Institutional Coordination and Sentencing Reform

Daniel C. Richman
Columbia Law School, drichm@law.columbia.edu

Follow this and additional works at: https://scholarship.law.columbia.edu/faculty_scholarship

Part of the Law Enforcement and Corrections Commons

Recommended Citation
Available at: https://scholarship.law.columbia.edu/faculty_scholarship/1402

This Working Paper is brought to you for free and open access by the Faculty Publications at Scholarship Archive. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of Scholarship Archive. For more information, please contact cls2184@columbia.edu.
March 2006

Institutional Coordination and Sentencing Reform

Daniel Richman
Professor of Law

Forthcoming in Vol. 84, Issue 7 of the Texas Law Review as part of Punishment Law and Policy Symposium

This paper can be downloaded without charge from the Social Science Research Network electronic library:
http://ssrn.com/abstract=889389
Deciding how much time a person should spend in prison for a serious crime is an inherently moral and political act. And it is certainly cold-hearted and philosophically problematic to view sentencing as just an agency problem, with criminal defendants as objects of a system in which prison terms are simply outputs. So I won’t even try to justify resort to a narrow institutional perspective as a normative matter. But, for better or worse, those political actors with the greatest influence on sentencing regimes have to think in aggregate terms. While there is considerable normative appeal to the idea of courtroom actors, and particularly judges, crafting an individualized sentence for each defendant, we need also to recognize that for elected officials at the top of the prosecutorial hierarchy, sentencing – particularly sentencing after a negotiated guilty plea – presents just another iteration of the classic problem of administrative law: how to limit the ability of agents to take advantage of informational asymmetries to slack off or import their own policy preferences.

One need hardly embrace the policy preferences of elected executive officials or the policy process to see the virtues of considering this internal perspective. The first is simply a matter of advocacy: However convinced one is that sentences are too high or that sentencing policy should be the exclusive province of judges, these are not likely to be effective starting points for conversations with sentencing hierarchs who can promote their sentencing preferences only by restricting the authority of line actors. And these sentencing hierarchs dominate the policymaking and legislative process. Budgetary considerations, at least at the state level, may lead statewide officials to reconsider levels of incarceration. But within these constraints, they still will

---

1Professor, Fordham Law School. Thanks to Richard Boylan, John Pfaff, Bill Stuntz, Ian Weinstein, Ron Wright, and attendees at the Punishment Law and Policy Symposium for extremely helpful comments. Thanks also to Susan Klein, Jordan Steiker, and the Texas Law Review for inviting me to participate in the symposium and for being such generous hosts.


have enforcement priorities. The second is a matter of accountability and transparency. Legislators, elected officials, and political appointees may be all too prone to go overboard in getting “tough on crime,” but our general expectation (at least outside the criminal justice area) that unregulated agents are prone to drift or shirk ought not be relaxed. Perhaps there is some optimal amount of drifting or shirking that will compensate for the failure of the political process to appropriately calibrate its response to crime. But even as one recognizes the value of “working group”\textsuperscript{4} justice – the justice meted out within the courtroom triad of prosecutor, defense attorney and judge – one can still find some counterbalancing value in the efforts of politically accountable officials to maintain a degree of control over the compromises reached by their unelected minions in the low visibility world of plea bargaining.

Yet how do these officials exercise such control? And what effects will the successful exercise of such control have on the rest of the system? These are the questions considered (but hardly resolved) here in an effort to start exploring how the internal prosecutorial monitoring project conflicts with or reinforces particular sentencing regimes. As a descriptive matter, this is a technical inquiry into the mechanics of institutional coordination and attempt to add to the literature that, when it does consider the politics of sentencing, tends to treat prosecutorial interests as monolithic. Treatments of prosecutorial discretion in the sentencing context also tend to focus on its challenge to horizontal equity and judicial discretion within sentencing regimes. The goal here is to reverse the arrow, and, using an internal executive perspective, start looking at how sentencing regimes and judicial enforcement of those regimes can be used as tools for the hierarchical control of line prosecutors in the plea bargaining process. To be sure, measures that promote hierarchical control will also promote horizontal equity. Indeed, the two goals have much in common. But the focus here, and perhaps in certain policymaking circles, will be on control.

First, we will consider a problem arising out of ostensibly successful regulation within a local prosecutor’s office. Then, we will consider issues relating to regulation from outside the office, to see how judicial supervision of plea bargaining through factually intensive sentencing inquiries can reflect (and perhaps occur because of) the interest of a state-wide or national centralized prosecutorial authority in controlling its own minions. Finally, we consider how the decentralization of authority in a local office can give the chief prosecutor of that office a perspective much like that of a state-wide official.

\textit{A. Internal Management and the Pitfalls of Unilateral Reform}

We start with a story that, at first blush, seems to be one in which shirking by line actors is successfully and unilaterally attacked by their politically accountable superior. On further analysis, however, it ends up looking a lot more like a case study in institutional disarray.

