Social Problems, 2024, **XX**, 1–15
https://doi.org/10.1093/socpro/spae060
Advance access publication 10 October 2024
Article

Assessing the Risks of Risk Assessments: Institutional Tensions and Data Driven Judicial Decision-Making in U.S. Pretrial Hearings

Sino Esthappan

Northwestern University, USA

ABSTRACT

Risk assessments, which are predictive technologies designed to augment organizational decision-making processes, are expanding globally. The use of risk assessments may lend legitimacy to official actors who routinely adjudicate highly consequential decisions alongside competing institutional pressures. But because these tools are laden with measurement errors, using them may also magnify institutional tensions and erode the legitimacy of official actors' decision-making practices. In this article, I use interviews with judges in four large U.S. criminal courts to reveal how they strategically engaged with risk assessment scores to navigate tensions within and among different institutional logics. In pretrial hearings, judges selectively invoked risk scores to legitimate punitive sanctions that mitigated tensions from bureaucratic logics to process high caseloads with limited resources, legal logics to protect public safety and impose the least restrictive conditions, and political logics to follow the law while facing public scrutiny. I discuss the implications of these findings for future research on penal change and the uses of discretion in the age of big data.

Key words: risk assessments; algorithms; decision-making; criminal courts; punishment.

Risk assessments are algorithmic tools that computationally analyze large volumes of digitized records to quantify the likelihood of future events (Burrell and Fourcade 2021; Christin 2017). They are increasingly used in institutions such as education, healthcare, and punishment to aid or inform official actors as they make highly consequential decisions. While risk assessments are generally regarded as objective and authoritative tools, the scores they produce can misrepresent risks in ways that undermine their perceived legitimacy and thus the legitimacy of the actors who use them (Espeland and Stevens 1998; Kellogg, Valentine, and Christin 2020; Porter 1995). How do these challenges to the legitimacy of risk scores shape how official actors use them to make highly consequential decisions?

In this article, I examine how criminal court judges' concerns about legitimacy affect how they use risk assessment scores to make pretrial release decisions. Each year in the United States, nearly five million jailed criminal defendants who are legally presumed innocent but often assumed guilty

I thank Miguel Chavez, Steven Epstein, Wendy Espeland, Ryan Fajardo, Sarah Lageson, James Mahoney, Karin Martin, Robert Nelson, Laura Beth Nielsen, Mary Pattillo, Wayne Rivera-Cuadrado, Kris Rosentel, Devin Wiggs, Karin Yndestad, my colleagues at the American Bar Foundation, the participants of Northwestern University Sociology Department's Culture and Urban & Community workshops, and the editors and anonymous reviewers at *Social Problems* for providing helpful comments on earlier drafts. Please direct correspondence to the author at the Department of Sociology, Northwestern University, 1812 Chicago Ave, Evanston IL 60208-001, USA or to the American Bar Foundation, 750 N. Lake Shore Drive, 4th Floor, Chicago IL 60611, USA; email: sinoesthappan2026@u.northwestern.edu; ORCID: 0000-0003-3638-6323.

Received 18 September 2023; revised 22 July 2024; accepted 24 September 2024

(Winter and Clair 2023) cycle through pretrial hearings (Sawyer and Wagner 2022). Pretrial hearings occur within the initial days after police officers arrest and detain defendants. They last from two to five minutes (TCF 2022), and judges assess each defendant's likelihood of flight and rearrest, deciding among three options to ensure reappearance at their next trial: (1) keep in detention, (2) release with restrictive conditions (i.e., cash bond, curfews, electronic monitoring, etc.), or (3) release without restrictive conditions (i.e., recognizance bonds).

Judges reproduce racial and socioeconomic inequalities in pretrial release outcomes by meting out sanctions in hearings where biases and stereotypes are routinized (Sudnow 1965; Ulmer 2019; Van Cleve 2016; Winter and Clair 2023). This is especially true in the kinds of large, urban courts featured in this study, which often manage intractably large caseloads (Kirk et al. 2022). To reform this pernicious social problem, most courts today use risk assessment tools that help judges assess and classify pretrial defendants' flight and rearrest risks by statistically comparing their case characteristics to large numbers of prior records (Christin, Rosenblat, and boyd 2015). State officials contract software developers to design risk scores that offer judges guidance in a highly consequential decision-making setting where they otherwise lack information and time (Zottola, Clarke, and Desmarais 2022). By leveraging mechanically objective procedures, risk scores ostensibly strengthen the legitimacy of judicial decision-making practices in pretrial hearings (Porter 1995).

Yet some scholars argue that scores or recommendations from risk assessments erode the legitimacy of judges' pretrial release decision-making practices by reifying existing inequalities (Angwin et al. 2016; Eubanks 2018; Harcourt 2006). Qualitative inquiries find consistent evidence of organizational decoupling, wherein court actors express mixed opinions about risk scores, and some use them in pretrial hearings more than others (Brayne and Christin 2021; Christin 2017; DeMichele et al. 2019, 2021; Terranova, Ward, Azari, and Slepicka 2020; Terranova, Ward, Slepicka, and Azari 2020). This small but growing body of evidence raises empirically underexplored questions about how judges perceive the legitimacy of risk scores, and, in turn, how risk scores affect the perceived legitimacy of judges.

Using qualitative data from four large U.S. criminal courts, this article asks: How do institutional legitimacy concerns affect how judges use risk scores to render pretrial release decisions? As organizations embedded in multiple, often competing, institutional logics (Battilana and Dorado 2010; Greenwood et al. 2011; McPherson and Sauder 2013), criminal courts offer a rich setting to study how highly consequential decision-makers evaluate and use algorithmic tools to manage their legitimacy. I show how judges consider and use risk scores in light of the bureaucratic, legal, and political institutional logics that embroil their decision-making practices.

PRETRIAL RISK ASSESSMENTS IN CRIMINAL COURTS

Pretrial detention accounts for a substantial share of the United States' contemporary mass incarceration system. Roughly three-quarters of detained individuals are held until trial because they cannot afford bail, violate release conditions, or because court actors consider them too "risky" to release for other reasons (Sawyer and Wagner 2022). Along with penalties to housing, employment, and health, pretrial detention causes future penal system involvement (Smith and Hu 2021). These punitive outcomes disproportionately impact racially and socioeconomically minoritized defendants (Demuth and Steffensmeier 2004; Schlesinger 2005). In large urban areas across the United States, Black felony defendants are more than 25 percent more likely to be held in pretrial detention than white defendants (Sawyer 2019), and the national median bail amount is two-thirds the median annual income of most pretrial defendants (Rabuy and Kopf 2016).

Scholars attribute these disparate outcomes to extralegal courtroom policies and practices that reflect structural inequalities deeply entrenched in the carceral system (Clair 2021; Ulmer 2019; Van Cleve 2016; Winter and Clair 2023). Yet, attending to the more narrowly observable problem of judicial bias, policymakers often tout pretrial risk assessments as effective tools to encourage uniformity in pretrial release decision-making practices (Kleinberg et al. 2018; Zottola et al. 2022). Advocates argue that, with sufficient training and resources, judicial adherence to risk scores can help inhibit the increasing uses of monetary bail and jail detention sanctions throughout the United States (Desmarais et al. 2021; Kleinberg et al. 2018).

