Agency Slack and the Design of Criminal Justice Institutions

Aziz Huq
AGENCY SLACK AND THE DESIGN OF CRIMINAL JUSTICE INSTITUTIONS

Aziz Z. Huq

THE LAW SCHOOL
THE UNIVERSITY OF CHICAGO

March 2015

This paper can be downloaded without charge at the Public Law and Legal Theory Working Paper Series:
http://www.law.uchicago.edu/academics/publiclaw/index.html
and The Social Science Research Network Electronic Paper Collection.
Agency Slack and the Design of Criminal Justice Institutions

Aziz Z. Huq*

Forthcoming in,
The Routledge Handbook of Criminal Justice Ethics
(J.P. Jackson & J. Jacobs, eds. 2016).

Introduction

Crimes are investigated, prosecuted and adjudicated by agents of the state such as the police, prosecutors, judges, and juries. In all jurisdictions, the legal authority exercised by each of these agents is circumscribed by constitutional rules, statutes, regulations, or internal guidelines. Legal constraints are a way to ensure that individual suspects’ constitutional and human rights are not violated. But it is tolerably clear that agents do not always observe the letter or spirit of those constraints. Police, for example, can and do target suspects on the basis of impermissible criteria, such as race, perceived sexuality, or religion. They occasionally employ unlawful methods for extracting information, such as illegal searches or coercive interrogations. And when a criminal offense comes to be adjudicated, prosecutors, judges, and juries may violate legal rules due to bad motives, negligence, or simple ignorance. The result of such abuse, neglect and error is the impairment of highly esteemed constitutional and human rights to privacy, equality and liberty from arbitrary state action.

This chapter analyzes the problem of remedies for state wrongs in criminal justice institutions. The problem of remedies is analytically distinct from questions of what the substantive rules should be in the first instance. Jurisdictions differ in how much authority they assign to police, prosecutors, and other key actors in the criminal justice system. French prosecutors, for example, have considerably more discretion when it comes to interrogations and searches than their American or British counterparts (Bradley 2001). Even if the scope of legal constraint bears on how much a remedial mechanism is needed, it is nonetheless analytically feasible to bracket the substance of legal rules, and to focus narrowly on the distinct problem of remedies, and in particular how remedies for state wrongs are appropriately built into the design of criminal justice institutions. Due to my own unfortunate parochialism, I will draw largely, albeit not exclusively, on examples from the American context to explore this question, although the general framework offered here should be portable across national borders.

To be clear at the threshold, my concern here is not with any and all sorts of errors that occur in criminal justice’s administration. I am not, for example, concerned with routine bureaucratic inefficiencies and errors. Like any other large bureaucracy, a criminal justice system must develop regularized administrative systems to triage information and correct administrative problems (Gaines and Worrall 2011). At times, the breakdown of this administrative system can lead to violations of individual rights, for example when paperwork errors lead the police to target the wrong person (e.g., Herring v. United States 2009). I do not address the very real problems of sound administrative design in criminal justice systems here. Nor am I directly

* Professor of Law and Herbert and Marjorie Fried Teaching Scholar, University of Chicago Law School. Thanks to Mike Seidman for helpful comments.
concerned with the development of rational, evidence-based policing strategies (Sherman 2013), although these too might have a large effect on rates of abuse and neglect.

I am also not concerned here with instances of comprehensive political failure at the system designer level. Democracies and, more commonly, autocrat regimes often fail to install formally robust protections of basic rights in their criminal justice systems (see, e.g., Tevaskes 2007). It is also possible that a polity enacts rights into formal law but then systematically and persistently lacks political will to enforce those rights. There is no serious question that such political failures raise complex and difficult problems—problems that warrant often political solutions—but such problems are distinct from the phenomenon of institutional failure. To bring the latter into crisp focus, I limit my analysis to liberal democracies in which a democratic principal (e.g., a parliament or a legislative coalition acting in tandem with a chief executive) wishes to secure public safety but, at the same time, also ensure that its agents do not engage in serious forms of official malfeasance and neglect leading to rights violations. To the extent that democratic polities entrench rights in formal law, for example through constitutionalization, but falter in their implementation, it is quite plausible to talk of a rights-remedy gap and to conceptualize the problem using the economic terminology of principal-agent relationships. Under this description, there is a democratic principal that seeks to reduce agency slack among criminal-justice agents (Moe 1984). This entails minimizing rates of illegal searches, coercive interrogations, discriminatory policing and prosecutions, and serious violations of due process. This analytic approach, focusing on the formal institutional context of delegations to agents, is hardly novel, but is employed routinely in economics (Tirole 2009), political science (Goodin 1998), and law (Vermeule 2007).

Further, while my analysis sets to one side systematic political dysfunction, I do assume that the democratic principal has limited time, energy, and willpower to devote to monitoring, such that its goal is not the strictly optimal institutional arrangement—an ambition that may illusive—but rather a satisfactory arrangement. In decision theoretical terms, I ask what a satisficing democratic legislative would do (Byron 1998). For American readers, this perspective should be familiar: The U.S. Supreme Court crafts rules for police and other government actors when interpreting provisions of the U.S. Constitution’s Bill of Rights. Like my analytic perspective, the Court’s rule-making is superficially styled as somewhat sub specie aeternitatis, rather than the product of distinct historical interest groups and ideas. At least ideally, the Court is supposed to weigh how best to install remedies for constitutional violations given not only the likely responses of state agents within criminal justice systems, but also given its own institutional limitations.