In their fascinating 2002 \textit{Stanford Law Review} article, Ron Wright and Marc Miller

\footnote{See James Eisenstein and Herbert Jacob, \textit{Felony Justice} (Boston: Little, Brown, 1977).}
presented and assessed the efforts of New Orleans District Attorney Henry Connick to take on a system in which, he claimed, “lazy” prosecutors were using the plea bargaining process to “‘move’ cases and avoid trial.” Connick “instructed his prosecutors not to engage in plea bargaining – particularly charge bargaining, except under very limited circumstances.” What interested Wright and Miller, however, was that Connick coupled this promulgation with a move calculated to reduce the incentives that drive deeply discounted plea offers: a significant commitment of resources to the screening of cases presented by the police for prosecution.

Attorneys in the screening unit – some of the most experienced in the office – would review each investigative file from the police, speak to all the key witnesses and victims and generally assess the strength of the case. Office policy then required them – if they decided to go forward – to charge the most serious crimes that the facts would support at trial. And office policy made it extraordinarily difficult for the assistants handling the case thereafter to retreat from the charges thus specified. The result, as intended, was that the office declined a relatively high number of cases – a fact it blamed on poor police work – and that bargaining played a extraordinarily limited role in the disposition of cases that it did pursue. Wright and Miller thus found that “[t]he data mostly support [] Connick’s claims to have implemented a screening/bargaining tradeoff over the last thirty years.” The lesson, they noted, is “that a committed prosecutor can implement the screening/bargaining tradeoff even without the conscious support of other actors in the system.”

Wright and Miller were careful to note that “[a] prosecutor who shifts to stronger screening risks more strained relations with the local police.” They optimistically suggested, however, that at least in theory, “an intense screening policy should encourage better police work,” particularly if the police insisted on “feedback from the prosecutor’s office both during and after the screening process,” and “raise[d] the political cost for recalcitrant prosecutors who continue to decline cases that the police give a high priority.”

In March 2002, after twenty-eight years in the post (five terms), Connick announced that

---


6Id. at 63.

7Id. at 65.

8Id. at 117.

9Id. at 97-98.
he would not seek reelection. His successor, Eddie Jordan, took office in January 2003 and soon drew on New Orleans Police Foundation money to commission a study by a consulting team led by Paul Schechtman. The team’s report is a portrait of an utterly dysfunctional system, with the District Attorney’s office at least acquiescing in and perhaps contributing to this dysfunctionality.

The consultants’ report noted the lack of coordination between the DA’s office and the police and bluntly set out the consequences of unilateral action: The conviction rate for cases actually pursued hovered around between 90% and 97% between 2002 and 2004. And, of the defendants who pleaded guilty, between 82% and 95% pleaded as charged. But only 40.3% of the Part I (FBI index) crime cases presented by the police were accepted for prosecution in 2002 and only 38.2% in 2004. It is difficult to put these acceptance rates in a precise comparative perspective. But in California, at around the same time, 60.7% of violent felony arrests and


12Id. at 7 (noting “lack of an arrest/prosecution strategy with the NOPD [New Orleans Police Department] to address the issue that many arrests are not prosecutable.”). A 2005 report by the Metropolitan Crime Commission in New Orleans noted: “The negative impact of the lack of coordination between the NOPD and DA’s Office is nowhere more apparent than in the case screening process[]. The current system was put in place in the 1980's and little has changed since then.” Metropolitan Crime Commission, New Orleans, La., Performance of the New Orleans Criminal Justice System, 2003-2004 (Aug. 2005), available at http://metrocrimeno.org/2Perf_of_the_NO_Criminal_Justice_System_2003-20041.pdf

13Toward a Fully Integrated Criminal Justice System, supra note __, at 16.

14Id. at 18.

15Id. at 12.

16Bureau of Justice Statistics, Felony Defendants in Large Urban Counties, 2002 (2006), reports that in forty large counties, selected to representative of the nation’s 75 most populous counties, 68% of all felony arrests ended in convictions; 60% of all violent felony arrests, and 72% of all property arrests. Id. at 24, tbl. 23. Discerning screening rates in the reporting jurisdictions is a challenge, however. In some of these forty counties, cases screened out by prosecutors are reported as “dismissals.” In others, cases in which someone was arrested on
76.4% of property arrests ended in convictions (not simply acceptances), with a 69.8% figure for all felony arrests. While one cannot easily quantify the point, the New Orleans acceptance rates were extraordinarily low.

The New Orleans police evidently weren’t responding to rigorous prosecutorial screening by bringing stronger cases. They were just continuing to bring weak ones and willing to suffer a refusal to prosecute. The lack of coordination between New Orleans’ police and prosecutors was surely not the city’s only criminal justice problem before Hurricane Katrina, and the structure and operations of the city’s criminal justice institutions will likely be different after it. Yet the episode offers a powerful reminder that transcends both the city’s idiosyncratic approach to criminal justice and the rationale behind DA Connick’s demand for more rigorous screening: For better or worse, the police are critical actors in the sentencing process.

Once one sees a casefile not as a given but as an artifact of a fact-gathering process that is primarily dominated by the police and that incorporates prosecutorial decisionmaking only to the extent that some political or bureaucratic mechanism mediates between the two coordinate entities, one sees the limitations of a sentencing literature that focuses on the results of the plea bargaining or formal adjudicative process. Before we get too caught up in sentencing differentials across the defendants who are actually prosecuted, we ought to give some thought to the defendants who get away because investigative or adjudicative resources are not expended on them. And, of course, it’s also helpful to get all murderers off the streets.

