritizing efficiency gains devalues or ignores these non-utilitarian features of adjudication. Efficiency-oriented analysis is consequentialist, and it tilts policymakers and courts toward a normatively thin, utilitarian, and instrumental conception of adjudication that works against hard-to-quantify values and claims for the intrinsic value of criminal process. Consequentialist analysis largely treats adjudicative procedure and decision-makers’ identity as non-substantive, or as non-constitutive of judgments. It assumes that law’s content lies mostly in substantive rules and that the process of its application is substantively and normatively insignificant. Most adjudication reforms motivated by efficiency disregard the effects that process changes have on substantive outcomes—on judgments, and justice. It is hardly controversial to note that plea bargaining is not simply a quicker way of reaching the same outcomes we could achieve with more trouble through trials (even holding aside the explicit discount from post-trial sentences that drives plea bargaining). Nor is it a process that equally serves—or serves at all—other constitutional or political purposes of the trial. The process of evidence evaluation is radically different, and the evidentiary record usually is as well.
So are the decision-makers; prosecutors, and secondarily judges, largely control the plea negotiation process.

By obscuring these differences and focusing on measures of productivity, the policy priority for efficiency distorts our understanding of adjudication’s structure and the full implications of efficiency-minded reforms. From this consequentialist perspective, efficiency gains are more readily viewed as unambiguous gains, because the consequentialist framework diminishes the significance of other, qualitative interests that are not included in, or served by, the efficiency calculus. To return to the motor vehicle analogy: Greater fuel efficiency can be taken as an unambiguous gain only if one ignores required tradeoffs in diminished vehicle power or size.

The problem of competing values is more difficult in domains such as adjudication, where certain values cannot be assessed on the same scale or dimension, such as price. In that sense they are not only competing interests but incommensurable values. Policymakers nonetheless must choose between them. Choices reveal preferences, in the economist’s terminology, but that does not mean the conflicts have been reconciled and the choice is fully satisfactory. This reflects the fact that value has a tragic dimension; one is sometimes required to choose between two goods, rather than between a good and a bad, so that something of value is inevitably lost.72 The price we pay for a good, or accept for its loss, does not always capture its full value, just as monetary awards do not fully compensate for some harms.

Such conflicts are well known in the context of adjudication. As in other domains, resource limits conflict with other interests—for example, with maximizing accuracy, defendant trial participation, and jury supervision of prosecutions. (Although, as noted, the severity of resource constraints depends on the demand for those resources—that is, caseloads—which are partly matters of policy discretion.) The accuracy goal might conflict with use of juries as decision-makers or with defendants’ confrontation rights.73 Similarly incommensurable tradeoffs lie be-

---

72 See Huigens, Skeleton of Value Fallacy, supra note 69, at 546.
73 Regarding juries and accuracy, see generally Harry Kalven, Jr. & Hans Zeisel, The American Jury 55–65 (1966) (assessing jury verdicts against judges’ decisions in the same cases). The Federal Rules of Evidence attempt to mediate the potential conflicts in accurate verdicts and jury decision-making by excluding relevant evidence that jurors seem likely to misinterpret or misuse.

Confrontation rights can in theory undermine accuracy by intimidating witnesses who then withhold testimony or because cross-examination is less effective than other means of scru-
hind many rules of adjudication practice. Rules specifying whether prosecutors must disclose certain kinds of evidence before trial, for example, rest on choices among competing interests or values. Non-disclosure lowers prosecutors’ costs and aids the confidentiality or security of state sources, yet it may burden the defense and trial court administration, undermine accuracy in the adversarial process, and conflict with some versions of fairness norms. Rules choose among competing interests, but they cannot reconcile or fully serve all of them. Such choices are rational yet also inevitably imperfect, and in that sense tragic. Consequentialist, efficiency-minded analysis readily values and favors some of these interests while obscuring others.