To take this posture toward criminal justice institutions, and to ask about best how to remedy serious violations of the law, is to reject the view, famously associated with Oliver Wendell Holmes, that the law simply is a prediction of the legal consequences of an act (Holmes 1897). On the Holmesian view, to speak of rights and remedies as distinct is incoherent: The observed pattern of official behavior is all there is, and complaining about the law being unobserved is whistling in the wind. This is neither a necessary nor an appetizing view of the law here. Instead, it is clear that the law can meaningfully specify a behavioral desideratum related to individual rights, such as an absolute ban on torture or racial profiling, even if that rule is not observed in every (or even most cases). Such law reflects an independent normative judgment
about what sort of official behavior is desirable. Even if violated, law can still serve as a lodestar for institutional designers. Law can be an instrument to mitigate the stresses imposed by the transient preferences of legislators suffering from occasional lapses of political will. It can also influence unelected officials’ selection, incentives and beliefs (Besley 2006; Shleifer and Vishny 1998; Tirole 2009).

The analysis presented here cleaves into roughly two halves. In the first half of the chapter, I investigate why criminal justice institutions should include remedies for serious malfeasance and neglect in the first instance. Why do these rights violations even provide a cause for remedial design expenditures in the first instance? I identify both consequentialist and deontic grounds for acting. But even assuming such grounds matter, I consider why we need institutional design solutions, as opposed to political solutions, or other responses exogenous to the structure and basic operation of criminal justice institutions. In the second half of the chapter, I then define and explore two pivotal institutional design choices that a democratic principal must make. First, should policy responses or remedies be ex ante or ex post? And second, should it be public or privatized? After showing that all four combinations of these two choices exist in practice, I explore the strengths and weaknesses of each approach. That analysis suggests, at a bare minimum, that some mix of remedial features is probably warranted in the criminal justice context, although the precise mix will vary with local circumstances.

I. Why Does Institutional Design Matter in Criminal Justice?

Agency slack—malfeasance, neglect, gross error and other species of lawlessness—is in all likelihood endemic within not just most criminal system systems, but also most governmental entities to greater or lesser degrees. That fact alone does not mean that an institutional designer must include mechanisms to impose formal, legal consequences when legal rules are violated and important individual rights are infringed. Indeed, there are two reasons for withholding such institutionalized responses to serious rule-breaking. First, it may be that there is no welfarist or other normative justification for the provision of such remedies. Second, although a remedy might well be justified, it may not need to be a formal, legal one given the existence of alternative channels for corrective action, in particular in the political domain. Both reasons might lead to a rationing of remedial responses, as well as outright elimination of remedies. Neither of these reasons suffices, though, to vitiate the need to consider remedies in the design of criminal justice institutions.

A. Normative Justifications for Institutionalizing Remedies

Foundational legal instruments such as the U.N. Convention on Civil and Political Rights and the U.S. Constitution reflect a belief in universally held individual entitlements against certain kinds of state action, including torture, rape, discrimination on the basis of race, religion, or gender, and arbitrary deprivations of somatic liberty. These entitlements rest upon an overlapping consensus amongst consequentialist and deontic normative theories rooted in different cultural and religious contexts (Rawls 1989). All are at least imperiled, if not necessarily impinged, when the state exercises its Weberian monopoly on force to achieve public order and minimize violence. From this simple fact flows the need to design criminal justice institutions with agency slack in mind and remedies to hand.
It is not adequate, moreover, to postulate that if these rights are indeed foundational and the object of an overlapping consensus, then there will be no need to build a remedial architecture since criminal justice institutions, being necessarily embedded in society, will respect those values. This fails to account for the hardly improbable scenario in which individual police and prosecutors have preferences at odds with legal norms. Nor does it address the possibility that they may simply not wish to exert the necessary effort to ensure that rights are respected, or the possibility that rights violations might be an illicit means to achieve some otherwise legitimate end, such as success in the electoral process. Unanimity over social ends, in short, does not refute the possibility of divergence over means.

But perhaps such rights violations are, as Wilson (1968) intimates, simply the price of maintaining order. On this view, it suffices to focus on crime-control goals, while writing off rights violations as regrettable, but necessary, costs of ‘doing business’ as a Weberian state. For a number of reasons, this view is not persuasive. At a minimum, state agents who use their authority in a discriminatory fashion, or who brutalize rather than investigate, are not always or as a consequence engaged in efficient law enforcement. Hence, it would be a mistake to assume that there is necessarily a simple trade-off between public order and individual rights.

Further complicating the relationship between public order and rights is the large empirical literature showing that expected compliance with legal rules is a primarily a function of the public’s perceptions of state legitimacy. This literature further suggests that legitimacy is a normative judgment derived from an evaluation of whether state agents comply with norms of procedural justice (Tyler 2006; Tyler and Huo 2002). Contra rational-actor models, it also suggests that normative evaluation of the state provides a more powerful predictor of future conformity to law than expected sanctions. As a correlative, empirical research in Britain demonstrates that decreasing police legitimacy is associated with increasing acceptance of private violence (Jackson et al. 2013). Given these externalities from official law breaking, the costs of rights violations are not well captured by a narrow focus on the discrete state-citizen transaction.

Especially powerful evidence that official malfeasance can have corrosive spillover effects derives from evidence of how race and policing interact in the American context. The U.S. criminal justice system is characterized by large racial disparities at both the policing stage (Epps et al. 2014) and the incarceration stage (Western 2006). Whether motivated by individual bad motives or structural forces, pervasive patterns of racially disparate policing and incarceration powerfully communicates a message of social stigma and hierarchy, one that resonates with America’s long and recent history of racial subordination (Epps et al. 2014; Goffman 2014). That message, at least one study suggests, may have criminogenic effects (Unnever et al. 2009). Stigmatic and demoralization effects have been identified as following from the disparate treatment of other ethnic minorities (Tyler et al, 2010).