Without pointing fingers at DA Connick – who after all may have been making the best of a bad situation – one can use the breakdown in relations between his office and the New Orleans Police Department to raise some larger questions. Where prosecutors and police have independent sources of authority and stand in a coordinate, not hierarchical relationship, what mechanisms ensure that each agency internalizes, or at least strongly considers the policies and preferences of the other? One might model the relationship as a bilateral monopoly and assume that some degree of negotiation occurs between the two parties necessary to the prosecution of a felony charges but prosecutors chose not to proceed are not reported at all. Id. at 39.


criminal case.\textsuperscript{19} The validity of this assumption, however, turns on the degree to which the two are trying to maximize their joint output, or that each is judged by a performance measure that implicates the others.

One could imagine an administrative mechanism that would prevent a prosecutors’ office from shifting costs to the police (for however laudable a reason) and/or prevent the police from shifting its costs back to the public (in the form of unprosecuted offenders). In the federal system, the Attorney General – who, along with the Deputy Attorney General, has hierarchical authority over federal prosecutions and those enforcement agencies housed in the Justice Department – at least potentially plays this role.\textsuperscript{20} But state systems generally lack even this small degree of structural coordination, with police chiefs reporting to mayors, and district attorneys directly elected. The degree to which alternative mechanisms will develop will ultimately be a matter of politics.

Bill Stuntz and I have argued that electoral accountability can play an important role in ensuring the health of state systems – at least in comparison to the federal system -- since voters can easily grasp and track the prosecutions of the murders, rapes and robberies that are staples of those systems.\textsuperscript{21} In 2003, for example, the DA in San Francisco, Terence Hallinan, also found himself at loggerheads with the police. Shortly before the 2003 election, the most recent data (from 2001) showed his office declining to file charges in 36.7% of all felony arrests (compared to a statewide figure of 13.6%) – a rate Hallinan attributed to sloppy police work. Just 29% of all adult felony arrests in 2001 ended in convictions.\textsuperscript{22} Hallinan lost the election to an opponent whose main platform was his incompetence and who made much of his low conviction rate.\textsuperscript{23}

\textsuperscript{19}Daniel C. Richman, Prosecutors and Their Agents, Agents and Their Prosecutors, 103 Colum. L. Rev. 749, 758 (2003).


\textsuperscript{22}Bill Wallace, San Francisco Ranks Last in Convictions; State Figures Show Relatively Low Rate for D.A.’s Office, San Francisco Chronicle, Oct. 17, 2003, at A1. [Note that conflicts between San Francisco’s police and the DA’s office were somewhat exacerbated by the obstruction of the justice investigation that Hallinan pursued against the police department’s leadership.]

\textsuperscript{23}Demian Bulwa, Harris Defeats Hallinan After Bitter Campaign, San Francisco Chronicle, Dec. 10, 2003, at A1; Maura Dolan, New D.A. Beat Incumbant While Embracing Positions; Defeat of Terence Hallinan May End Regular Clashes Between Prosecutors and
Indeed, two political scientists, Sanford C. Gordon and Gregory A. Huber, have recently gone so far as to predict that “voters will always reward prosecutors [electorally] for obtaining convictions and punish them for acquittals. This strategy holds irrespective of either how tough on crime voters want prosecutors to be, or how much information voters have about individual cases.”

But political accountability won’t always do the trick – either because so many elections are largely uncontested or because of deeper democratic failures. The New Orleans report noted that although the city had the highest homicide rate among the nation’s 71 largest cities in 2002 and 2003, the vast majority of homicides reported and screened in those years did not end in convictions. Indeed in 2002, only 14.3% of the 258 homicides reported and only 15.0% of the homicide cases screened ended in conviction. Why was this tolerated? Perhaps the rule is that (1) prosecutors are more likely to be held politically accountable than the police where there is a dysfunctional relationship between the two agencies – which would make sense, given that police chiefs are politically accountable only through elected mayors, and mayors have responsibilities (and sources of popularity) that go far beyond the criminal justice system. And (2) prosecutors will not always be held so responsible, either because they are held blameless or because of some larger failure of electoral accountability.

At any rate, the New Orleans program ought to be taken as an object lesson in the pitfalls of unilateral regulation of prosecutorial bargaining discretion.

B. Executive Regulation from Outside the District

For all its perverse systemic effects, the New Orleans D.A.’s screening initiative was a success story as an experiment in the regulation of line prosecutors from within their district. And this is not particularly surprising. Using a centralized intake mechanism and essentially freezing the valuation of every case that is allowed to go forward may not serve the interests of sentencing equity, broadly defined, but it will promote the internal management of the sentencing process within a prosecuting office. Managing the process from outside the office can pose very different (even insurmountable) challenges, however.