We see this effect in another body of law in which efficiency has recently served as a prominent rationale—strict criminal liability. The Supreme Court, widely followed on this point by state courts, commonly interprets certain kinds of criminal offenses, notably “public welfare”


75 Burdens to the defense may take the form of inadequate preparation for trial, confrontation of state evidence, or efforts to independently discover evidence that the state possesses but conceals. Trial court administration bears the burden of necessary delays to allow defendants time to prepare to respond to evidence that the government does not reveal until trial. For a rule authorizing government nondisclosure until mid-trial, see 18 U.S.C. § 3500 (2006).

and regulatory crimes, to lack requirements of proof regarding a defendant’s culpable mental state. The rationale is consequentialist: Strict liability makes conviction easier, which is deemed essential to effective enforcement in the context of these offenses. Proof of culpable mental state is not required for “prosecutions under statutes the purpose of which would be obstructed by such a requirement,”77 or when “necessary to effective administration of the statute.”78 Strict liability’s “purpose and obvious effect . . . is to ease the prosecution’s path to conviction.”79 Strict liability certainly makes enforcement more efficient, in the sense that it makes convictions quicker and cheaper to achieve (as does plea bargaining). Equally obviously, this instrumental rationale conflicts with the norm that criminal liability should be imposed only on the morally deserving, as determined by a finding of fault. The instrumental priority for efficient imposition of sanctions pushes aside the incommensurable value of culpability as a prerequisite for punishment.80

C. Efficiency, Measureable Interests, and Policymaking

Part of efficiency’s effect on adjudication, then, is that it encourages a consequentialist mode of analysis that tends to ignore or devalue certain kinds of competing qualitative interests. Some of this effect may come from a problem familiar in other policy contexts but little noted in adjudication literature. Some interests are more easily measured than others, and policy decisions often pay more attention to considerations that are easily measurable, even if they are poor or incomplete proxies for the real, underlying interests. This problem is familiar with respect to measures of gross domestic product (“GDP”), where it has drawn in-

creasing attention in recent years. For decades governments and researchers have used GDP as a proxy for national wealth, and per-capita GDP as a proxy for individual wealth and wellbeing. GDP aggregates a wide range of economic data in order to provide a measure of the total value of all final production in a certain period. That measure is expressed as a single, bottom-line figure of total economic output. Yet by design GDP relies on accessible data of market values and excludes many non-market values, including measures of human capital such as education, environmental damage, public health, and relative income inequality. As a measure of social progress or average wellbeing, governments and policy analysts have increasingly come to view GDP as inadequate. Worse, it can mislead policymakers, by encouraging measurable activity even though the activity causes harms or undermines interests that are unmeasured and harder to recognize. As a report by a panel of leading economists put it, “What we measure affects what we do; and if our measurements are flawed, decisions may be distorted.”

Much the same problem is apparent in the reform of adjudication practice oriented overwhelmingly toward greater efficiency. A handful of important factors are easy to measure, especially public agency budgets, criminal court caseloads, times to disposition, and the portion of prosecutions ending in convictions. Other values are qualitative and difficult or impossible to measure—the intrinsic or political purposes of the trial process, the value of process participation, and differences from altered processes in the accuracy or sentencing terms of convictions. The effects of adjudicative efficiency on the real purposes of criminal law

---

enforcement—crime prevention, public order, and the like—are nearly as hard to measure. Moreover, even if one holds those unmeasured values aside, courts and other policymakers lack data on how various procedural alternatives actually affect adjudication costs. That leaves courts to rely on little more than intuitions and policy preferences. In recent decades, the Supreme Court’s assumptions have closely tracked the model of efficiency-oriented adjudication that prioritizes concerns about resources-per-case and gives little credence to competing qualitative purposes. Hence the Court makes confident assertions about the absolute necessity not only of plea bargaining, but of particular bargaining procedures such as the waivability of evidence disclosure mandates or unlimited plea discounts.\(^{82}\) The Court is equally certain that ill effects would follow from judicial inquiry in prosecutors’ reasons for charging decisions or for withholding evidence during plea negotiations.\(^{83}\) Yet no real evidence supports those conclusions. Some, in fact, belies it, such as rules in England and in some U.S. states that require more evidence disclosure from prosecutors, England’s rule limiting the magnitude of plea bargain discounts, and the success of lower courts in the United States, before the Supreme Court’s 1978 decision \textit{Bordenkircher v. Hayes}, at developing constitutional doctrines to do much the same thing.\(^{84}\)