Racial disparities in criminal justice do not merely impede public order goals. In the U.S., they also have large collateral repercussions on social outcomes. Those disparities have been found to deepen economic inequalities by systematically limiting blacks’ access to employment (Pager 2007) and by diminishing income over the life course (Kerney et al. 2004). As a result,
disparate enforcement of the criminal law also impoverishes black communities in particular (Clear 2009), and leads to political disenfranchisement through subtle demoralization effects, not just due to incarceration but also due to contact with police (Lerman and Weaver 2014). Perhaps most troublingly, those disparities reproduce and deepen historical patterns of social and economic inequalities across generational lines. The one in four black children within a cohort born in the 1990s who experienced paternal incarceration suffered sharp increases in infant mortality, more mental and behavioral problems, and also higher rates of homelessness, all as a probable result of their father’s incarceration (Wakefield and Wildeman 2013). Finding effective remedies for racial discrimination in law enforcement therefore has the potential to generate not only a more just state, but also gains to public order and social equity.

B. Alternatives to Institutionalizing Remedies

The second strategy for resisting formal legal remedies in the design of criminal justice institutions would not deny either the deontic or the consequentialist case for respecting rights against state violence and discrimination. It would instead insist that the solution for these problems is exogenous to the design of criminal justice institutions. Rather, the argument would go, the democratic framework of elections provides ample opportunity for retrospective accountability exercised through the ballot. Just as voters use the franchise to punish other abuses of state power, so too they can deploy it to discipline criminal justice actors. This is a serious argument. Impressive theoretical and experimental evidence supports the underlying hypothesis of retrospective voting (see, e.g., Powell and Whitten 1993; Woon 2012; but see Berry and Howell 2007 for some cautionary countervailing evidence). In the U.S. Constitution, one key element of constitutional design—the sweeping power of presidents to pardon crimes, perhaps even their own (Kalt 1996)—also implies a general reliance on retrospective voting, rather than tailored remedies, as a strategy for controlling agency slack in criminal justice. Should not the same hold for police, prosecutors and their ilk?

Some have suggested that the availability of retrospective accountability in the democratic sphere already inflects the institutional design of criminal justice institutions. Barkow (2008), following Davis’s canonical treatment (1980), points out in the American context that prosecutors possess a uniquely large measure of discretionary authority. Departing from Davis, Barkow postulates that this discretion is viewed as normatively unproblematic because prosecutors are uniquely subject to democratic control via elections. In other work, Barkow has suggested she shares Davis’s concern with the breadth of prosecutorial discretion (2009, 2013), but it is worth considering, and rejecting, the possibility that electoral accountability of prosecutors should suffice as a mechanism to regulate rights violations in criminal justice.

A threshold reason for skepticism of political responses to discrete wrongs is a mismatch in scale between the problem and the putative solution. That is, rights violations in the criminal justice context are usually individualized in effect, and dispersed in space and time. Political accountability mechanisms, and especially elections, are by contrast periodic and bundled in nature: An electoral is asked to evaluate many policy decisions via a single vote (Gersen 2010). Public-order maintenance is often the responsibility of city mayors, who are typically responsible for a heterogeneous array of distinct policies. It is certainly possible to point to examples in which violations by criminal justice officials furnish the central point of contention in an
election—consider the New York City election of 2013, where the “stop and frisk” policy of the New York Police Department (NYPD) was a pivotal object of public debate—but such instances seem the exception rather than the rule. At a minimum, the all-or-nothing character of democratic elections mean that electorates have no way to punish low to moderate rates of rights violations.

Stepping back and accounting for the historically contingent aspects of politics only deepens the case against any reliance on electoral checks. At least in pervasively racially divided societies such as the United States, where minorities bear a disproportionate share of the costs from public-order maintenance (Alexander 2012), it is hardly clear that majoritarian institutions can be relied upon to generate sufficient remedies. To the contrary, in the U.S. context at least, recent empirical work suggests that the polity may be all too content that the costs of crime control fall on groups widely viewed as less deserving of compassion and solicitude (Enns 2014). Unnever and Cullen (2010), for example, demonstrate that public preferences for punitive policies are predicted by negative views of racial minorities, and not fear of crime. Similar factors predict punitive sentiment in the British context (King and Maruna 2009). The problem of electoral bias is compounded by systematic epistemic deficiencies in public policy debates. Pettit (2012) shows that the young, low-skill black men who are overrepresented in the criminal justice system, are also “categorically and systematically” excluded from the data used to frame social policy. Finally, Garland’s (2002) pathmarking analysis of the ideological framings of crime policy implies that majoritarian institutions will not respond rationally to perceived crime risk in ways that mitigate effectively the rights-related consequences of criminal justice (see also Beckett 1997). The democratic politics of crime, that is, may be distinctively (if not uniquely, see Gilens 1999) opaque to non-racial conceptions of justice.

These objections to reliance on political remedies hinge on the distinctive role of race in the criminal justice context. Another argument, developed by Marie Gottschalk (2008, 2014), focuses on the path-dependent development of “carceral politics” in the United States. Gottschalk points out that punishment-related policies in the American context are characterized by a positive feedback effect, one that both entrenches existing punitive policies against reform, and also catalyzes the subsequent adoption of even more punitive policies. The positive feedback loop arises from the complementary operation of two dynamics. On the one hand, she points out, prison unions and the prison industry have become increasingly powerful lobbies. On the other hand, legislative appointment rules that assign prison populations to districts with prisons, rather than districts of original residence, deflate urban communities’ influence at the expense of rural communities’ power. This political economy means that correcting excessively punitive policies, to say nothing of discrete rights violations, becomes more difficult with time, even as the scale of criminal justice institutions—and thus the numerical volume of related errors and abuses—expands. It is worth noting that Gottschalk’s powerfully pessimistic argument for the positive feedback loops in the carceral state coexists with more optimistic prognoses (Clear and Frost 2013).