Perhaps the best evidence of how difficult it is for a prosecutorial hierarch to manage the process from beyond the district level comes from the story of Alaska’s plea-bargaining ban. Because of its late start, Alaska has the most centralized prosecutorial organization of all the fifty states. The governor appoints the attorney general, who in turn appoints and maintains authority


25Toward a Fully Integrated Criminal Justice System, supra note __, at 1.
over all district attorneys and assistant district attorneys. In 1975, Attorney General Avrum Gross famously announced a ban on plea bargaining that, like DA Connick’s initiative in New Orleans, relied on heightened screening standards at intake and severe limitations on plea dispositions thereafter. From the start, however, the ban “was not implemented uniformly throughout the state,” with “local legal culture” “shap[ing] the contours of the policy in each area.” And its bite diminished over time as successor attorneys general gave more discretion to local district attorneys. Indeed, by 2003, complaints from line prosecutors in Alaska were less about centralization and more about the lack of sufficient policy guidance from the above.

There are obviously many unique aspects to the Alaska experiment (beginning with its geographic setting), but it raises a more general issue: To what extent can a translocal prosecutorial hierarch get a handle on, let alone control, the way line prosecutors present cases to sentencing judges? Or put more provocatively: To what extent can such a hierarch do this without the assistance of sentencing judges? It may sound strange to think of judges as handmaidens of an executive agenda, but the issue is raised by recent developments – nicely highlighted in two pieces by Ron Wright – in one of the few states that comes close to Alaska in the extent it gives the attorney general hierarchical control over the state’s entire corps of prosecutors: New Jersey.

Wright tells how, in response to an increase in mandatory minimum statutes that it (understandably) saw as increasing the risk of prosecutorial manipulation of sentencing outcomes, the New Jersey Supreme Court interpreted these statutes to require the attorney general to draft statewide charging guidelines. And it specified that the attorney general, not the

---


28 Carns & Kruse, supra note __, at 34.

29 Id. at 35.


32 For an effort to model prosecutorial manipulation of mandatory sentencing regimes, see David Bjerk, Making the Crime Fit the Penalty: The Role of Prosecutorial Discretion Under Mandatory Minimum Sentencing, 48 J. L. & Econ. 591 (2005).
prosecuting attorneys for each county, keep effective control of these guidelines.\textsuperscript{33} The attorney general responded by promulgating just such guidelines for the covered offenses (mostly in the narcotics area),\textsuperscript{34} and the Supreme Court has worked hard to ensure compliance. Under the new guidelines regimes, prosecutors had to explain to trial judges precisely why they were invoking or not invoking enhanced sentencing provisions. The trial judges “reviewed these reasons to assure that the prosecutors were following the guidelines.”\textsuperscript{35}

The New Jersey Supreme Court’s readiness to readiness to take a laboring oar in institutional reform (“activism” is such a loaded term) might be a sufficient explanation for the implementation of this scheme. But one might still wonder whether it is merely a coincidence that this judicial initiative occurred and has flourished in a state in which county prosecutors are not elected but, like the attorney general, appointed by the governor, with senate confirmation. Did the New Jersey court commandeer the attorney general to serve its (ostensibly the statute’s) purposes? Did the attorney general readily enlist in an effort that served his own institutional goal of limiting the discretion of line prosecutors? These seem like questions worth pursuing, particularly since Wright points out that after promulgating the judicially mandated guidelines, the New Jersey attorney general created guidelines in other areas on his own initiative.\textsuperscript{36}

Noting how the “lack of monitoring or enforcement by other actors” reduces the value of internal prosecutorial guidelines “as a way to achieve consistency and accountability,” Wright suggests: “Together, the sentencing actors in New Jersey may be creating a partnership that could thrive in many places.”\textsuperscript{37} Yet Wright might be underestimating the role played by New Jersey’s unusual executive hierarchical structure. The “partnership” in New Jersey is not between prosecutors and judges generally, but between the judiciary and an attorney general who, but for the guidelines and their judicial enforcement, would be hard pressed to regulate how his scattered minions use sentencing concessions. Elsewhere, an attorney general would not, and would not be expected to, consider local prosecutors as his minions.

Among states, Alaska and New Jersey are outliers when it comes the interaction of hierarchical prosecutorial authority and sentencing schemes. They are among the few states whose prosecutorial regime offers at least a clear structural possibility of direct control and

\textsuperscript{33}Wright, Prosecutorial Guidelines, supra note __, at 1095, discussing State v. Gerns, 678 A.2d 634 (N.J. 1994).

\textsuperscript{34}For the most recent version of the guidelines, see http://www.njdcj.org/agguide/directives/section_1.pdf

\textsuperscript{35}Wright, Sentencing Commissions, supra note __, at 1031.

\textsuperscript{36}Wright, Sentencing Commissions, supra note __, at 1034.