\textbf{CONCLUSION: EFFICIENCY AND CRIMINAL LAW ENFORCEMENT POLICY}

More efficient adjudication lowers the costs of convictions and makes criminal law enforcement cheaper and more tempting to use. In this way, it contributes to greater use of criminal law and punishment. And just as efficiency subtly facilitates a reordering of the competing values

\(^{82}\) United States v. Ruiz, 536 U.S. 622, 633 (2002) (approving nondisclosure of exculpatory evidence to facilitate fast-track plea bargaining); Bordenkircher v. Hayes, 434 U.S. 357, 364–65 (1978) (concluding the Constitution does not regulate disparities in trial versus plea liability and sentencing outcomes created by prosecutors to encourage guilty pleas); Blackledge v. Allison, 431 U.S. 63, 71 (1977) (“[T]he fact is that the guilty plea and the often concomitant plea bargain are important components of this country’s criminal justice system.”); Santobello v. New York, 404 U.S. 257, 260 (1971) (“[Plea bargaining] is an essential component of the administration of justice. . . . 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.”).


\(^{84}\) See, e.g., Cooper v. United States, 594 F.2d 12, 18–19 (4th Cir. 1979) (requiring the government to honor terms of a plea agreement); United States v. Bowler, 585 F.2d 851, 853–54 (7th Cir. 1978) (same). On England’s limitation of discounts for plea bargains, see supra note 7.
within adjudication, it contributes also to the larger-scale norms and policy choices regarding what social problems should be addressed with criminal law, and whether criminal law or an alternative policy intervention is most appropriate to address a wide range of problems related to personal and national security, social order, public health, and economic risk. Cheaper convictions make for cheaper criminal law enforcement, which is likely to mean more criminal punishment. And as more criminal punishment (either more frequent or more severe) becomes the norm, punishment as a social practice and policy gains the bias of the status quo. It starts to appear less problematic, and perhaps more necessary.

One implication of this argument is that there is an upside to the relatively high cost of some goods. That is especially so for goods such as criminal punishment that carry substantial costs as well as benefits, both of which are hard to measure. Punishment is not in the class of goods that have no negative externalities. Even the most ardent defenders of the idea that criminal punishment is an affirmative good rather than a necessary evil concede that it has some bad effects, including ancillary harms to third parties or communities. Imprisonment policies can have criminogenic effects, leading to more offending rather than less. Independent of deterrent effects or offenders’ reintegration after serving sentences, incarceration causes injuries to offenders’ families, offenders’ post-prison life prospects, and the social capital of communities in which criminal sanctions are highly concentrated. On the other side of the balance sheet, the benefits of punishment are hard to assess. There is intractable uncertainty about the optimal amount of punishment for the goals of public safety, economic order, and public health that punish-

85 Dan Markel et al., Privilege or Punish: Criminal Justice and the Challenge of Family Ties, at xi–xx (2009) (recognizing punishment’s value as an affirmative good but expressing a willingness to tailor its contours in special circumstances to mitigate unnecessary third party harms); Darryl K. Brown, Third-Party Interests in Criminal Law, 80 Tex. L. Rev. 1383, 1386–96 (2002) (describing harms to others from punishment and legal recognition of those interests). For a view that punishment is only a necessary evil, see Douglas Husak, Over-criminalization: The Limits of the Criminal Law (2008).

ment serves. A few reject any role for criminal law and punishment, but deterrent effects elude sound measurement, and the relative effectiveness of criminal law compared to other policies that can address the same social problems is often deeply uncertain.