The argument developed here against reliance on political checks in the criminal justice context might be pressed even further, so that it yields a critique of my analytic strategy in this chapter. The critique, which has a familial resemblance to the Holmiesian position rejected earlier, goes as follows: If democratic mechanisms are inadequate responses to retail rights
violations in the criminal justice context, it is because there is insufficient democratic will to mitigate them. And if that is so, then the whole project of institutional design to minimize such costs is misbegotten. It will always be the case that majoritarian preferences prevail, and minorities fold, by dint of the sheer weight of numbers (an argument anticipated most eloquently in Stephen 1873). The strong do what they can, and the weak bear what they must,

I do not think this critique ought to succeed for several reasons, but delineate here only one response: Even if elective democratic institutions are not persistently attentive to the rights-related costs of crime control, this does not mean they are never so responsive. It seems plain that there are at least moments of broadly shared concern. Moreover, it also seems plausible to think that in those moments elective institutions may choose to install durable institutional responses to agency slack and abuse in criminal justice. Without succumbing to whiggish fantasia of inexorable progress, it is surely plausible to think that sometimes we work improvements in our institutions.

In the American context, the task of institutional repair fell first and foremost to the U.S. Supreme Court under Chief Justice Earl Warren, which installed a suite of rights under the Bill of Rights to the U.S. Constitution through path-marking rulings in the 1950s and the 1960s (Amar 1998). To be sure, these rulings have proven to be vulnerable to erosion and backlash (Steiker 1996), and may have had the counterintuitive effect of legitimating a certain volume of rights violations (Seidman 1992). Any institutional remedy can be limited by legislation, or it can be defunded to the extent it depends on the public fisc. But an institutionalized remedy coupled to some backsliding may still be more desirable than no institutionalized remedy at all. To foreclose the possibility of democratically catalyzed responses to the agency and abuse problems of crime control is either to succumb to a Nirvana fallacy or to adopt a mistakenly static view of political institutions. One can, in short, be generally pessimistic about democratic institutions, while still insisting upon, and perhaps working to realize, their occasional capacity for deep and enduring good.

There is a second way of arguing against the need for formal legal institutional responses to agency slack and abuse in the criminal justice system. Setting to one side the electoral mechanisms of control and sanctioning discussed above, it might be argued that formal legal remedies of the sort considered below are unnecessary given the correct mechanisms for selecting criminal justice officials. That is, assuming the correct sorts of screening and selection mechanisms are in place, the task of the democratic principal designing criminal justice institutions is at an end.

Yet again, this alluringly easy response is also insufficient. To begin with, it seems clear that selection, as well as retention and promotion, rules will play an important part in any institutional strategy to address agency slack, lawlessness and error. An institution that persistently hires individuals who are unable or unwilling to account for important goals is unlikely to achieve these goals. Nevertheless, there are several reasons to resist the temptation to rely exclusively on personnel-related instruments. As discussed below, criminal justice systems (police, prosecution offices, and courts) tend to be large organizations. As the size of an organization expands, the more costly and difficult it is at the margin to find good personnel. At the outer margin of organizational growth, it is likely to become increasingly difficult to hire
effective personnel. Simply put, it is easier to find the first good police officer than the one thousandth.

In addition, employment decisions are characterized by asymmetric information, and, as a result, a need for employees to find a signal that separates good from bad types in the candidate pool. This signal is a quality good types find less expensive to acquire than bad types (Spence 1973). At least with respect to the propensity to violate rights, it is not clear what that signal would be. One possibility, at least for prosecutors, is that lawyers who had been more successful in legal education and during the early stages of their legal careers are more likely to be highly motivated and hence effective prosecutor. Empirical studies of prosecutorial choice in the U.S. context, create complications for this account by demonstrating that prosecutorial ambition tends to distort charging decisions, by eliciting indictments that are likely to gain public attention rather than social gains (Rasmusen et al. 2009; Glaeser et al. 2000)

Finally, even if there is a signal that can be employed to sort employees who will respect rights from those less likely to do so, that signal may be undermined or distorted by other dynamics within the hiring process. Consider here the example of police. Generally, police officers will be permitted to use force in situations when individuals are not so allowed (Harmon 2008). That is, electing to join the police is a way of enlarging one’s right to use force. There is likely variance within the populace in respect to the taste of violence. An implicit wage of police employment, therefore, is the license to use violence that would otherwise be unlawful (see Heywood et al. 2007 for the concept of an implicit wage). It follows that police employment will be more attractive—because it carries a larger net explicit and implicit wage—for those in the population who have more of a taste of violence. Of course, whether this implicit wage effects hiring, or whether it is offset by other factors (e.g., the implicit cost accepted by police of entering stressful and abusive situations), is an empirical question. Resolution of that question will nevertheless at least complicate, and perhaps seriously distort, the operation of signaling in the job market for police.

Finally, it is worth noting here that conduct rules for state actors might also interact with efforts to select good types only for criminal justice institutions. Rules for the use of force, for example, may influence the composition of police departments (Vermeule 2004). That is, the larger the margin of discretionary authority police have to use force, therefore, the larger the net wage differential between those with and those without a taste for violence. The more discretion police have to use force, therefore, the more the job will disproportionately attract those least inclined to use it wisely. The ensuing tension is perhaps sharpest when a jurisdiction faces significant public order problems.