While only twelve U.S. jurisdictions had implemented pretrial risk assessments at the turn of the 21st century (Scott-Hayward and Fradella 2019), they are now used statewide in eleven states and in 228 additional counties in other states (MPI 2022). In felony cases, which are the current study's focal point, court-appointed pretrial officers gather risk assessment data by conducting background checks and interviews with defendants. The kinds of data used in risk assessment tools vary, but they typically include defendants' criminal histories as well as their sociodemographic, family, and community profiles. Pretrial officers record these data in the risk assessment software, which uses statistical techniques to compare defendants' cases with digitized records of hundreds of thousands of previous cases to quantify their flight and rearrest risks. These risks are represented as separate or combined scores. Some pretrial risk assessment tools represent risk scores categorically (i.e., high, medium, and low risk), others ordinally (i.e., 1–10, 1–5, 0–3), and still others probabilistically.

The latest pretrial risk assessments use unsupervised machine learning methods that are computationally trained to calculate risk scores by automatically optimizing statistical associations between case inputs and risk outputs (Christin et al. 2015). But the most prevalent tools, including the ones used in the current study sites, employ analyst-driven, regression approaches in which developers preselect inputs (MPI 2022). While each approach has technical strengths and shortcomings, they both construe cases using comparative ways of knowing that often eclipse narrative and noncomparative context (Kiviat 2023).

Advocates, policymakers, and scholars debate whether and how using risk scores improves the perceived legitimacy of court outcomes. Some problematize the use of arrest data in risk assessments, which they argue contaminate risk scores with measurement biases from historically unequal police surveillance practices (Brayne 2020; Eubanks 2018; Harcourt 2006). Others raise ethical questions about the predictive uses of protected identity categories such as age and sex, since they may be statistically associated with flight and rearrest risk outcomes but are otherwise unconstitutional release factors (Garrett and Monahan 2020; Starr 2015). Moreover, local political controversies and discourses affect whether state policymakers adopt or roll back policies requiring the use of risk assessments in criminal courts (König and Wenzelburger 2021; Wenzelburger and König 2022).

While the larger corpus of quantitative evaluations of the predictive validity and reliability of pretrial risk assessments is largely mixed (Desmarais et al. 2021; Dobbie, Goldin, and Yang 2018; Lowder et al. 2020; Viljoen et al. 2019), some studies suggest that judges' decisions to engage with risk assessments may be motivated by extralegal organizational factors. For instance, Stevenson's (2018) evaluation found equally sharp increases and declines in Kentucky judges' uses of risk assessments as initial bureaucratic support for the tool waned over time. Angelova, Dobbie, and Yang (2023) reported that judges were more likely to issue punitive sanctions that deviated from risk scores in the period after highly salient events such as when an unrelated locally released defendant was arrested for a violent crime. Building on these insights, I investigate how institutional legitimacy concerns affect how judges use risk scores to make pretrial release decisions.

TENSIONS IN THE LEGITIMACY OF JUDICIAL DECISION-MAKING

Organizational theory points to logics as significant constraints to and enablers of legitimacy in decision-making (Suddaby and Greenwood 2005). Institutional logics refer to taken-for-granted assumptions that provide actors with context, meaning, and scripts to interpret and respond to specific situations (Lounsbury et al. 2021; Thornton and Ocasio 1999). Logics reflect an organization's guiding principles, and "complex" or "hybrid" organizations such as criminal courts contain multiple competing logics that can undermine the legitimacy of actors' decisions (Battliana and Dorado 2010; Greenwood et al. 2011; McPherson and Sauder 2013; Pache and Santos 2010; Reay and Hinings 2009). Zozula's (2019) ethnography revealed how court actors justified the use of increased supervision and punitive sanctions under the guise of rehabilitation to circumnavigate tensions between punishment and therapeutic logics.

Institutional perspectives are compatible with how criminologists view legitimacy as stemming from a procedural justice logic, which compels law enforcement officials to listen to criminal defendants and to convey trustworthiness, respect, and neutrality (Nagin and Telep 2017; Tyler 1990). From this view, judicial legitimacy comes from the public, whose assessments and evaluations take the form

of a *political institutional logic*. But related frameworks, such as organizational justice (Taxman and Gordon 2009) and self-legitimacy (Tankebe 2019), which consider how penal actors assess their own workplace authority and rules, also attribute legitimacy to *bureaucratic* and *legal institutional logics*. This scholarship, largely developed in studies of policing, suggests a multidimensional view of legitimacy. The current article integrates these criminological and institutional approaches to examine how criminal court judges use risk scores to manage their legitimacy against competing bureaucratic, legal, and political logics.

A growing number of studies examine the contexts in which penal actors interpret and use risk assessments. Some research suggests that actors across the criminal legal system view risk assessments as threats to their professional expertise and legitimacy and, therefore, resist them in different ways (Brayne 2020; Cheliotis 2006; Lynch 2019; McNeill et al. 2009; Werth 2017, 2019). In one Southern County courtroom, judges characterized risk scores as discretionary oversight mechanisms that professionally deskilled them, and thus they effectively ignored them (Brayne and Christin 2021; Christin 2017). Court officers also sometimes "criteria tinker" with or "game" risk assessments to generate outputs they consider more consistent with their own evaluations (Brayne and Christin 2021; Hannah-Moffat, Maurutto, and Turnbull 2009).

Other studies indicate that some judges embrace risk assessments, albeit tepidly. Judges interviewed in Hartmann and Wenzelburger's (2021) study valued using risk scores in pretrial hearings in large part because "the context of the situation changes from one of fundamental uncertainty to one of statistical probability" (284). DeMichele et al.'s (2019) national survey reported that almost two thirds of judges (63 percent) always or often agreed with pretrial risk assessment scores, and over three quarters (79 percent) always or often made decisions consistent with their recommendations. In subsequent interviews, judges clarified that while these scores were valuable, they still harbored concerns about their validity and accuracy, which corresponded with how they used risk scores on a case-by-case basis (DeMichele et al. 2021).

Whether judges resist or embrace risk scores, it stands to reason that they engage with them in strategic ways. This is especially likely in pretrial hearings, where judges regularly face tensions within or among competing institutional logics that threaten their legitimacy (Winter and Clair 2023). This explanation is consistent with descriptions of judicial decision-making as context-dependent and situational (Clair 2021; Feeley 1979; Kohler-Hausmann 2018; Van Cleve 2016; Zozula 2019). For instance, although judges and attorneys in McPherson and Sauder's (2013) courtroom ethnography were guided by their own "home" logics, they also "hijacked" each other's logics when they otherwise lacked evidence to bolster the legitimacy of their claims. Clair and Winter (2016) interviewed judges in a northeastern state who uniformly disapproved of racial disparities, but selectively contested perceived racism based on the situational and institutional contexts of each stage of trial.

These studies suggest that judges may strategically use risk scores to legitimate their pretrial release decisions. But how they conceptualize and leverage risk scores across different cases, and the ways these practices affect processes of judicial decision-making in pretrial hearings, are not well explained. I address this gap by examining how judges think about and use risk scores in the context of the bureaucratic, legal, and political logics that govern their pretrial release decision-making practices.

METHODS

Pretrial hearings are highly structured proceedings. Although speaking turns vary across courtrooms, each party follows a scripted temporal order. Pretrial officers present risk scores, providing technical context as needed. Prosecutors and public defenders make "proffers," that is, they summarize relevant evidence and recommend outcomes on behalf of the state and their clients, respectively. Finally, judges openly discuss the factors they consider in each case before ruling on sanctions. The mechanical nature of these proceedings accommodates the high volumes of cases that many courts must process. But they also render pretrial hearings at an extraordinarily fast pace Court actors reported that they lasted anywhere from 30 seconds to fifteen minutes with a modal reported length of five minutes.