This is one theme in the large literature on the causes of the dramatic rise in incarceration rates and use of criminal law since the 1970s, much of which points to political explanations and stresses the tenuous relationship of crime rates to incarceration rates. The more political criminal enforcement policy is, the more likely it is that enforcement costs regulate political decisions. (Political in multiple senses, including discretionary policymaking and, more negatively, actions intended for partisan or electoral gain, or decisions based more on ideology than good-faith assessments of evidence.) The U.S. political system is recognized as one in which expertise and relatively neutral policy analysis play smaller roles in policymaking than in most other advanced democracies. On top of that, the U.S. criminal justice system is unique in its preference for politically responsive prosecutors and judges rather than politically insulated ones. In this context in particular, lower adjudication costs are not necessarily an unmitigated good. By providing more resources for politicized policymaking, efficiency makes room for more ill-conceived policies. In other words, we expect undesirable rebound effects from adjudicative efficiency: an excess of adjudication’s prod-

87 See Braithwaite, supra note 32, at 124–25, 230–31, 248 (describing disadvantages and comparative ineffectiveness of command-control enforcement compared to other regulatory strategies).


90 See generally Stuntz, Constitution of Criminal Justice, supra note 2, at 820–21, 838 (describing the local election of criminal justice officials and arguing for reform).
ucts—convictions and sanctions. *Less efficient* adjudication could work like a Pigouvian tax to reduce negative externalities, like a fuel tax designed to reduce but not eliminate fuel consumption. Higher costs lead decision-makers to reassess demand and to ensure that the benefits of the good are sufficiently high to justify its full costs. They force a harder look at value—in the criminal context, a harder look at the efficacy of some charges compared to other ways of reducing wrongful harms.

In the United States, with the world’s highest incarceration rate and political processes that channel populist enforcement preferences into policy, it is plausible to worry that the costs of criminal adjudication can be too low. Adjudication’s efficiency contributes to excessive enforcement and punishment policies. This sounds heretical in an age that valorizes efficiency and is skeptical of both public expenditure and public-sector inefficiency. Yet high costs—or, the same thing, resource limits—are uncontroversial means to force better spending choices in many contexts, from Pigouvian taxes on fuel to legislative constraints on prosecutors’ budgets that function as one tool to limit charging discretion. (U.S. policymakers are sometimes quite open about cutting enforcement funding, notably for tax evasion, securities fraud, and environmental regulation, in order to limit prosecutors’ enforcement capacity.) State budget constraints in recent years have prompted modest changes in sentencing laws to reduce incarceration rates, which many scholars applaud.

---

91 6 The New Palgrave Dictionary of Economics 435–36 (Steven N. Durlauf & Lawrence E. Blume eds., 2d ed. 2008) (defining Pigouvian taxes as “taxes designed to correct inefficiencies of the price system that are due to negative external effects”).


93 See Rachel E. Barkow, Federalism and the Politics of Sentencing, 105 Colum. L. Rev. 1276, 1291–97 (2005) (arguing that tight state budgets have deliberation-enhancing effects on legislative policymaking about criminal sentencing that help correct for distortions in the
recognizing that an efficiency improvement is distinct from efficiency’s effects reminds us that there is no ex ante reason to assume that lower adjudication costs necessarily lead to better criminal justice policy outcomes, or that higher costs are inevitably linked to worse ones. Those judgments are difficult, and they are at bottom political. But it can improve political decision-making—as a Pigouvian tax can improve decision-making—to acknowledge that the price of goods with negative externalities can be too low. American courts, which have played the lead role in making criminal process more efficient, could improve American political decision-making by developing the law and practice of adjudication with less priority for its efficiency and more for its traditional, qualitative public interests.