In short, it is not persuasive to write off abuse, neglect, and malfeasance as merely the price of engaging in crime control. Nor is it sufficient to relegate the management of those costs to the democratic political process, or to hope that the right people will somehow always fill critical posts. Instead, just as in the corruption context (Shleifer and Vishny 1998), there is a need for more particularized attention to institutional design, and the role of remedies therein.

II. The Institutional Circumstances of Criminal Justice Remedies
If institutional rather than political remedies are warranted, then what form should they take? There is likely no global answer to this question. The political, sociological, and historical circumstances of criminal justice diverge widely from jurisdiction to jurisdiction. The ensuing variance precludes across-the-board responses. Rather, the best way to understand how a democratic principal might go about regulating her criminal agents so as to limit rights violations is to specify the institutional circumstances that characterize criminal justice in ways that are salient to remedial design. With these firmly in view, we can start to think most clearly about remedy-related design choices.

Most importantly, the delivery of criminal justice is almost always a highly dispersed phenomenon. Unlike other state functions such as legislating or fashioning monetary policy, the delivery of criminal justice necessarily occurs in a large number of geographically dispersed sites. This is most obviously the case with policing, which is necessary decentralized down to the street level. But it is also true of the prosecutorial and adjudicative functions. In the United States, for example, most prosecutions are initiated by state and local actors, not federal U.S. attorneys (Glaeser et al. 2000). Each county or city may have its own prosecutor, one who often serves only upon election, as well as its own system of criminal courts. The next result is a great deal of necessary variance in the preferences and behavior of prosecutors and criminal judges across a wide variety of physical sites for the delivery of criminal justice outcomes. In addition to this, theories of policing emphasize the need for only “loose coupling” between an agency’s goals and the day-to-day actions of its officers (Crank 2003). Given this complexity and unpredictability, that is, a space between instruction and application is almost inevitable.

Geographic dispersion and decentralization has a number of implications. As a threshold matter, it means that a democratic principal cannot easily maintain ongoing surveillance over its agents in the criminal justice domain. Agency relations are almost always characterized by informational asymmetries (Moe 1984). But the epistemic gap in the criminal justice context will tend to be especially large for a number of reasons. Most importantly, many criminal justice encounters occur with only agents of the state and putative victims present. If a legal violation occurs, it is usually brief or instantaneous but at the same time generate evidence that will endure into the adjudicative process. That enduring evidentiary trace, however, usually not indicate whether it was secured through illegal means. Determinations of what happened when police use force—or when a confession is elicited in the station-house, or when prosecutors squeeze out a plea deal—require an evaluation of competing testimony from state agents and putative victims. Those determinations, even if reached through the formalized mechanism of the individual criminal trial, must be based on often systematically untrustworthy information. As a result, it will almost never be enough simply to announce a rule and expect it to be self-executing: The likelihood of detection—which is the main determinant of deterrence effects (Nagin 2013)—is inevitable small, or practically nonexistent.

Worse, variation in local conditions means that observed fluctuations in official behavior cannot always be immediately characterized in positive or negative terms. For example, American jurisdictions close to the Mexican border often have “fast track” programs with streamlined procedures and reduced punishments for immigration-related offenses. These programs are responses to a local problem. They may be either ranked as pragmatic accommodations or problematic deviances from the law, depending on one’s baseline (Bibas
Decentralization has other implications for the institutional context of criminal justice, particularly in large polities such as the United States. There, geographic variation is associated with wide fluctuations in fiscal arrangements. As Thatcher (2011) has demonstrated, there can be an order-of-magnitude difference in the fiscal resources available to police in a given U.S. jurisdiction. Further, in many jurisdictions, law enforcement budgets are a function of the magnitude of assets seized pursuant to criminal forfeiture laws (Baicker and Jacobson 2007). Again, the presence of these mutable background conditions of policy implementation likely impinge on officials’ incentives and also complicate efforts to identify certain forms of official behavior as per se problematic, at least on an ex ante basis.

In summary, criminal justice presents a particularly acute epistemic problem from the perspective of a democratic principal due to decentralizations, wide variation in local conditions, and the weak epistemic signals. These are not the only constraints on monitoring in law enforcement. For example, Richman (2003) has pointed out that prosecutors and law enforcement agents in the United States often operate in collaborative teams. This implies the existence of a team production problem for a principal, where the bundled nature of outputs makes it difficult to reach separate judgments about each participant in the team (Holmstrom 1982). Even without accounting for these additional concerns, it should be tolerably clear that any mechanism to identify and remedy abuse and neglect in the criminal justice face faces considerable epistemic hurdles.

III. Two Design Decisions in Institutionalizing Remedies

The acute epistemic difficulties a democratic principal faces in designing remedies in the criminal justice context render two distinct institutional choices especially critical: the election between ex ante and ex post remedies, and the distinction between public enforcement mechanisms and their private analogs. These design margins matter because selecting between them can dramatically alter the magnitude and accuracy of information available for remedying abuse and neglect. Altering these parameters furthermore has implications for the efficacy of a remedial regime. To illuminate these diverse implications, I first set forth the two design margins, and then identify some consequences of the choices they imply. As will become clear, most of the examples I supply from the American context concern remedying police misconduct for the simple reason that American law does not bring significant institutional resources to bear on the problem of prosecutorial or judicial misconduct (Barkow 2010). In response to that shoal of wrongs, American law maintains a profound silence.