Given these restrictions, observations lend a somewhat limited view of judicial decision-making practices at pretrial. Pairing fieldwork with interviews helps overcome these limitations, capturing how judges account for both their symbolic beliefs and material courtroom practices (Clair 2021;

Van Cleve 2016). In this article, I draw primarily on interviews with 27 judges across four criminal courts to examine how the reception and uses of risk scores affect judicial decision-making practices in pretrial hearings. This qualitative approach is well suited to reveal how judges "position themselves with respect to algorithms" and "enroll them in their institutionalized ways of doing things" (Christin 2020:906).

From July 2022 to July 2023, I conducted virtual, open-ended, semi-structured interviews with judges in four large criminal courts in urban Northwest, Midwest, South, and West U.S. counties. I selected the study sites partly based on convenience and to maximize variation on geography, politics, and risk assessment tools. I obtained approval from agency administrators in each site, who sent flyers to their listservs, facilitated access to directories, and/or made direct referrals to staff. I also contacted some interviewees using directories I obtained from requests to county FOIA officers. I recruited and enrolled judges using e-mail and phone, and I administered the interviews on Zoom for an average of 52 minutes, ranging from 28 to 94 minutes. I audio-recorded and transcribed interviews with all but two judges who allowed me to take notes.

The 27 judges in my sample represent 44 percent of all 61 judges across the study sites who were assigned to preside over pretrial hearings when I collected data. Although I was unable to contact all 61 judges due to the sampling method, the final proportion I sampled is still high considering elite interviewees, especially judges, are normally reluctant to participate in research (Nir 2018). The average age of judges was 54, and they were diverse in terms of gender and race and ethnicity (Table 1).

My interview questions covered the pretrial process, what judges thought about risk assessments, how they used them, and how they believed other court actors used them. I asked questions such as "when do you find risk scores more or less useful?" and "how and with whom do you discuss risk scores in pretrial hearings?" to gain insight into how risk scores figured in judges' decision-making practices. I modified the interview questionnaire based on insights I gained from preceding interviews. For instance, midway through the study, I began asking judges from Northwest and Midwest counties their opinions about upcoming bail reforms when I learned that they were highly politicized in those jurisdictions.

I supplemented my primary dataset with three additional sources. First, I interviewed 11 pretrial officers, 18 public defenders, and 17 prosecutors at these four sites using questionnaires I adapted for each court role. These interviews offered reference points that helped explain variation in judges' risk assessment-based pretrial release practices, and I used them to contextualize results from my analyses of judicial interviews. Second, I intermittently conducted 50 hours of in-person and virtual observations in the Midwest County bond court throughout the study. This data source enriched

Table 1. Sociodemographic Characteristics of Judges

	Count (%)	Mean (SD)
Location		
Midwest	5 (18.5)	
Northwest	7 (25.9)	
South	8 (29.6)	
West	7 (25.9)	
Age		53.9 (7.9)
Gender		
Woman	12 (44.4)	
Man	15 (55.6)	
Race/Ethnicity		
Black	4 (14.8)	
Latinx	3 (11.1)	
Asian	2 (7.4)	
White	18 (66.7)	
Total	27 (100)	

my understanding of court norms and terms and how judges dealt with institutional tensions in situ. Third, I analyzed a wide range of documents (i.e., local news coverage, research reports, and redacted case files) to better understand the political climates and pretrial risk assessment tools used in each site.

Triangulating these diverse sources helped me better parse through differences in site-specific pretrial practices and risk assessment tools. I incorporated the insights I gleaned from this process into both my interview protocols and the subsequent analyses. Throughout the analysis, I directly reference field observations that correspond with and effectively illuminate findings from the interviews. However, I do not explicitly reference data from non-judicial interviews or other documents in the findings.

After initially reviewing the full dataset, I determined that institutional processes were sufficiently consistent across the sites to combine the interviews by court role for analysis. However, a few differences are worth noting. Both Northwest and Midwest counties had recently passed laws that would eliminate the use of bail, or monetary sanctions, once in place. Bail reforms did not seem to be a priority in South and West counties, where there were no reported plans to implement them. Additionally, each site used risk assessments that were developed by different agencies and included different input factors (Table 2). Notably, Midwest County's risk assessment included only a battery of criminal history input measures, but pretrial officers nonetheless added other discretionary data from intake interviews to defendants' case files.

I used Atlas.ti 23 to abductively build and test emerging concepts; that is, I iteratively compared my data to prior theories to explain unexpected empirical results (Timmermans and Tavory 2012). Following Deterding and Waters (2021), I indexed the dataset, and then applied analytic codes to those indexed excerpts. I derived these analytic codes from my initial theoretical assumptions that judges would either exclusively welcome or dismiss risk scores. Over multiple rounds of coding, I refined the analytic codebook based on observations that these behaviors were patterned by distinct institutional logics. This iterative process surfaced tensions within which judges situated their perceptions and uses of risk scores. In the final round of coding, I used three parent codes ("bureaucracy," "the law," and "politics") linked to three subcodes ("tensions," "risk score views," and "risk score practices").

FINDINGS

In this section, I show how judges strategically used risk scores to legitimate their decisions when faced with tensions stemming from three institutional logics with different legitimating referents. Bureaucratic logics were upheld by local and state rules that governed courtroom procedures, and they demanded that judges efficiently process high volumes of cases with limited information and time. Legal logics were upheld by statutory rules, and they demanded that judges protect public safety while imposing the least restrictive conditions. Finally, political logics were upheld by the media and

Table 2. Risk Assessment Inputs across the Study Sites

	Midwest	West	Northwest	South
Age[1, 2]	✓	✓	✓	✓
Current and pending charges[1]	✓	✓	✓	✓
Prior arrests, convictions, incarceration, and FTAs[1, 3]	✓	✓	✓	✓
Alcohol and/or drug use[1, 2]		✓	✓	✓
Education and/or employment status[2]		✓	✓	✓
Home and family life factors[2]			✓	✓
Attitudinal factors[2]				✓

Notes: [1]data from criminal records databases; [2]data from interview; [3]failures to appear

members of the public who assessed judicial performance in municipal elections, and they demanded that judges follow the law while managing public scrutiny.

In each subsection that follows, I illustrate how tensions within and among the bureaucratic, legal, and political logics constrain judges' pretrial release decisions. I also show how judges mobilize risk scores to manage these tensions and maintain their legitimacy. I conclude by arguing that the institutional logics and tensions in which their decisions are embroiled can reveal how official actors strategically engage with algorithmic risk scores to make highly consequential decisions.

Bureaucratic Tensions and the Uses of Risk Assessments

Judges face growing uncertainty about their decisions as cases are piled onto their desks each day. During pretrial hearings, presiding judges who oversee the courthouse expect the judiciary to efficiently process high volumes of caseloads. Judges reported struggling to meet this bureaucratic demand using the limited time and minimal evidence at their disposal.

