Tipping the Balance: Predictive Algorithms and Institutional Decision-Making in Context

Simone Zhang*

August 31, 2024

Abstract

Predictive algorithms inform institutional decisions about individuals that shape the distribution of benefits and burdens in society. How do these technologies influence decision-making practices? This article argues that predictive algorithms can disrupt the balance of multiple goals central to many institutional decisions by requiring specific, measurable outcomes to model. When incorporated into deliberations among decision-making actors, algorithms add a voice that endorses a narrowed set of objectives, anchoring attention and empowering actors whose perspectives align with the algorithm's own. I develop this argument through the case of pretrial risk assessment algorithms. Using court hearing transcripts and administrative data from a county that implemented such a tool in a randomized controlled trial, I show risk assessments heighten concern about an adverse outcome they model – missed court dates – and serve as more effective resources for prosecutors seeking harsher pretrial conditions than for defense attorneys. These findings suggest that predictive algorithms can skew the balance of power and objectives in decision-making.

1 Introduction

Institutions and organizations routinely make decisions that affect individuals' life chances. In recent decades, expansions in personal data collection and the growth of machine learning accelerated the use of predictive algorithms to guide choices of all kinds, from whom to hire and help to whom to investigate and incarcerate. The basic idea behind these models is simple: we can infer things about you by learning from the patterns of similar others we have previously encountered. Will you be a top employee (Ajunwa, 2023)? A reliable tenant (Rosen et al., 2021)? A good bet for showing up to court (Goel et al., 2021)? By promising answers to such questions, predictive algorithms are

*Department of Sociology, New York University, simone.zhang@nyu.edu

often positioned as improvements over biased humans and arbitrary bureaucratic rules (Eubanks, 2018; Rona-Tas, 2020; Johnson and Zhang, 2022; Wang et al., 2023).

How do predictive algorithms shape institutional decision-making? Influential macro-level accounts argue that predictive algorithms alter how people are morally evaluated (Simon, 1988; Feeley and Simon, 1992; Fourcade and Healy, 2017). Instead of seeing individuals as moral agents capable of change, institutions come to view people as "locations in actuarial tables" (Simon, 1988, p. 772) determined by models that treat all aspects of peoples' lives – from where they live to their marital status – as risk factors for which they alone are responsible (Starr, 2014; Fourcade and Healy, 2017; Hirschman and Bosk, 2020). Predictive models, devoid of sociological imagination, thus challenge conventional ways of reasoning about merit and blame (Kiviat, 2019, 2023).

Meanwhile, research on how algorithms shape decision-making in practice emphasizes processes of resistance that temper this technical logic. In high-stakes decision-making, algorithms typically assist, rather than replace, human decision-makers (Levy et al., 2021). Past scholarship has highlighted how these decision-makers counter or adapt algorithmic information in ways that preserve their professional discretion and established evaluation practices (Kiviat, 2017; Werth, 2019a; Bosk, 2020; Brayne and Christin, 2021), consistent with local reactions to past reforms that put faith in numbers and science to improve institutional decision-making (Ulmer and Kramer, 1998; Espeland and Vannebo, 2007).

Prior literature thus presents accounts that stand in tension, with some suggesting the possibility of broad transformation and others highlighting the stickiness of the status quo. To make progress amid this tension and facilitate the tracking of incomplete transformation, this article further disaggregates and specifies components of evaluation regimes and decision-making structures that predictive models may intervene in. To date, sociological studies have focused on one main component – whether decision subjects are primarily evaluated via their biographical narratives or as instances of algorithmically determined cases – and observed sustained reliance on the former as a means to realign decisions with conventional moral intuitions (Heimer, 2001; Kiviat, 2017; Werth, 2017; Lynch, 2019).