\textsuperscript{37}Id. at 1035-36.
perhaps therefore a degree of political accountability this regard.\textsuperscript{38} In other states, the political independence of district offices appears to be a conversation stopper. Even where state-wide elected officials champion mandatory sentencing measures – with California’s three-strikes statute being a particularly famous example – they leave ample room for district variance, or at least stand ready to acquiesce in it.\textsuperscript{39} A ten-year retrospective study of the California law by the state’s District Attorneys Association cited data from the state’s most populous counties indicating that prosecutors exercised their discretion to ask for dismissal of felony strikes in 21-40\% of all three strikes cases.\textsuperscript{40} (Having already availed himself of this license, and limited the use of the statute within his district to violent third-strikes, the Los Angeles district attorney is now leading a campaign to amend the law and so limit use of the statute throughout the state.\textsuperscript{41}) This pattern has been echoed in other states with three strikes legislation.\textsuperscript{42}

Perhaps this acquiescence stems from a substantial harmony in the viewpoints of state officials and local prosecutors – who, if anything, may well be more prone to incarceration than the state officials who have to actually pay for prisons.\textsuperscript{43} Perhaps it reflects acceptance of the independent political status of prosecutors. Perhaps it reflects a perceived institutional inability to regulate. Or maybe some combination of these, and other factors. But outside those states with a formal statewide regime, statewide elected officials have been largely untroubled with what an

\textsuperscript{38}See Robert L. Misner, Recasting Prosecutorial Discretion, 86 J. Crim. L. & Crim. 717, 732(1996) (noting that in Alaska, Connecticut, Delaware, and Rhode Island, “local prosecutors are part of the statewide state attorney general’s office and are controlled by the state attorney general”).


\textsuperscript{40}California District Attorneys Association, Prosecutors’ Perspective on California’s Three Strikes Law: A 10-Year Retrospective (Summer 2004), available at http://www.cdaa.org/WhitePapers/Three Strikes.pdf


\textsuperscript{43}See Robert L. Misner, Recasting Prosecutorial Discretion, 86 J. Crim. L. & Crim. 717, 719-20 (1996); see also Rachel E. Barkow, Administering Crime, 52 UCLA L. Rev. 715 (2005); Rachel E. Barkow, Federalism and the Politics of Sentencing, 105 Colum. L. Rev. 1276 (2005). The Los Angeles D.A.’s campaign to limit the three-strikes law is good evidence, however, that ideology can loom larger than this fiscal dynamic.
outsider might consider district drift but what we take for granted, even celebrate, as local prosecutorial discretion.\textsuperscript{44}

The story has been quite different in the federal system. Here, the existence of a unitary Justice Department (at least on paper) offers at least the structural possibility of extra-district regulation. And the nature of federal criminal enforcement brings a far greater risk of policy drift than one finds in state systems. With no “federal crime rate” to provide an external performance metric, a level of resources dwarfed by the range of criminal jurisdiction allowed by Congress, and few clear public expectations of how their caseload should be selected and handled,\textsuperscript{45} federal prosecutors pose unique supervisory challenges to distant sentencing hierarchies.

In the federal system, the effort from outside the district to deploy judges as regulators of prosecutorial charging and bargaining has been quite clear, albeit ill-fated. This is the story of “relevant conduct.” And it is a story of a attempt that did not just fail, but really backfired. As Julie O’Sullivan has explained, the best justification for the modified “real-offense” sentencing approach of the Federal Sentencing Guidelines (which often required sentencing judges to consider criminal conduct alleged in counts that were dropped as part of a plea deal) is that it endeavored to limit prosecutor’s ability to undercharge or to otherwise understate the seriousness of a defendant’s conduct.\textsuperscript{46} To be sure, the approach promised to limit prosecutorial leverage in plea negotiations, but the goal was also to have a defendant’s sentence would be based on all of his “relevant conduct,” not just the subset of it the line prosecutor chose to identify. That judges would actually learn about all of a defendant’s “relevant conduct” was simply a matter of faith, faith that was probably misplaced – notwithstanding the efforts of probation officers to serve judges in this regard.\textsuperscript{47} As the Sentencing Commission’s Fifteen Year Report makes clear, the extent to which prosecutors – with the cooperation of defense attorneys – understate offense

\textsuperscript{44}Decentralization also means that efforts to cut sentences are best aimed at state legislators or state legislation, as has occurred in California and Arizona when voter initiatives aimed at low-level drug possession sentences. See K. Jack Riley, et al., Just Cause or Just Because?: Prosecution and Plea Bargaining Resulting in Prison Sentences on Low-Level Drug Charges in Californian and Arizona (2005) (RAND study), available at http://www.rand.org/pubs/monographs/2005/RAND_MG288.pdf

\textsuperscript{45}See Richman & Stuntz, supra note __, at 613.


conduct as part of a negotiated settlement may be had to quantify, but appears to be quite
significant.\textsuperscript{48} (It is also entirely possible that judges, uncomfortable with the severity of guidelines
sentences, consciously overlooked such factual omissions.) Indeed, as Judge Nancy Gertner has
written, “real offense” sentencing “in fact proved to be a boon for prosecutors rather than a limit
on their power.”\textsuperscript{49}

Questioning O’Sullivan’s justification of real-offense sentencing as a way to restrain
inappropriate prosecutorial leniency, David Yellen suggested in 1997 that if such leniency really
was a problem, the Department of Justice could find other ways to regulate charging decisions.\textsuperscript{50}

Not long thereafter, in 2003, the Ashcroft Justice Department endeavored to do just that through
the famous “Ashcroft Memorandum.” In the wake of Congressional efforts to further limit
judicial sentencing discretion under the (then mandatory) sentencing guidelines,\textsuperscript{51} Attorney
General Ashcroft directed that “in all federal criminal cases, federal prosecutors must charge and
pursue the most serious, readily provable offense or offenses that are supported by the facts of the
case, except as authorized by [designated supervisory officials]” under certain limited
exceptions.\textsuperscript{52}

It is far from clear that the Ashcroft Memorandum had any effect on line prosecutors in the

\textsuperscript{48}\textsuperscript{See U.S. Sentencing Comm’n, Fifteen Years of Guidelines Sentencing: An Assessment


\textsuperscript{50}\textsuperscript{Yellen, supra note __, at 1438.