Until pretrial hearings began, judges had access only to narrative arrest reports written by law enforcement officials. These reports alone did not offer sufficient grounds for their decisions, as judges also relied on statements from pretrial officers and proffers from attorneys to learn the necessary context of each case. But in Northwest, West, and South Counties, judges pointed to a "public defender crisis" that frequently left defendants unrepresented and judges without the facts they needed. Judge Clark, for instance, lamented that "I have a lot of self-represented litigants, and I get concerned that they're not getting a fair result because they're just not well equipped to navigate the courtroom or our procedures [...] I can't be their attorney or give them legal advice because that's not fair."

Judges also reported experiencing decision fatigue from the limited time they had to evaluate cases. Judge Smith characterized these hearings as "fast and furious," explaining that "I get handed the reports as the case is being called. So, I would say, we learn to process a lot of information quickly." Judge Johnson took note of these bureaucratic tensions when he was appointed to preside over pretrial hearings in Midwest County: "I think the real focus was, 'there are a lot of cases on the calendar today, and we've got to get through this.' And when we came in, we said, 'let's be more efficient,' but efficiency shouldn't just be focused on speed, but rather if there's a process that we can look at." These observations led Judge Johnson to champion the use of risk assessments in his court to improve the efficiency of pretrial hearings. He believed that risk assessments were essential in such a resource-constrained environment:

> You're making detention decisions that have an impact way down the line. So, then every tool that you have to assist you is important. The [risk assessment] then is a tool that assists us. Not like that's the only tool. The [prosecutor] is presenting their side, the defense is presenting their side, and now we've got this neutral tool. It's not saying, "Judge, you've got to hold this guy, he's a bad guy." Or, "Judge, you don't understand, my guy is a really good guy and goes to church all the time." This is right down the middle. Here's how we assessed this individual's risk. Take that into account, as you will.

Risk scores gave Judge Johnson a much-needed efficient process for obtaining objective guidance he otherwise would not have. Other judges were warier about the objectivity of this guidance. But like Judge Johnson, they found utility in risk scores, especially when they lacked time and information to evaluate cases. Judge Williams was skeptical of the risk assessment in her county, which calculated scores based on interview data: "Obviously, that information is not perfect because some if it is from the defendant themselves, so it may or may not be accurate. And we also may or may not have the full criminal history available at the time. So, in the end, it doesn't fill in all the gaps." Despite these criticisms of processual bias and error, Judge Williams selectively relied on risk scores to quell her uncertainties in pretrial hearings:

> It just gives you a baseline shot of where this person might fall on... I think it's most useful when the score is either really high or really low. But if somebody's kind of in the middle, I don't know if that's super helpful, just because they're falling in the middle of where everybody would expect to be.

Pretrial risk assessments consider the entire population inherently "at risk" of flight and rearrest based on multiple input measures (Werth 2019). Recognizing that fact, Judge Williams exclusively paid attention to high and low scores, which she felt better informed her release decisions. Perhaps because of this bureaucratic utility, many judges were adamant that risk scores were not only helpful but also necessary, even (and often) when they were skeptical about a specific score's legitimacy.

Judge Brown emphasized the need to contextualize risk scores when interpreting them: "When I look at the risk assessment, and it says it's a [high score] that they should probably not be released, then I'm going to be looking for why. What is the situation that's giving this person a high score? What I'm looking for is to understand how the assessment came up with that number, to see if I agree with its recommendation." But Judge Brown also reported that risk scores were most useful precisely when he had little time to scrutinize other information in case files:

> So, if you're in a hurry and you haven't done any pre-work, you just ran into the court, and you open up the person sitting right in front of you. You open up their file and you jump to this section where there is a number. And you go, "oh, this person's a [high score]. I don't have time to go through all the file. But I do know that the [prosecutor] is likely not going to ask that the person be held. Right, so that's one less thing that I have to worry about."

Like Judge Brown, most judges were cognizant of errors in the production of risk scores, but they reconciled these concerns by situating them against the backdrop of a bureaucratically inefficient pretrial process. When presented with bureaucratic tensions that threatened to undermine their legitimacy, judges selectively used risk scores to make pretrial release decisions. Even when they questioned the accuracy and validity of risk scores, judges selectively relied on them to help justify their decisions in cases when they were hampered by information and time restrictions.

Legal Tensions and the Uses of Risk Assessments

Judges unanimously cited two statutory factors that governed how they assessed flight and rearrest risks. On one hand, judges were to assign the "least restrictive conditions" necessary to ensure that defendants safely returned to court (American Bar Association 2007). On the other hand, they had to protect public safety by using detention or release measures that all restrict the mobility of defendants, save for recognizance bonds. These demands posed a vexing challenge to the legal bases of judges' decisions, which they referred to as a "complex calculation" and "a careful balancing act."

Judges often cited defendants' social service needs when considering their legal obligation to impose least restrictive conditions, but they also worried about how their decisions would affect public safety. Judge Jones explained that "We have to figure out, 'Is it punishment, or is it needs?' Because you can't punish someone into their needs, but we're not addressing anyone's needs by just letting them out on [recognizance bonds]." Like Judge Jones, many judges attributed the bulk of cases they heard in pretrial hearings to the criminalization of poverty, that is, they believed these defendants' crimes would be best addressed through social services rather than punishment. Judge Davis felt that her county did not have sufficient resources to impose least restrictive conditions as often as she wanted to; instead, she reported more frequently using punitive sanctions intended to protect public safety:

> In the long run, we're hopefully setting this person up so that they're not going to be committing property crimes or domestic violence. The problem, of course, is resources. We don't have a mental health clinic, or a place for homeless people to go. So sometimes it's frustrating, but as a judge, you have to step back and say, "Well, my role is not to help in that way." My role is to call balls and strikes. I'm not a social engineer. If I start trying to social work it, then I veer into advocate.

Judges' difficulties with balancing the law's demands were magnified by their challenges with discerning flight and rearrest risks from the limited evidence at their disposal. Judges believed that risk scores measured some legally pertinent case input factors better than others. Notably, they expressed concerns about risk scores overestimating or underestimating the "true" risks in each case. Indeed, judges most often reported exercising their discretion in cases when they deemed risk scores under- and over-representative of legally defined risks.

Judges believed that risk scores sometimes overestimated the legal risks of cases by inflating the importance of case factors they considered trivial. Many also questioned high risk scores when cases included outdated criminal records, which they viewed as potential proxies for race and socioeconomic status. Judge Williams elaborated, "We've had concerns about, 'Are we overvaluing things like unemployment or hasn't graduated? Because certain types of populations are inherently going to have a higher percentage of people that just don't have those things. But does that necessarily mean they're a danger? No."

Judge Garcia routinely dismissed risk scores that she believed overestimated risks:

> The person comes before the court charged with the lawful possession of a controlled substance. It's like 0.1 gram of crack. The defendant is 62 years old, and the risk score is [high]. Pretrial's recommendation is maximum conditions or electronic monitoring if release. And I'm like "no, because that's based on a long criminal history" [....] You cannot be rote when using the [risk assessment]. You really have to weigh everything, and you have to bring your life experience in too.

Judge Garcia also dismissed risk scores that she believed underestimated risks. She cited several pertinent legal factors such as juvenile records and the presence of victims and weapons, which risk scores omit as inputs. Judge Garcia was especially concerned with how risk scores overlooked cases involving loaded firearms known as "extended magazines" with higher-than-normal ammunition capacities. She called them "straight up killing machines." Several months after our interview, I observed Judge Garcia's courtroom, where she ruled on six pretrial cases within thirty minutes using a standard script that I heard many other judges repeat. By the seventh case I observed, in which a defendant faced felony battery charges for using an extended magazine, Judge Garcia broke frame. During the public defender's argument to release the defendant based on his low-risk score, Judge Garcia interjected to issue detention. In justifying her decision, Judge Garcia reminded the court about the hazardous weapon the defendant carried, adding that the nature of the charge outweighed his low-risk score.