In this article, I identify two further sites where predictive algorithms can intervene: decision-making goals and power among decision-making actors. Situating these algorithms in their organizational context, I first highlight their potential for ***goal funneling*** – narrowing the goals guiding decision-making. I generalize this as a concern that recurs across many settings in which important decisions about individuals are made. Whereas decision-makers in these organizations typically grapple with multiple ambiguous, often competing objectives (Wilson, 1989; Lipsky, 2010; McPherson and Sauder, 2013), prediction models reduce decisions to specific measurable outcomes (Barocas and Selbst, 2016; Johnson and Zhang, 2022). Indeed, across many contemporary applications, they prescribe decisions based on predictions of a single outcome (e.g. Kleinberg et al., 2015; Chouldechova et al., 2018; Obermeyer et al., 2019). As metrics steer attention (Espeland and Stevens, 1998; Espeland and Sauder, 2007), predictive algorithms can skew an organization's working balance of goals by encouraging decision-makers to optimize for the narrowed set of goals that the tools endorse (Harcourt, 2007; Rona-Tas, 2020; Green and Chen, 2021).

Second, I elaborate on predictive algorithms' capacity to relocate power through a process of ***selective empowerment*** in contexts where decision-making involves multiple actors. While prior cross-disciplinary research on such tools commonly analyzes how solo decision-makers engage with algorithms, many predictive decision aids enter organizational contexts where horizontal and vertical relationships structure decision-making. Drawing on insights from scholarship on quantification and technology in organizations, I conceptualize algorithms in these multi-actor settings as rhetorical resources that actors can harness to add a new voice to decision deliberations. This approach moves beyond the common focus on asking whether algorithms constrain professional discretion. Instead, it opens up sociological questions about how the addition of an algorithmic perspective might reconfigure interactional dynamics among decision-makers, such as by enabling new ways to substantiate claims about decision subjects and shore up influence (Lynch, 2019). Critically, however, predictive algorithms are not neutral resources: they embed particular goals and understandings of social problems. As a result, they may empower decision-making actors unevenly, affording more leverage to those whose roles and positions more closely align with the

viewpoints that the algorithms encode.

I develop these arguments and demonstrate these processes empirically through an analysis of pretrial risk assessment algorithms, a popular yet controversial technology used in criminal legal systems across the United States. Designed to guide decisions about whether and under what conditions an arrested individual may be released before trial, these tools are intended to reduce over-reliance on cash bail and pretrial detention (Goel et al., 2021). However, they have been criticized for their potential to exacerbate or legitimize discriminatory and punitive decisions (Starr, 2014; Van Cleve and Mayes, 2015; Robinson and Koepke, 2019). This article highlights the contrast between ideals – that decisions carefully balance concerns about preventing crime and assuring court appearance against interests in preserving defendants' liberty and minimizing burdens (Stevenson and Mayson, 2021) – and the reality that most pretrial algorithms only predict crime and missed court. Examining courtrooms as multi-actor settings where prosecutors, defense attorneys, and judges interact to process cases, this article further considers how risk assessments can shift power as they endorse optimizing for public safety and preventing flight, aligning with a prosecutorial perspective while remaining silent on countervailing imperatives that could moderate the use of restrictive pretrial conditions.

Accordingly, I ask: how do algorithmic pretrial risk assessments shape attention to different decision-making goals? How do they impact power dynamics in the courtroom? Which actors' positions do they strengthen and, ultimately, does this lead to harsher or more lenient decisions?

To address these questions, I examine a widely used pretrial risk assessment algorithm, the Public Safety Assessment (PSA), and its impact in a Midwestern county where the tool was implemented within the framework of a randomized controlled trial. In this county, the court randomly received PSA reports for around half of all cases. I augment this pre-existing field experiment with data on decision processes and outcomes, including administrative court records and an original set of court hearing transcripts containing verbatim records of courtroom deliberations.

Comparing cases where the PSA was provided and withheld, I demonstrate that risk assessments empower prosecutors to more easily secure their requests for more restrictive conditions and enable

judges to more often depart from their default of rendering decisions consistent consensuses between the prosecution and defense. When the tool recommends cash bail, judges are more likely to order it even when both sides deem it unnecessary. I further show that the algorithm causes courtroom actors to pay greater attention to the risk of missed court dates without offsetting increases in attention to competing aims. While this ostensibly aligns deliberations with the formal legal basis for cash bail in the state – assuring court appearance – I illustrate how it casts defendants in a harsher light on net. Overall, risk assessments can have an asymmetric effect: defendants are punished for high risk ratings but receive minimal benefit from low risk ratings.