U.S. attorneys’ offices.\textsuperscript{53} When it came to capital cases, which have long been given intensive attention at the highest levels of the Department,\textsuperscript{54} the Department could indeed preclude lenient dispositions (at least until the case went to the jury).\textsuperscript{55} But ordinary cases appear to have been more immune to this policy intervention. The Sentencing Commission’s Fifteen Year report found: “While charging and plea bargaining are officially regulated by nationwide DOJ policies, researchers reported that in practice these policies were less determinative of prosecutorial conduct than internal U.S. Attorney’s offices policies.”\textsuperscript{56}

That the Department has had so much trouble regulating line prosecutors does not render its efforts illegitimate. Even one who thinks federal sentencing generally too high can recognize the interest of politically responsible officials in implementing an Administration’s policy preferences in the trenches. And an Administration’s concerns are likely to grounded as much in institutional perspective as in policy preference.

In their insightful effort to explain why federal drug sentences had declined between 1992 and 1999, Bowman and Heise suggested:

Part of the explanation for the continued downward drift of federal drug sentences is surely that some of the front-line actors in the federal criminal system feel passionately that drug sentencing rules are too harsh. But a far more important consideration may be that a critical mass of those front-line actors are simply unconvinced of the imperative to commit the time, institutional resources, and emotional capital necessary to defend strict interpretation of drug sentencing rules.\textsuperscript{57}

\footnotesize{
\begin{itemize}
\item \textsuperscript{53}See G. Jack King Jr., NACDL Survey: USAOs Deny Ashcroft Memo Affecting Plea Bargaining, 27 Champion 6 (2003).
\item \textsuperscript{56}Fifteen Year Report, supra note __, at 84.
\item \textsuperscript{57}Frank O. Bowman, III & Michael Heise, Quiet Rebellion? Explaining Nearly a Decade of Declining Federal Drug Sentences, 86 Iowa L. Rev. 1043, 1133 (2001).
\end{itemize}
}
One factor that might have influenced the way line prosecutors made these judgments during the 1990s was that a growing careerism in U.S. attorneys’ offices during this period. By Todd Lochner’s calculation, “the average tenure time for assistant U.S. attorneys increased from roughly three-and-a-half years in the 1960s to just over eight years” by 1996-98.\(^{58}\) Lochner noted that these figures reflect the tenures only of people who left their positions as AUSAs, and that it appears that turnover rates declined from 6% to 2%.\(^{59}\) With this statistical backdrop, Lochner drew on personal interviews with numerous U.S. attorneys and assistants to conclude:

The growing length of tenure among assistants tends to discourage compliance with changes in office or department priorities because (1) assistants know that short of egregious behavior they will “outlive” both the attorney general and their respective U.S. attorney; (2) their careerist status has dramatically altered their personal incentive structures, making them less eager to take complex cases involving extensive discovery and overtime hours, and more eager to take cases that can be disposed of quickly and with little effort; and (3) their longevity as federal prosecutors tends to reinforce the view that they, rather than the department or the U.S. attorney, are in the best position to set an office’s prosecutorial agenda.\(^{60}\)

This growing careerism can only accentuate the inherent tension between line assistants and the political leadership of the Justice Department. The department, of course, can be well served by plea bargaining that results in efficient use of its investigative and adjudicative resources. And it can obtain such peak efficiency only by relying on the local knowledge of prosecutors and agents in the field. Yet with such reliance comes the inevitable risk of agency costs and policy drift, which, given the extent of federal enforcement discretion, presents far more of an oversight challenge in the federal system than it does in the states.\(^{61}\)

Hard-pressed to implement its controls through administrative directives, where will the Justice Department turn? Prediction is hard here, particularly given the varying extents to which different Administrations have been committed to centralization. Yet there are only a few

\(^{58}\)Todd Lochner, Strategic Behavior and Prosecutorial Agenda Setting in United States Attorneys’ Offices: The Role of U.S. Attorneys and Their Assistants, 23 Justice Sys. J. 271, 282 (2002). Offices in the largest metropolitan areas, such as New York and Los Angeles, appeared to be exceptions to the trend. Id. at 285. This is consistent with the finding by Boylan and Long that AUSA turnover is higher in high-private-salary districts than in low-private-salary districts. Richard T. Boylan & Cheryl X. Long, Salaries, Plea Rates, and the Career Objectives of Federal Prosecutors, 48 J. L. & Econ. 627 (2005).

\(^{59}\)Id. at 282-83.

\(^{60}\)Id. at 288.