Whether it was because they believed the scores overestimated or underestimated risks, many judges described the need to exercise their discretion with using risk scores. They reported vaguely "pairing" risk scores, or "throwing them in the mix," with their own legal professional evaluations of each case, which tended to rely more heavily on prosecutorial charges at the pretrial stage (Frederick and Stemen 2012). To this point, most judges deemed it acceptable to exercise their discretion when risk scores did not align with their perceptions of serious charges. Several judges evoked driving metaphors like "speed bumps," "bumpers," and "yellow tape" to stress the need for discretion around certain higher-level charges they viewed as exceptions. Judge Miller explained using charges as a legal benchmark to decide when to rely on risk scores:

> If it's somebody that you're going to give a [recognizance bond] to, listen to the [risk assessment]. There's a vast amount of data involved in that recommendation, and that's going to be more accurate than your gut, and so definitely follow that. Other than that, it's like striped, yellow lines on an express lane. You should stay within it, but sometimes you need to go outside the line because what you're seeing doesn't match. It's like driving a Tesla. Keep your hands on the wheel, even though auto pilot's doing its job.

Many judges deployed their own moral views about specific charges as yardsticks to decide when risk scores were and were not legitimate in the eyes of the law. These practices run in sharp contrast to the goals of risk assessment, which are to institutionalize data-driven decision-making practices in pretrial hearings while deinstitutionalizing practices based on judges' moralistic evaluations of specific charges. By selectively invoking risk scores, judges nevertheless exercised their discretion to define what did and did not constitute risks, even in cases when they were presented with evidence to the contrary. Judges' selective uses of risk scores then helped maintain their legitimacy in the face of legal institutional tensions.

Political Tensions and the Uses of Risk Assessments

While judges described themselves as arbiters of the law, they also recognized their roles as public servants. These twin commitments reflect a political tension that mired judges' pretrial release decisions.

Judges' public service roles were most concretely evident in South and West counties, where they stood for reelection. In organizational terms, judges' legitimacy and survival depended on favorable public assessments of their performance. Many judges conveyed that heightened public scrutiny exacerbated their concerns about false-negative outcomes, that is, when they released high-risk defendants. But there appeared to be no similar public accountability process to address false-positives, that is, when judges sanctioned restrictive conditions such as detention or conditional release on low-risk defendants. These concerns reportedly made some judges think twice about public responses to their decisions.

Judge Anderson made a point to emphasize that she "makes [her] own decisions," but she also described the difficulties of escaping public opinion:

> So, you make a decision that somebody doesn't like, fair enough. But now they can show up on your front lawn, yelling at you. They can put on social media that you're a racist. And then judicial ethics says we can't respond to that [...] and I've noticed very much an uptick in the kind of disrespectful disregard for court authority from litigants, but lawyers too. Using this claim of you're a racist as a litigation tactic to, you know, get rid of you as the judge.

Judge Anderson's concerns reveal how decisions that judges consider legally legitimate may be in tension with decisions that the public considers politically legitimate. For similar reasons, Judge Thomas took extra care to prepare for pretrial hearings: "There's a little level of paranoia that if you just go out there blind, you don't know that the [prosecutor] is going to give you everything about the case. And I don't care what the public says about my decisions, but I do care if they think I'm not paying attention."

Several judges pointed to the news media as a key source of fear about public scrutiny. Judge Jones explained that "Your worst nightmare is you let someone out on a lower bond and then they go and hurt someone. I mean, all of us, when I see those stories on the news, I think that could have been any of us." Judges in South County referred to the media's political influence on judicial decision-making as "the [local newspaper] rule," which Judge Ryan defined as "not wanting to end up in the newspaper tomorrow for the decision you made today."

Within this context of heightened public scrutiny, risk scores represented a valuable resource that judges could use to legitimate their decisions. By aligning their decisions with risk scores, judges could effectively shift accountability for case outcomes to the tool, shielding against public criticism. Take, for instance, Judge Hall, who characterized risk scores as an essential guard rail for his decisions: "Evidence-based decision-making is an important goal. You want to make sure that when you're setting bonds, you're making your decisions based on facts, not emotions, not some other inappropriate considerations." But more than valuing risk scores for their purported objectivity, Judge Hall reported using them to justify politically difficult choices:

> If I'm being honest, sometimes we'll use it as a justification to do something that a victim doesn't want us to do. Say you've got a victim who wants someone locked up. Sometimes, what you'll do as the judge is say, "We're guided by a risk assessment that scores for success in the defendant's likelihood to appear and rearrest." And, based on the statute and this score, my job is to set a bond that protects others in the community.

Judge Hall's remark brings into sharp focus the tendency for judges to publicly invoke risk scores as a reference point for their decisions. Most pretrial hearings I observed began with a ceremonial nod to risk scores. In their opening statements, judges recanted: "The decisions I make today are guided by a statistically validated risk assessment." On one hand, when rendering pretrial release decisions, judges affirmatively invoked risk scores when they appeared politically advantageous. On the other hand, when risk scores were politically unfavorable, they habitually repeated the phrase "despite the [risk score]," followed by perfunctory statements indicating moral assessments of each case.

Judges' pretrial release decisions are rife with political tensions. Judges make politically risky decisions that can cost them public trust and, in some cases, reelection. Notably, judges face political risks for making decisions that are at odds with risk scores, should those defendants who are released be

rearrested or fail to reappear in court. Judges do not, however, pay a political price for making harsh decisions that are at odds with risk scores. Judges navigated this tension by selectively invoking risk scores when they were politically advantageous and ignoring them when they were not.

DISCUSSION AND CONCLUSION

Pretrial risk assessments are among a growing wave of new technologies that are now increasingly central to highly consequential adjudication responsibilities traditionally governed solely by official actors. Past studies provide limited accounts of how these actors deal with technological change. Using data from a qualitative study of data-driven pretrial release decision-making practices in four large U.S. criminal courts, I reveal how judges conceptualized and used risk scores with concerns about their legitimacy in mind.

Judges faced competing demands from bureaucratic logics to efficiently dispose of cases, legal logics to protect public safety and impose least restrictive conditions, and political logics to manage their commitments to pretrial statutes and the public. Although they harbored mixed opinions about the tools, judges selectively invoked risk scores when they effectively managed tensions within and among these logics and ignored them when they did not. These strategic uses of risk scores effectively insulated judges from institutional critique and legitimated their punitive sanctions. These findings can help explain more general mechanisms of penal change and organizational decision-making writ large; I address each mechanism in turn.

Implications for Penal Change

Policymakers feature risk assessments as a solution to criminal court outcomes that perpetuate racial and socioeconomic inequalities, suggesting that they rationalize pretrial release practices by augmenting judges' decision-making repertoires. Some scholars wage criticisms against these claims, asserting that risk scores instead reflect and thereby reproduce biases in the underlying data they measure (Brayne 2020; Eubanks 2018; Harcourt 2006). My findings showcase an additional source of post hoc bias in risk scores introduced by judges who invoke them when they mitigate institutional tensions and dismiss them as invalid when they do not. In contexts such as pretrial hearings, where the use of risk assessments is advisory, risk scores may constrain decision-making practices to some extent, but judges still preserve their discretion by selectively invoking risk scores to circumnavigate institutional tensions.