This study builds on efforts to understand how predictive algorithms operate on the ground in social context (Selbst et al., 2019; Brayne and Christin, 2021; Joyce et al., 2021; Anthony et al., 2023). It speaks to enduring questions on the relationship between formalization, accountability, and institutional bias (Weber, 1978; Espeland and Vannebo, 2007; Kalev, 2014), elaborating on the distinctive tensions that an ascendant type of technologically enabled formalization presents. By juxtaposing the goal specificity required by predictive algorithms with the goal multiplicity and ambiguity that characterize many sites of people-processing, I position algorithms as interventions that influence both the means and ends of decision-making. Algorithms can empower actors most aligned with the goals the tools endorse, shaping what dimensions of complex individuals become salient in decision deliberations. These findings highlight the downstream consequences of upstream formal design choices and underscore the value of examining how algorithms reconfigure power and the balance of goals across implementation contexts.

Methodologically, this research illustrates the value of pairing records of organizational deliberation with experiments to unpack the mechanisms behind aggregate outcomes. From a policy standpoint, this research speaks to the impacts of a popular reform affecting the legal experiences and lives of people awaiting trial. It illuminates how such tools shape how arrested individuals are evaluated in the bail process. It further suggests that simulation studies showing that algorithms could curb unwarranted uses of cash bail and pretrial detention (Kleinberg et al., 2018) may be overly optimistic because decision-makers can incorporate algorithmic information unevenly in

ways that magnify institutional risk aversion.

2 Decision-Making and Predictive Modeling Amidst Goal Multiplicity

Efforts to improve institutional decision-making have long depended on crafting rules, guidelines, and procedures to assess and sort people in more standardized ways (Porter, 1995; Espeland and Stevens, 1998; Pager and Shepherd, 2008; Hirschman and Bosk, 2020). From this standpoint, modern predictive decision-making algorithms have much in common with cases of classification and quantification that have animated sociological inquiry across substantive domains, from the U.S. federal sentencing guidelines to scoring systems used in college admissions and employment decisions (Espeland and Vannebo, 2007; Dobbin et al., 2015; Hirschman et al., 2016).

Yet, predictive models present a distinctive paradigm for how rule systems are constructed. Many conventional formalization projects rely on human experts or policymakers to manually define rules and guidelines for making decisions about people. For instance, earlier criminal justice risk assessments were devised by experts who drew on theory, research, and political judgment to identify factors criminal justice actors should weigh when evaluating legal subjects (Gottfredson and Moriarty, 2006). In hiring, an employer might craft a rubric that details desirable candidate characteristics and how much each should count in a candidate's favor (Uhlmann and Cohen, 2005; Blair-Loy et al., 2022).

The broad class of models that are the focus of the present paper, by contrast, construct rules inductively according to how much they help to predict a particular chosen outcome – commonly called the target variable – in examples in historical data.[1] In criminal justice, models commonly predict outcomes like a defendant re-offending or returning to court. In hiring, models might forecast some measure of job performance. Using records of earlier defendants or employees,

[1] "Actuarial" is often used in economic sociology and criminology to refer to similar approaches. This paper uses the more general term of "predictive" given its broad usage in interdisciplinary literatures and because "actuarial" is commonly associated with managing risk, whereas this approach is also applied to tasks beyond characterizing risk.

statistical or machine learning algorithms identify attributes most associated with the outcome of interest and learn rules for accurately predicting the target. Model-generated predictions, in turn, are used to prescribe decisions.[2] Given enough examples of hires tagged as high performers, for instance, predictive algorithms promise, as one software company claimed, to help employers "clone your best, most reliable people" (Ajunwa, 2023). Emerging in the post-War era, this predictive modeling approach has grown in popularity. Recent decades have seen it permeate wide-ranging domains of decision-making, buoyed by advances in machine learning methods adept at prediction using large, high-dimensional datasets (for a review, see Wang et al., 2023).