A. Ex ante versus ex post remedies

Consider first the choice between ex ante and ex post remedial mechanisms. American law resorts to both options. On the one hand, the Fourth Amendment to the U.S. Constitution requires warrants for most home searches. The warrant rule is a procedural check: It does not directly speak to the scope of the search, or to the evidentiary justification required before an officer can proceed. Instead, it operates as a sort of licensing scheme, inserting a independent
magistrate before a legal violation occurs rather than after the fact (Stuntz 1991). The prophylactic effects of licensing emerge, to be sure, partly from the evidentiary and explanatory burdens imposed on police by the warrant rule. But the warrant system also helps limits illegal searches because it operates as a “costly screen” (Milgrom and Roberts 1986). In simple terms, it imposes an effort tax on police, forcing them to “stop and think” whether a search is really appropriate, lawful and needful (Bar-Gill and Friedman 2012). That cost incentivizes police to choose among potential searches, and to proceed only with those that are truly worthwhile.

Another ex ante remedy is the requirement that a suspect be permitted to have legal representation prior to being included in a post-charge police line-up (Wade v. United States 1967). In the interrogation context, counsel does not need to be provided, but suspects are given a warning of their right to have counsel present, as well as their right to remain silent without that being used against them at trial (Miranda v. Arizona, 1966).

Ex post remedies come in at least three flavors under U.S. law. First, policing illegality often has trial consequences. Starting in the United States, but increasingly across the world, physical and testimonial evidence obtained in violation of the law will be excluded from a subsequent criminal trial. This is done at least in part to elicit compliance with ex ante rules such as the warrant requirement. Bradley (2001) argues that main difference between the U.S. rule and other jurisdictions’ is the latter’s mandatory character (Mapp v. Ohio 1961). Recent expansions to a ‘good faith’ exception to the exclusionary rule in the American context (see, e.g., Davis v. United States 2011), however, may be narrowing that gap. Another class of illegality that has a trial remedy is courtroom misconduct by prosecutors, which is often addressed through the use of, arguably fictive, curative function of instructions from judge to jury (Sklansky 2013).

Second, federal law provides a damages remedy for individuals whose constitutional rights are violated. Damages actions against state actors proceed under a late nineteenth century statute, now codified as 42 U.S.C. § 1983, while the damages action against federal officials has been recognized as matter of constitutional common law (Bivens v. Six Unknown Named Agents 1971). The law of constitutional torts is complicated by the existence of “qualified” immunity, which police enjoy, and “absolute” immunity, which prosecutors often and judges always enjoy (Rudovsky 1989). Under qualified immunity doctrine, liability does not attach unless it would be very clear to a reasonable officer that he or she was violating a constitutional norm. Under absolute immunity, even the most egregious of errors do not generate liability. In consequence, prosecutors and trial judges can almost never be held liable for money damages, even for the most egregious and harmful species of constitutional violations. Even in fairly extreme instances of discrimination and abuse rising to the level of torture, monetary relief therefore tends to be only rarely available (Huq 2009).

Finally, American law has adopted the English common-law remedy of habeas corpus, and transformed it into a vehicle for the postconviction review of criminal convictions for constitutional error. A prisoner seeking habeas relief identifies a constitutional error in their trial, and seeks a vacatur of the conviction or (in capital cases) the sentence. Most frequently, the prisoner will focus upon alleged trial errors such as defense counsel that failed to meet constitutionally required standards of competence or prosecutorial failures to produce exculpatory evidence. The state then faces the option of releasing the prisoner or retrying them. Since the 1970s, postconviction habeas has accrued a positively baroque encrustation of
procedural rules that operate to titrate out in small increment (for a summary, see Huq 2014). For example, police violations of Miranda can be raised on habeas, but unlawful searches and seizures cannot. The net result of these rules, which are a product of both legislative hostility to prisoners and judicial hostility to the habeas caseload, is that federal courts must assign a large amount of effort applying complex procedural rules to pro se prisoner pleadings, some of which might be meritorious, even though those same rules virtually guarantee that the rate of merits success will be de minimus. Absent major reform, it is hard to see how the postconviction review stands on a rational or defensible footing.

B. Private versus public enforcement

The second important design margin is the election between private and public remedies. All of the aforementioned ex post remedies exception for suppression motions are private, in the sense that it is a private individual (typically the person who has suffered constitutional harm), who must file suit and prosecute a case against the state. It is also possible to imagine an ex ante private remedy in the form of an injunction, sought on behalf of a given individual or a class, against a certain kind of state conduct, and enforced via contempt sanctions. As a matter of practice, ex ante private remedies are rarely observed. In 1983, the U.S. Supreme Court held that person lacked “standing” (i.e., a constitutional right to sue) to enjoin a police chokehold policy in the absence of sufficient evidence that he would be subject to a future chokehold (City of Los Angeles v. Lyons 1983). Because police and prosecutors tend not to announce whom they intend to target with illegal measures before the fact, the 1983 decision in Lyons staunched the flow of private ex ante challenges. After Lyons, therefore, class actions against police and prosecution offices are rare, although, as recent litigation challenging the NYPD’s stop-and-frisk policy shows as racial discrimination, not impossible (Kalhan 2014; Steiker 2013; Geller and Fagan 2010).

A key characteristic of all private remedies is that their use turns on the initiative—and by implication the legal, epistemic and fiscal resources—of the individual harmed by the state. Given inequalities in the distribution of legal knowledge and resources across the populations, in practice this means that the ability to deploy a private remedy turns on the quality of counsel that a person can obtain, or that the state furnishes. American law permits some contingent fee arrangements and occasionally imposes fee shifting, but distributional factors remains salient.

A public remedy, by contrast, is one sought and administered by the state. Ex ante state intervention and supervision of a sort is at stake in warrant rules. It is possible to read the U.S. Supreme Court’s rules requiring warrants for some classes of searches as a departure from the text of the Fourth Amendment, which seems to limit the use of warrants rather than requiring them for any class of cases (Amar 1998). On that view, the Court suo moto has delegated to the judiciary at large responsibility for policing the state’s use of search authority ex ante. That delegation is public in character because it comprises a requirement by one state actor imposed on another state actor, and operationalized without any triggering by a private party. As with any other delegation, there is an attendant risk of agency slack.