These findings help situate risk assessments as technologies that amplify the U.S. pretrial system's enduring social control function. Page and Scott-Hayward (2022) argue that the system incentivizes risk aversion and punitiveness because defendants are assumed risky, not only to flight and rearrest, but also to judges' political careers. These incentives are not new: judges' pretrial release decisions have been under public scrutiny since at least the 1960s when politicians and advocates alleged that judicial leniency contributed to rising crime rates (Zimring, Hawkins, and Kamin 2001). When selectively invoked, risk scores serve as legitimating resources that judges can harness to stymie political blowback for decisions that result in unfavorable outcomes (i.e., when defendants commit new crimes or fail to reappear in court). Whereas policymakers tout their ability to curb judicial discretion, in practice, risk scores expand the uses of discretion among judges who strategically use them to justify punitive sanctions.

In recent years, computational scholars have proposed redeveloping risk assessment tools by minimizing biases built into the algorithms' constitutive data and code (Berk and Elzarka 2020). While these reforms show some promise for addressing the tool's underlying biases, they overlook how judges can appropriate these tools (Brayne and Christin 2021; Christin 2017; Lynch 2019). My findings suggest that, even if technical changes to risk assessments were to reduce biases related to measured inputs or statistical modeling approaches, they would not address biases introduced by the kinds of discretionary judicial practices revealed in this study.

Instead, by lending legitimacy to judges who may otherwise continue to exercise biased discretion, the use of risk assessments can further obscure decision-making practices that reinforce racial and socioeconomic inequalities in pretrial hearings. This finding aligns with recent scholarship that suggests risk assessments increase the political costs of lenient pretrial release decisions, nudging judges

toward punitive ones instead (Angelova et al. 2023). It also suggests that, by further analyzing the relationship between pretrial risk assessments and the legitimacy of court actors who use them, future studies can more fully appreciate the effects these technologies have on penal institutional change.

Implications for the Institutionalization of Risk Assessments and the Use of Discretion

Risk assessments are culturally authoritative tools that instill legitimacy in organizational decision-making processes (Espeland and Stevens 1998; Porter 1995). They are valuable in publicly scrutinized institutions like criminal courts (Lynch 2019), whose legitimacy is undermined by inefficiencies and inequities in institutional outcomes (Zottola et al. 2022). Today's risk assessments are considered especially credible because they use algorithms that process massive amounts of data at tremendous speed (Burrell and Fourcade 2021). Yet, they are also impaired by measurement errors that call their legitimacy into question.

This study asks how judges navigate these legitimacy concerns in the context of using risk scores. I show how judges strategically reference risk scores in pretrial hearings to manage tensions within and among institutional logics that govern their decisions. These findings defy a binary view of algorithmic tools as either constraints to or enablers of organizational decision-making. Instead, I show how risk scores may be better conceptualized as resources that actors situationally use to navigate institutional tensions and legitimate their decisions. This is especially likely in "unsettled" contexts such as pretrial hearings where judges have limited access to decision-making resources (Swidler 1986). A common refrain among judges in this study was in fact that risk scores were "simply another tool in their toolkit."

While the long-term implications for how algorithmic risk scores affect the work of official actors are yet to be seen, the current analysis helps elucidate how the uses of data-driven discretion do and do not change what organizations fundamentally do and how they do it. Pretrial risk assessments do not seem to exclusively move judges from a non-comparative evaluation approach to a comparative one (Kiviat 2023). These officials rather selectively invoke comparative and non-comparative judgments depending on the situation at hand, and risk assessments offer yet another tool to legitimately exercise discretion in the process.

The uses of risk scores also notably varied across criminal courts. Judges in South County expressed more skepticism about the bureaucratic legitimacy of risk assessments than in other counties where judges had stronger relationships with pretrial officers. In other words, judges had greater doubts about risk scores in a site where they were largely unfamiliar with the pretrial officers who prepared them. Judges also appeared to express greater concerns about the political legitimacy of risk scores in Northwest and Midwest counties, where pretrial practices were under heightened public surveillance due to the politicization of local bail reforms. Taken together, these findings highlight the salience of local organizational contexts, and thereby further qualitative inquiry, for understanding the uses of algorithmic risk scores.

While this study sheds light on judges' self-reported accounts of using risk scores, it does not systematically capture how these practices affect pretrial release outcomes. Textual analyses of courtroom transcripts are well suited to more precisely measure how risk scores may be causally linked to judicial decisions and court outcomes. Moreover, although this study primarily features judges' accounts of using pretrial risk assessments, criminal court decisions are often the outcome of interactions between judges and other court actors (McPherson and Sauder 2013; Ulmer 2019). Future work incorporating perspectives from attorneys and pretrial officers can uncover how differences in organizational status and role, and interactions among these groups, shape processes of data-driven decision-making.

Despite these limitations, this qualitative study of pretrial release decision-making practices presents a novel account of how judges selectively invoked risk scores to mitigate tensions within or among institutional logics that legitimated their decisions. Studying risk assessments as legitimating resources in other highly consequential decision-making domains can help reveal a wider range of strategies actors use to mobilize these tools and navigate institutional tensions. Moreover, as risk assessment technologies proliferate in institutions around the world, the use of global and comparative approaches to studying them can uncover important insights into how they structure the uses of professional discretion and life outcomes.

REFERENCES

American Bar Association. 2007. *ABA Standards for Criminal Justice*, 3rd ed. Chicago: American Bar Association.

Angelova, Victoria, Will S. Dobbie, and Crystal Yang. 2023. *Algorithmic Recommendations and Human Discretion*. National Bureau of Economic Research. Working Paper 31747. https://doi.org/10.3386/w31747.

Angwin, Julia, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2016. "Machine Bias." *ProPublica*, May 23. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Battilana, Julie, and Silvia Dorado. 2010. "Building Sustainable Hybrid Organizations: The Case of Commercial Microfinance Organizations." *Academy of Management Journal* 53(6):1419–40. https://doi.org/10.5465/amj.2010.57318391.

Berk, Richard, and Ayya A. Elzarka. 2020. "Almost Politically Acceptable Criminal Justice Risk Assessment." *Criminology & Public Policy* 19(4):1231–57. https://doi.org/10.1111/1745-9133.12500.

Brayne, Sarah. 2020. *Predict and Surveil: Data, Discretion, and the Future of Policing*. New York: Oxford University Press.

Brayne, Sarah, and Angele Christin. 2021. "Technologies of Crime Prediction: The Reception of Algorithms in Policing and Criminal Courts." *Social Problems* 68(3):608–24. https://doi.org/10.1093/socpro/spaa004.

Burrell, Jenna, and Marion Fourcade. 2021. "The Society of Algorithms." *Annual Review of Sociology* 47:213–37. https://doi.org/10.1146/annurev-soc-090820-020800.

Cheliotis, Leonidas K. 2006. "How Iron is the Iron Cage of New Penology? The Role of Human Agency in the Implementation of Criminal Justice Policy." *Punishment & Society* 8(3):313–40. https://doi.org/10.1177/1462474506064700.

Christin, Angèle. 2017. "Algorithms in Practice: Comparing Web Journalism and Criminal Justice." *Big Data & Society* 4(2). https://doi.org/10.1177/2053951717718855.