The rules that predictive models learn may be simple or complex, transparent or opaque; what sets them apart is that they are set inductively to maximize prediction accuracy with respect to specifically selected targets, rather than devised manually by people (Breiman, 2001; Kleinberg et al., 2015; Wang et al., 2023). I argue that this shift – from specifying rules themselves to specifying the ends that rules should serve – is critical to understanding the distinctive intervention that predictive algorithms present for institutional decision-making. It brings to the fore a key question when we analyze such tools in their organizational context: among all the various goals that organizational decision-makers might pursue or juggle, which ones do predictive algorithms endorse?

I derive this concern from the key insight from organizational sociology and public administration that people-processing organizations often manage multiple ambiguous, sometimes conflicting mandates when making decisions (Wilson, 1989; Lipsky, 2010; McPherson and Sauder, 2013; Zacka, 2017). For example, child welfare agencies are charged with protecting children from maltreatment and with preserving and strengthening families (Berrick, 2017; Fong, 2020). College admissions strives to uphold academic standards and reputation, enhance university finances, and expand equitable access to opportunity (Rona-Tas, 2020; Chu, 2021; Engler, 2021). Courts making

[2]This class of models is distinct from other interventions commonly associated with "automated decision-making," such as the automated application of pre-existing human-specified decision rules (e.g., screening out job applicants lacking a required credential) and models that aim to replicate the decisions that human decision-makers have made previously (e.g., a model that selects applicants to interview based on whom that employer has previously interviewed). Other uses of prediction beyond institutional decision-making, such as in organizing social media feeds or product recommendations, also fall outside the case discussed here.

pretrial decisions aim to balance interests in public safety and the rule of law against a defendant's right to liberty (Stevenson and Mayson, 2021). Good decision-making is commonly thought to achieve a balance among competing imperatives (Zacka, 2017).

Whereas manual human-generated rules can be devised even when decision-making goals are under-specified and heterogeneous, predictive modeling requires explicitly selecting measurable outcomes to model. Indeed, its stated promise and appeal lies in distilling complex decisions down to concrete prediction tasks (Kleinberg et al., 2015). Across a wide array of contemporary examples, this distillation entails a narrowing of goals: child welfare algorithms predict maltreatment, not the toll of unnecessary family separation or the gains of added family supports (Chouldechova et al., 2018; Bosk, 2020); admissions and financial aid algorithms model yield and tuition, not student success (Engler, 2021); and pretrial tools predict non-appearance and future crime, not the harms that pretrial detention could inflict on defendants (Gouldin, 2016).[3]

As quantification steers attention toward what is measured (Espeland and Vannebo, 2007; Espeland and Stevens, 1998; Espeland and Sauder, 2007; Chu, 2021), I contend that predictive algorithms can initiate *goal funneling*, elevating the goals that target variables represent while potentially overshadowing other important objectives in decision-making (Rona-Tas, 2020; Green and Chen, 2021). Presenting a particular architecture of evaluation (Rivera and Tilcsik, 2019), predictive models serve as symbolic anchors that orient decision-makers toward model-endorsed goals and lower the costs of seeing people through the lenses of the outcomes they promote. In the case of pretrial, nearly all risk assessment algorithms predict outcomes like "failure to appear," "new criminal activity," or a generalized "pretrial failure" (Gouldin, 2016). They enter high-volume, low-information decision-making environments as courts must typically make pretrial decisions shortly after arrest. Offering readily digestible summaries of an arrested person's future propensity for missing court or committing criminal offenses, these tools can facilitate a narrowed focus on flight and public safety by making it easier to evaluate individuals according to those objectives.

[3]A related issue is that available data only approximate goal constructs. In criminal justice, complex concepts like flight risk are proxied by metrics like missed court dates, which thinly capture the underlying legal constructs (Gouldin, 2016, 2018). See Barocas and Selbst (2016) and Jacobs and Wallach (2021) for a broader discussion.