\(^{61}\)See Richman & Stuntz, Al Capone’s Revenge, supra note __, at 599-618.
options: Concede defeat. Continue at the present level of regulation, in full knowledge that the program is more for public and congressional consumption than internal use. Increase the extent and rigor of administrative supervision, through some combination of direct bureaucratic and measures that drew on the monitoring potential of the FBI and other enforcement agencies. Or do more, presumably through sentencing legislation, to recruit judges as monitors.\footnote{For a proposal to have judges reject plea bargains that result in substantial sentencing concessions, see Oren Gazal, Partial Ban on Plea Bargains, 27 Cardozo L. Rev. __ (2006) (forthcoming).}

To be sure, this last option may not be workable. The Sentencing Commission’s Fifteen Year report noted: “Judicial scrutiny of plea agreements, supported by probation officers’ independent presentence investigations, were often inadequate to control plea bargaining because both judges and probation officers were heavily dependent on the information provided by the prosecutor in a given case. In addition, resource limitations and a reluctance to reject agreements [] made judicial rejection of plea agreements that undermined the guidelines relatively rare.”\footnote{U.S. Sentencing Commission, Fifteen Year Report, supra note __, at 84.} Moreover, as the First Circuit recently noted in a refreshingly candid opinion about fact-bargaining, the “costs of monitoring compliance” with “a system of mandatory disclosure of all possible information at a plea hearing” “are high and would come with the loss of “many of the efficiencies created by plea bargaining.”\footnote{United States v. Yeje-Cabrera, 430 F.3d 1, 28 (1st Cir. 2005).} The court also worried that such a system would “lead to the blurring of roles” by effectively involving judges in plea bargaining.\footnote{Id.} Still, one could imagine legislation bolstering the judicial oversight capability with measures of the sort recently suggested by Nancy King\footnote{Nancy King, Judicial Oversight, supra note __.} and significantly restricting the sentencing options available on any given facts.

One can have one’s own policy intuitions or preferences about the extent to which federal judges should have sentencing discretion. The point here, however, is that any realistic policy proposal ought to take into account the likely Justice Department response to it. A world of untrammeled judicial discretion but intensive administrative monitoring of prosecutors by Washington might ostensibly respect the prestige of the judiciary but at considerable cost to the rest of the system – at least to the extent one values local discretion in the federal system.\footnote{Compare Dan M. Kahan, Is Chevron Relevant to Federal Criminal Law?, 110 Harv. L. Rev. 469 (1996) (promoting more centralized control as mean of better internalizing costs of federal prosecutorial policies), with Richman, Federal Criminal Law, supra note __ (noting risks...
hard (but not impossible) to imagine Washington ever regulating line decision-making with any degree of success. Replicating the intensive supervision of death penalty cases outside the capital docket, where the death penalty has been authorized against a total of 371 defendants since 1988 – would be prohibitively expensive and politically difficult. But it is certainly possible to imagine Washington doing a lot more in this direction. While Congress has traditionally preferred to commit resources to the districts, rather than Washington, there is enough play in the budget and in existing bureaucratic structures for substantially more monitoring than is the rule today.

To what extent ought one to value and promote local discretion in the federal system? There are no easy answers here. Stephanos Bibas recently sought to draw some lines: Conceding that “local crime problems, caseloads, and knowledge vary and require varied responses,” he observed: “Our democratically elected representatives have decided to enact uniform national criminal laws to address national problems and enforce them with one voice through one agency – the U.S. Department of Justice. Locales that disagree cannot in effect secede from federal criminal law any more than they can secede from the Union.” He concluded the “justified variation is grounded in tactical decisions about localized crime problems – particularly, transient crime waves. Unjustified variation, in contrast, stems from value disagreement; from legally irrelevant factors such as race, ethnicity, sex, and class; or perhaps from strategic choices, especially concerning enduring crime problems.”

Even were it easier to distinguish between localized and enduring crime problems, Bibas’s distinction would still be normatively contestable. Not right or wrong, just contestable. After all, the notion of ostensibly uniform national criminal laws long predates the Justice Department (not created until 1870). And while Congress did decide to create such a central authority and give it powers over the U.S. attorneys’ offices, it also has also endeavored (sometimes) to insulate U.S attorneys offices from domination or at least to limit Washington’s ability to control those offices. Moreover, to the extent that Bibas worries that “[v]ariations also make the law seem

______________________________


70 Id. at 141.

71 See Richman, Federal Criminal Law, supra note __, at 805-10.
arbitrary, undercutting its perceived fairness and legitimacy, he underestimate Americans’
tolerance in this regard. There is, after all, considerable variation in the priorities of district
attorneys offices within a state, notwithstanding the uniformity of the prevailing penal code.
Because each district attorney is politically accountable to a different electorate, this variation may
be more or less defensible than in the federal system. The point, however, is that the mere
existence of national set of criminal laws does not, by itself, imply some optimal allocation of
authority between Main Justice and the U.S. attorneys’ offices.