Christin, Angèle. 2020. "The Ethnographer and the Algorithm: Beyond the Black Box." *Theory and Society* 49(5): 897–918. https://doi.org/10.1007/s11186-020-09411-3.

Christin, Angèle, Alex Rosenblat, and danah boyd. 2015. *Courts and Predictive Algorithms*. New York: Data & Civil Rights.

Clair, Matthew. 2021. *Privilege and Punishment: How Race and Class Matter in Criminal Court*. Princeton, NJ: Princeton University Press.

Clair, Matthew, and Alix S. Winter. 2016. "How Judges think about Racial Disparities: Situational Decision-Making in the Criminal Justice System." *Criminology* 54(2):332–59. https://doi.org/10.1111/1745-9125.12106.

DeMichele, Matthew, Peter Baumgartner, Kelle Barrick, Megan Comfort, Samuel Scaggs, and Shilpa Misra. 2019. "What Do Criminal Justice Professionals Think about Risk Assessment at Pretrial?" *Federal Probation* 83(1):32–41.

DeMichele, Matthew, Megan Comfort, Kelle Barrick, and Peter Baumgartner. 2021. "The Intuitive-Override Model: Nudging Judges toward Pretrial Risk Assessment Instruments." *Federal Probation* 85(2):22–31.

Demuth, Stephen, and Darrell Steffensmeier. 2004. "The Impact of Gender and Race-Ethnicity in the Pretrial Release Process." *Social Problems* 51(2):222–42. https://doi.org/10.1525/sp.2004.51.2.222.

Desmarais, Sarah L., Samantha A. Zottola, Sarah E. Duhart Clarke, and Evan M. Lowder. 2021. "Predictive Validity of Pretrial Risk Assessments: A Systematic Review of the Literature." *Criminal Justice and Behavior* 48(4):398–420. https://doi.org/10.1177/0093854820932959.

Deterding, Nicole M., and Mary C. Waters. 2021. "Flexible Coding of In-Depth Interviews: A Twenty-First-Century Approach." *Sociological Methods & Research* 50(2):708–39. https://doi.org/10.1177/0049124118799377.

Dobbie, Will, Jacob Goldin, and Crystal Yang. 2018. "The Effects of Pretrial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges." *American Economic Review* 108(2): 201–40. https://doi.org/10.1257/aer.20161503.

Espeland, Wendy Nelson, and Mitchell L. Stevens. 1998. "Commensuration as a Social Process." *Annual Review of Sociology* 24:313–43. https://doi.org/10.1146/annurev.soc.24.1.313.

Eubanks, Virginia. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin's Press.

Feeley, Malcolm M. 1979. *The Process Is the Punishment: Handling Cases in a Lower Criminal Court*. New York: Russell Sage Foundation.

Frederick, Bruce, and Don Stemen. 2012. *The Anatomy of Discretion: An Analysis of Prosecutorial Decision-Making*. New York: Vera Institute of Justice.

Garrett, Brandon L., and John Monahan. 2020. "Judging Risk." *California Law Review* 108:439–93. https://doi.org/10.15779/Z38B56D515.

Greenwood, Royston, Mia Raynard, Farah Kodeih, Evelyn R. Micelotta, and Michael Lounsbury. 2011. "Institutional Complexity and Organizational Responses." *Academy of Management Annals* 5(1): 317–71. https://doi.org/10.5465/19416520.2011.590299.

Hannah-Moffat, Kelly, Paul Maurutto, and Sarah Turnbull. 2009. "Negotiated Risk: Actuarial Illusions and Discretion in Probation." *Canadian Journal of Law and Society/La Revue Canadienne Droit et Société* 24(3):391–409. https://doi.org/10.1017/S0829320100010097.

Harcourt, Bernard E. 2006. *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age*. Chicago: University of Chicago Press.
Hartmann, Kathrin, and Georg Wenzelburger. 2021. "Uncertainty, Risk and the Use of Algorithms in Policy Decisions: A Case Study on Criminal Justice in the USA." *Policy Sciences* 54:269–87. https://doi.org/10.1007/s11077-020-09414-y.
Kellogg, Katherine C., Melissa A. Valentine, and Angele Christin. 2020. "Algorithms at Work: The New Contested Terrain of Control." *Academy of Management Annals* 14(1):366–410. https://doi.org/10.5465/annals.2018.0174.
Kirk, Gabriela, Kristina J. Thompson, Beth M. Huebner, Christopher Uggen, and Sarah K. S. Shannon. 2022. "Justice by Geography: The Role of Monetary Sanctions across Communities." *RSF: The Russell Sage Foundation Journal of the Social Sciences* 8(1): 200–20. https://doi.org/10.7758/RSF.2022.8.1.09.
Kiviat, Barbara. 2023. The Moral Affordances of Construing People as Cases: How Algorithms and the Data They Depend on Obscure Narrative and Comparative Justice." *Sociological Theory* 41(3):175–200. https://doi.org/10.1177/07352751231186797.
Kleinberg, Jon, Jens Ludwig, Sendhil Mullainathan, and Adesh Rambachan. 2018. "Algorithmic Fairness." *AEA Papers and Proceedings* 108:22–27. https://doi.org/10.1257/pandp.20181018.
Kohler-Hausmann, Issa. 2018. *Misdemeanorland: Criminal Courts and Social Control in an Age of Broken Windows Policing*. Princeton, NJ: Princeton University Press.
König, Pascal D., and Georg Wenzelburger. 2021. "When Politicization Stops Algorithms in Criminal Justice." *The British Journal of Criminology* 61(3):832–51. https://doi.org/10.1093/bjc/azaa099.
Lounsbury, Michael, Christopher W. J. Steele, Milo Shaoqing Wang, and Madeline Toubiana. 2021. "New Directions in the Study of Institutional Logics: From Tools to Phenomena." *Annual Review of Sociology* 47:261–80. https://doi.org/10.1146/annurev-soc-090320-111734.
Lowder, Evan M., Spencer G. Lawson, Eric Grommon, and Brad R. Ray. 2020. "Five-County Validation of the Indiana Risk Assessment System–Pretrial Assessment Tool (IRAS-PAT) using a Local Validation Approach." *Justice Quarterly* 37(7):1241–60. https://doi.org/10.1080/07418825.2020.1829006.
Lynch, Mona. 2019. "The Narrative of the Number: Quantification in Criminal Court." *Law & Social Inquiry* 44(1):31–57. https://doi.org/10.1111/lsi.12334.
McNeill, Fergus, Nicola Burns, Simon Halliday, Neil Hutton, and Cyrus Tata. 2009. "Risk, Responsibility, and Reconfiguration: Penal Adaptation and Misadaptation." *Punishment & Society* 11(4):419–42. https://doi.org/10.1177/1462474509341153.
McPherson, Chad M., and Michael Sauder. 2013. "Logics in Action: Managing Institutional Complexity in a Drug Court." *Administrative Science Quarterly* 58(2):165–96. https://doi.org/10.1177/0001839213486447.
MPI (Mapping Pretrial Injustice). 2022. *Pretrial Risk Assessment Tools*. https://pretrialrisk.com
Nagin, Daniel S., and Cody W. Telep. 2017. "Procedural Justice and Legal Compliance." *Annual Review of Law and Social Science* 13:5–28. https://doi.org/10.1111/1745-9133.12499.
Nir, Esther. 2018. "Approaching the Bench: Accessing Elites on the Judiciary for Qualitative Interviews." *International Journal of Social Research Methodology* 21(1):77–89. https://doi.org/10.1080/13645579.2017.1324669.
Pache, Anne-Claire, and Filipe Santos. 2010. "When Worlds Collide: The Internal Dynamics of Organizational Responses to Conflicting Institutional Demands." *Academy of Management Review* 35(3): 455–76. https://doi.org/10.5465/amr.35.3.zok455.
Page, Joshua, and Christine S. Scott-Hayward. 2022. "Bail and Pretrial Justice in the United States: A Field of Possibility." *Annual Review of Criminology* 5(1): 91–113. https://doi.org/10.1146/annurev-criminol-030920-093024.
Porter, Theodore M. 1995. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press.
Rabuy, Bernadette, and Daniel Kopf. 2016. *Detaining the Poor: How Money Bail Perpetuates an Endless Cycle of Poverty and Jail Time*. Northampton, MA: Prison Policy Initiative.
Reay, Trish, and C. Robert Hinings. 2009. "Managing the Rivalry of Competing Institutional Logics." *Organization Studies* 30(6): 629–52. https://doi.org/10.1177/0170840609104803.
Sawyer, Wendy. 2019. *How Race Impacts Who is Detained Pretrial.* Northampton, MA: Prison Policy Initiative.
Sawyer, Wendy, and Peter Wagner. 2022. *Mass Incarceration: The Whole Pie 2022*. Northampton, MA: Prison Policy Initiative.
Schlesinger, Traci. 2005. "Racial and Ethnic Disparity in Pretrial Criminal Processing." *Justice Quarterly* 22(2): 170–92. https://doi.org/10.1080/07418820500088929.
Scott-Hayward, Christine S., and Henry F. Fradella. 2019. *Punishing Poverty: How Bail and Pretrial Detention Fuel Inequalities in the Criminal Justice System*. Oakley, CA: UC Press.
Smith, Sandra S., and Cathy Hu. 2021. "Exploring the Causal Mechanisms Linking Pretrial Detention and Future Penal System Involvement." Pp. 88–109 in *Handbook on Pretrial Justice*, edited by Christine S. Scott-Hayward, Jennifer E. Copp, and Stephen Demuth. New York: Routledge.
Starr, Sonja B. 2015. "The Risk Assessment Era: An Overdue Debate." *Federal Sentencing Reporter* 27(4):205–06. https://doi.org/10.1525/fsr.2015.27.4.205.
Stevenson, Megan. 2018. "Assessing Risk Assessment in Action." *Minnesota Law Review* 103: 303–84. https://scholarship.law.umn.edu/mlr/58/.