Ex post public remedies, at least of a certain kind, are more common. On the one hand, civil and criminal suits against individual police or prosecutors for constitutional violations are
technically available, but in practice are rarely observed (Barkow 2010; Harmon 2010). Perhaps unsurprisingly, prosecutors are not disposed to investigate or charge peers or law-enforcement colleagues. The recent failure of state and federal prosecutors to bring charges against the police officer responsible for the shooting death of Michael Brown in Ferguson, Missouri, is symptomatic of this. More generally, there are remarkably few instances in which police or prosecutors are subject even to internal discipline either for constitutional violations or for culpable obfuscation of such violations.

On the other hand, a powerful and even more wide-ranging public remedy is available: the “pattern or practice” lawsuit that the federal Department of Justice can lodge against a municipal or county police department (Harmon 2010). Enacted in 1994, the statute, located at 42 U.S.C. § 14141, allows the federal government in effect to place police departments in receivership in the wake of repeated serious constitutional violations, and to impose new management, training, and internal rule-making §14141 requires the Department of Justice to maintain statistics on rates of police violence and excessive force. The ensuing suits, if relatively rare, can force local municipalities to prioritize policing investments, install ongoing monitoring mechanisms to elicit rule-compliance from frontline officers, and provide police leadership with political cover for reform.

As a final caveat, the public/private distinction refers to ideal types. Public defenders, who are funded by the state but tasked with acting on behalf of private individuals, arguably reflect a hybrid category. In practice, it is also possible to imagine yet other mixtures of private and public enforcement. For example, notionally private law suits seeking institutional reform can in practice operate as sites for negotiation and joint action by both private and public actors aimed at reforming problematic criminal justice institutions (Sabel and Simon 2004). And extending those possibilities, Gilles (2000) argues that § 14141 should be modified to allow for the delegation of enforcement powers to private individuals, who could sue in lieu of the Department of Justice.

IV. Trade-Offs in Remedial Design in Criminal Justice Administration

How should a democratic principal select between the diverse options—public versus private, and ex ante versus ex post—developed here? To anticipate my conclusion here, there is no one right answer: In most instances, some mix of remedial pathways will instead be desirable. Not all pathways, however, will be utilized, and under inevitable conditions of institutional scarcity, some remedies must be prioritized over others. But which? A starting point for analysis is how different remedial strategies measure up to the distinctive epistemic problem faced by a democratic principal in the criminal justice context.

Consider first the distinctive ex ante and ex post remedies have different epistemic strengths and weaknesses. On the one hand, it would seem obviously true that ex post remedies will be epistemically superior to ex ante remedies: The action in question has occurred, and hence can be investigated and clarified when the remedy is ex post. In the ex ante context, the regulator or judge must necessarily guess at the consequences of a state action. But this may not be so. Official decision-makers may suffer from cognitive biases that cause them to make persistent Type I errors. For example, in American law the ex post regulation of illegal search
occurs in the context of pretrial suppression hearings in which a judge is confronted with illegally obtained, but inculpatory, evidence. Even aside from the judiciary’s likely inclination to align itself with other official actors such as the police, judges will tend to view the suppression question through a cognitive lens distorted by hindsight bias, which presses toward the conclusion that police action was surely reasonable because it did, in fact, yield good evidence (Stuntz 1991). Worse, police may be inevitably tempted in suppression hearings to at best color their testimony to render their own actions seem lawful, or, worse, engage in outright perjury so as to prevent a perceived offender walking free. Given these systematic distortions in the ex post remedial context, the ex ante device of warrants may in fact be preferable even if epistemically more pinched.

The choice between ex ante and ex post rules also has selection effects, especially when the ex post remedy is available only in the criminal trial context. A thoroughgoing ex ante system for intrusive and coercive—say, that required a warrant or some other sort of approval, whether from a judge or a departmental superior—operates equally in cases where the object of state attention is innocent or guilty. Indeed, the very purpose of a costly screen is to elicit the decision-maker to leverage more fully available information to determine when an intrusive state action is in fact warranted (Milgrom and Roberts 1986). Ex ante screening may also prevent actions that would otherwise have been illegal, without any large spillover deterrence on lawful searches. The net result of an ex ante regime, therefore, will tend to a lower rate or Type I errors, at least in comparison to a world without remedies. Further, the suppression remedy used ex post for illegal searches and coercive interrogations selects for only the guilty objects of state attention, while leaving innocent objects of police attention without a remedy. Faced with a stream of culpable and unsympathetic victims of state depredation, a judge may be ill-disposed to fashion capacious privacy or dignity protections. That same judge, however, will never see the stream of cases in which innocent citizens are subject to state harassment or coercive prosecution threats. Exclusive reliance on ex post remedies, therefore, may both yield a distorted body of law and a radically incomplete remedial landscape.

Similarly, the choice between public and private remedies is not as clear as might first seen. On the one hand, the geographically dispersed operation of criminal justice institutions means that information about law violations reposes in the first instance in private hands (although it remains to be seen whether the emergence of algorithmic ‘big data’ tools could change that). Private remedial litigation affects a kind of Hayekian aggregation. Moreover, individuals have stable and strong incentives to pursue remedies at least once a violation has occurred. The epistemic advantage of private enforcement tools, on the other hand, may be outweighed by distortive wealth effects. Invocation of adversarial corrective mechanisms turns on the availability of fiscal and cognitive resources on the victims’ part. Where background allocations of such resources are unequal, an adversarial mechanism is likely to yield highly uneven and even inequitable results (Defains and Demougin 2008; Stuntz 2008). Where background inequalities in resources correlate with racial or ethnic stratification, as in the United States, it may well be that reliance on private enforcement has a perverse effect, especially where publically funded defense counsel tend to be under-resourced: It would mean that the expected rate of corrective process when an African-American’s rights are violated will be lower than when a Caucasian person’s rights are violated. Under these (rather too plausible) conditions,
private enforcement mechanisms create a tax that presses officials away from racially equitable policing.