Against this institutional backdrop, sentencing issues may take on a different complexion.
Once one sees the sentencing process as a means whereby Main Justice can monitor and constrain
the behavior of line prosecutors, then the hamstringing of judicial discretion so condemned in the
sentencing literature may (to some at least) be more attractive than its alternatives. If one, for
example, values local discretion (because, say, one prefers the district’s approach to the death
penalty or drug enforcement to the Administration’s, or because one thinks a district’s
commitment to attacking fraud or corruption will be greater than that of a particular
Administration or any national bureaucracy more generally), one may well prefer that Main
Justice recruit local district judges as monitors rather than engage in more direct bureaucratic
intervention. And it is quite possible that judicial monitoring, in the context of a relatively
inflexible sentencing regime, would sufficiently address Washington’s interest in limiting the
discretion of its prosecutors. One might, under this analysis, embrace a rather inflexible
sentencing scheme patrolled by judges charged with looking beyond bargained results not because
it is normatively worthy of celebration but as a second- (or third-) best alternative.

Do any of these speculations provide a blueprint for reform in the federal system? I doubt
it, but do think that reformers need to keep them in mind as the rough cease-fire that we now have
courtesy of Booker breaks down (a pessimistic but not unrealistic prediction). The last twenty-
five years – a time of the largest growth in federal enforcement activity in history – have been a
period of spectacular flux in both the allocation of power between Washington and U.S. attorneys’
offices and the extent of sentencing discretion allowed to federal judges. Conversations about the
two issues tend to stay on separate tracks. Yet although the two issues need not be linked, they
are, when it comes to federal prosecutorial discretion, two different ways of talking about some of
the same things.

When it comes to sentencing policy outside the federal context, the federal experience,
particularly when combined with the more equivocal evidence from Alaska and New Jersey
suggests that more attention should be paid to the way the existence of, or desire for, prosecutorial

72Bibas, Regulating, supra note __, at 139.

73See Frank O. Bowman, III, The Year of Jubilee . . . or Maybe Not: Some Preliminary
Observations about the Operation of the Federal Sentencing System After Booker, 43 Houston L.
Rev. __ (2006); Susan R. Klein, The Return of Federal Judicial Discretion in Criminal

centralization (or the lack thereof) influences sentencing policymaking. The suggestion here, albeit quite tentative, is that a statewide executive authority interested in promoting or maintaining a degree of control over local prosecutorial offices will be attracted to sentencing regimes that, say, encourage or require judges to scrutinize the congruence of plea bargains with case facts, or mandate other judicial monitoring measures. There may of course be countervailing considerations counseling against reliance on judges in this regard. But a state attorney general (or governor, or legislature acting at the behest of the attorney general or governor) will have few other options if the goal is limiting prosecutorial drift. As a general matter, state executive authorities are the dogs that haven’t barked very loudly when it comes to the overall management of the prosecutorial side of state criminal justice systems. That may well remain the case. But if it changes, the change will likely be seen on the sentencing side.

C. Toward New Issues of Intra-District Prosecutorial Control

The dominant pattern in state jurisdictions has statewide officials seeking – to various extents but always with less zeal than their federal counterparts – to restrict judicial sentencing discretion, but with little interest in restricting prosecutorial bargaining positions. These, they have left to the regulation of district authorities who, should they want to do so, have had recourse to screening and monitoring regimes of the sort deployed by D.A. Connick in New Orleans, with little need to draft judges to help in the endeavor. Such monitoring regimes, however, require a considerable degree of centralized control. And we are in period in which “community prosecution” approaches – which are inherently less centralized – are becoming all the rage.

As Catherine Coles noted: “Working closely with citizens who view their problems

74That many statewide policymakers have, for budgetary reasons among others, see Rachel Barkow & Kathleen O’Neill, Delegating Punitive Power: The Politics and Economics of Sentencing Commission Formation, __ Texas L. Rev. __ (2006) – become interested in using sentencing law to reduce overall incarceration levels and does not mean they are not also interested using sentencing procedures to limit prosecutorial concessions.

locally, by neighborhood, puts pressure on prosecutors to decentralize their operations. Many prosecutors are exploring how this can be achieved, even in the realm of screening and prosecuting cases." To the degree that prosecutors do move in this direction, we may well see a new interest in judicial monitoring, and in sentencing regimes that demand such monitoring, on the part of district attorneys. Even pretty flexible sentencing guidelines can promote a D.A. office’s “bureaucratic control over its staff of assistants,” as Jeffrey Ulmer and John Kramer found in their study of three offices in Pennsylvania. But with decentralization and consequent loss of alternative means of bureaucratic control, D.A.’s offices might become a lot more interested in less flexible sentencing regimes, and even in judicial monitoring. This is very much a story in progress.

Conclusion

Even as we try to figure out the best allocation of power between judges and prosecutors in the sentencing process, we ought not forget that prosecutorial interests are not monolithic and that for prosecutorial hierarchs, sentencing judges are as much potential monitoring partners as they are potential rivals. Whether judges will be enlisted (or, as in New Jersey, will volunteer) in this monitoring project will vary by jurisdiction and by the degree of prosecutorial centralization within the jurisdiction. Those who would engage prosecutorial policy-makers in a sentencing reform project ought at least be aware of this, and perhaps can even make use of it.

---