Suddaby, Roy, and Royston Greenwood. 2005. "Rhetorical Strategies of Legitimacy." *Administrative Science Quarterly* 50(1):35–67. https://doi.org/10.2189/asqu.2005.50.1.35.
Sudnow, David. 1965. "Normal Crimes: Sociological Features of the Penal Code in a Public Defender Office." *Social Problems* 12(3):255–76. https://doi.org/10.2307/798932.
Swidler, Ann. 1986. "Culture in Action: Symbols and Strategies." *American Sociological Review* 51(2): 273–86. https://doi.org/10.2307/2095521.
Tankebe, Justice. 2019. "In their Own Eyes: An Empirical Examination of Police Self-Legitimacy." *International Journal of Comparative and Applied Criminal Justice* 43(2):99–116. https://doi.org/10.1080/01924036.2018.1487870.
Taxman, Faye S., and Jill A. Gordon. 2009. Do Fairness and Equity Matter? An Examination of Organizational Justice among Correctional Officers in Adult Prisons. *Criminal Justice and Behavior* 36(7):695–711. https://doi/10.1177/0093854809335039.
TCF (The Civic Federation). 2022. *Observations of Cook County Felony Bond Court Hearings*. Chicago: TCF.
Terranova, Victoria A., Kyle Ward, Anthony M. Azari, and Jessie Slepicka. 2020. "Practitioner Perceptions of the Use and Utility of Pretrial Risk Assessment: Focus Group Analysis." *Federal Probation* 84(3):34–40.
Terranova, Victoria A., Kyle Ward, Jessie Slepicka, and Anthony M. Azari. 2020. "Perceptions of Pretrial Risk Assessment: An Examination across Role in the Initial Pretrial Release Decision." *Criminal Justice and Behavior* 47(8):927–42. https://doi.org/10.1177/0093854820932204.
Thornton, Patricia H., and William Ocasio. 1999. "Institutional Logics and the Historical Contingency of Power in Organizations: Executive Succession in the Higher Education Publishing Industry, 1958–1990." *American Journal of Sociology* 105(3): 801–43. https://doi.org/10.1086/210361.
Timmermans, Stefan, and Iddo Tavory. 2012. "Theory Construction in Qualitative Research: From Grounded Theory to Abductive Analysis." *Sociological Theory* 30(3):167–86. https://doi.org/10.1177/0735275112457914.
Tyler, Tom R. 1990. *Why People Obey the Law*. New Haven, CT: Yale University Press.
Ulmer, Jeffrey T. 2019. "Criminal Courts as Inhabited Institutions: Making Sense of Difference and Similarity in Sentencing." *Crime and Justice* 48(1):483–522. https://doi.org/10.1086/701504.
Van Cleve, Nicole Gonzalez. 2016. *Crook County: Racism and Injustice in America's Largest Criminal Court*. Stanford, CA: Stanford University Press.
Viljoen, Jodi L., Melissa R. Jonnson, Dana M. Cochrane, Lee M. Vargen, and Gina M. Vincent. 2019. "Impact of Risk Assessment Instruments on Rates of Pretrial Detention, Postconviction Placements, and Release: A Systematic Review and Meta-Analysis." *Law and Human Behavior* 43(5):397. https://doi.org/10.1037/lhb0000344.
Wenzelburger, Georg, and Pascal D. König. 2022. "The Liberal Dream of Smart Detention? Algorithms and the Politics of Pretrial Detention in the US States." *Law & Policy* 44(4):325–47. https://doi.org/10.1111/lapo.12197.
Werth, Robert. 2017. "Individualizing Risk: Moral Judgement, Professional Knowledge and Affect in Parole Evaluations." *British Journal of Criminology* 57(4):808–27. https://doi.org/10.1093/bjc/azw025.
Werth, Robert. 2019. "Risk and Punishment: The Recent History and Uncertain Future of Actuarial, Algorithmic, and 'Evidence-Based' Penal Techniques." *Sociology Compass* 13(2): e12659. https://doi.org/10.1111/soc4.12659.
Winter, Alix S., and Matthew Clair. 2023. "'The Roughest Form of Social Work:' How Court Officials Justify Bail Decisions." *Criminology* 61(4):904–28. https://doi.org/10.1111/1745-9125.12350.
Zimring, Franklin E., Gordon Hawkins, and Sam Kamin. 2001. *Punishment and Democracy: Three Strikes and You're Out in California*. New York: Oxford University Press.
Zottola, Samantha A., Sarah E. Duhart Clarke, and Sarah L. Desmarais. 2022. "Bail Reform in the United States: The What, Why, and How of Third Wave Efforts." Pp. 143–69 in *Handbook of Issues in Criminal Justice Reform in the United States*, edited by Elizabeth Jeglic and Cynthia Calkins. New York: Springer.
Zozula, Christine. 2019. *Courting the Community: Legitimacy and Punishment in a Community Court*. Philadelphia, PA: Temple University Press.