Yet the case in favor of public enforcement is not without significant weaknesses either. Public enforcement renders corrective process vulnerable to ‘capture’ by interest groups hostile to subordinated groups that may be more frequently targeted for wrongful criminal justice attention. That is, well-organized interest groups such as law enforcement may expend resources gaining control over oversight mechanisms. Hence, it is perhaps no surprise that former prosecutors numerically dominate federal courts (Kalhan 2014). Even in the absence of such capture, public enforcement may be vulnerable to the political cycle and the associated fluctuations in public priorities. Some administrations, often on the right of the political spectrum, are likely to be less sympathetic to the classes of individuals who tend to suffer more frequently from legal violations in criminal justice contexts. But there is no reason to believe that partisan cycles of officeholders will correspond to peaks and troughs in the demand for remedies in the criminal justice context. To the contrary, reliance on public enforcement alone may be perversely countercyclical, if administrations most prone to remedial underenforcement are also most prone to violations in the first instance.

Conclusion

The bottom line of this analysis is simple: Institutionalized remedies for serious abuse and neglect in the criminal justice context are necessary yet difficult. On the one hand, political and selection-based remedies are bound to fail. On the other hand, any effort to embed remedial resources within the criminal system faces substantial epistemic challenges given the dispersed and hard-to-monitor ways in which police, prosecutors, and trial courts work. In that regulatory enterprise, a democratic principal must select among ex ante and ex post options, and also between public or private remedies. A pure strategy—i.e., one that is all public or all private, or one that is all ex ante or all ex post—is likely to fail. Each such approach has substantial drawbacks standing alone. In consequence, a mixed approach, that draws in diverse tools in response to the observed distribution of abuse and neglect, as well as the expected responses by regulated actors, is likely necessary. The framework articulated in this chapter may guide the analysis of mixed approach, but it cannot resolve the local question of the satisfactory mix for a given jurisdiction—let alone the problem of how to install that mix by legislation or court decision. On these counts, the design of remedies in the criminal justice context necessarily presents persistently difficult questions to which no clear or quick response is feasible.
Bibliography


**Cases**

*City of Los Angeles v. Lyons*, 461 U.S. 95 (1983)
Readers with comments may address them to:

Professor Aziz Z. Huq
University of Chicago Law School
1111 East 60th Street
Chicago, IL 60637
huq@uchicago.edu
402. M. Todd Henderson, Voice versus Exit in Health Care Policy, October 2012
404. Lee Anne Fennell, Resource Access Costs, October 2012
405. Brian Leiter, Legal Realisms, Old and New, October 2012
407. Brian Leiter and Alex Langlinais, The Methodology of Legal Philosophy, November 2012
408. Alison L. LaCroix, The Lawyer’s Library in the Early American Republic, November 2012
409. Alison L. LaCroix, Eavesdropping on the Vox Populi, November 2012
411. Alison L. LaCroix, What If Madison had Won? Imagining a Constitution World of Legislative Supremacy, November 2012
413. Alison LaCroix, Historical Gloss: A Primer, January 2013
415. Aziz Z. Huq, Removal as a Political Question, February 2013
416. Adam B. Cox and Thomas J. Miles, Policing Immigration, February 2013
417. Anup Malani and Jonathan S. Masur, Raising the Stakes in Patent Cases, February 2013
421. Lior Jacob Strahilevitz, Toward a Positive Theory of Privacy Law, March 2013
422. Eric A. Posner and Adrian Vermeule, Inside or Outside the System? March 2013
426. Lee Anne Fennell, Property in Housing, April 2013
427. Lee Anne Fennell, Crowdsourcing Land Use, April 2013
430. Albert W. Alschuler, Lafler and Frye: Two Small Band-Aids for a Festering Wound, June 2013
432. Aziz Z. Huq, Tiers of Scrutiny in Enumerated Powers Jurisprudence, June 2013
438. Brian Leiter, Nietzsche against the Philosophical Canon, April 2013
441. Daniel Abebe, One Voice or Many? The Political Question Doctrine and Acoustic Dissonance in Foreign Affairs, September 2013
442. Brian Leiter, Why Legal Positivism (Again)? September 2013
443. Nicholas Stephanopoulos, Elections and Alignment, September 2013
444. Elizabeth Chorvat, Taxation and Liquidity: Evidence from Retirement Savings, September 2013
445. Elizabeth Chorvat, Looking Through' Corporate Expatriations for Buried Intangibles, September 2013
446. William H. J. Hubbard, A Fresh Look at Plausibility Pleading, March 2015
448. Lee Anne Fennell and Eduardo M. Peñalver, Exactions Creep, December 2013
449. Lee Anne Fennell, Forcings, December 2013
451. Nicholas Stephanopoulos, The South after Shelby County, October 2013
453. Tom Ginsburg, Political Constraints on International Courts, December 2013
454. Roger Allan Ford, Patent Invalidity versus Noninfringement, December 2013
459. John Rappaport, Second-Order Regulation of Law Enforcement, April 2014
460. Nuno Garoupa and Tom Ginsburg, Judicial Roles in Nonjudicial Functions, February 2014
461. Aziz Huq, Standing for the Structural Constitution, February 2014
462. Jennifer Nou, Sub-regulating Elections, February 2014