Does goal funneling improve institutional decision-making or exacerbate existing pathologies? On the one hand, specifying a single or a limited set of clear objectives can resolve ambiguities in mission that open room for favoritism and bias, creating concrete sites for contestation, oversight, and advocacy (Abebe et al., 2020; Johnson and Zhang, 2022). Some legal scholars, for instance, have posited that well-designed pretrial risk assessment may curb arbitrary or discriminatory decision-making by explicitly linking decisions to specific legally recognized grounds for restrictive pretrial conditions (Gouldin, 2016). On the other hand, predictive models may oversimplify complex problems and distort goals to fit measurable technical knowledge (March and Simon, 1958; Levy et al., 2021). Punishment scholars, notably, have warned that reliance on tools emphasizing incapacitating those deemed statistically likely to re-offend could crowd out other theories of just punishment like retribution, rehabilitation, and deterrence (Feeley and Simon, 1992; Harcourt, 2007). In the case of pretrial, some scholars have raised concerns that widely used target variables frame people through the lens of failure, systematically increasing individuals' perceived riskiness (Werth, 2019b; Robinson and Koepke, 2019) while diverting attention from countervailing considerations that might pull decisions in a less punitive direction (Green and Chen, 2021).

More broadly, while goal funneling describes a generic process applicable to most predictive decision tools, its impact on bias and arbitrariness in decision-making will depend on how the particular goals it elevates interact with specific adoption environments. Despite the open-ended nature of such interactions, however, it is worth noting that target variable choices exhibit regularities across real-world applications. Although prediction is a versatile framework that could serve a variety of objectives in theory (O'Malley, 1992; Maurutto and Hannah-Moffat, 2006), in practice, the technologists, policymakers, and organizational leaders who have a say in modeling choices frequently select target variables that mirror status quo approaches to societal and organizational dilemmas (Barabas et al., 2020). This tendency is structured in large part by the availability of data. As organizations selectively and strategically decide what information to collect, the data infrastructures on which predictive models rely reflect organizations' pre-existing commitments, interests, and preoccupations (Scott, 1998; boyd and Crawford, 2012; Gitelman, 2013; Brayne,

2020; Levy et al., 2021). These infrastructures shape what is understood as "measurable" and constrain the plausible pool of target variables (Passi and Barocas, 2019). The result is that predictive algorithms commonly spotlight the risk or value that individuals present to organizations and systems, rather than the risk or value that organizations and systems present to individuals (Barabas et al., 2020). Thus, as predictive models are commonly constructed today, goal funneling may tend to magnify existing forms of institutional risk aversion and loss orientation.

3 Predictions and Selective Empowerment in Multi-Actor Decision-Making

Interdisciplinary research on the practical impacts of predictive decision aids commonly analyzes decision-making as an individualistic process, with many studies focusing on how a lone decision-making agent engages with tools meant to "nudge" their choices (Chouldechova et al., 2018; Rosen et al., 2021; Lebovitz et al., 2022). Many recent studies of criminal justice risk assessments exemplify this approach, investigating how individual judicial officers interact with these algorithms (Kleinberg et al., 2018; Stevenson, 2018; Lowder et al., 2021; DeMichele et al., 2021; Imai et al., 2023; Albright, 2023).

However, this focus on individual focal decision agents commonly neglects the multi-actor nature of much institutional decision-making (Laufer et al., 2023), which can implicate relationships between decision-makers and their lateral peers, decision-makers and their superiors and subordinates, decision-makers and decision subjects, and in settings of collaborative decision-making, team members with differentiated roles. A rich prior literature on technologies and organizations underscores the importance of analyzing how predictive tools might refract and reconfigure the social dynamics among these parties (Barley, 1986; Orlikowski, 2000; Christin, 2020; Anthony et al., 2023). Ethnographies reveal, for instance, how these tools can reinforce organizational hierarchies and heighten perceptions of managerial control, generating internal resistance (Brayne and Christin, 2021). Thus, to analyze predictive algorithms in organizational context, scholars should

attend to the tools' influence on the interactional dynamics between decision-making actors, where the balance between abstract goals is continually negotiated (Ulmer, 2019).

In most jurisdictions, pretrial decisions are made as multiple strategic actors – commonly judges, prosecutors, and defense attorneys – convene in hearings to assemble information, interpret facts, and advocate for specific pretrial conditions guided by their formal roles. Prosecutors aim to protect public safety, defense attorneys to safeguard their client's interests and rights, and judges to make decisions that are seen as fair and impartial. However, while these actors have distinct roles and goals, courtroom ethnographies consistently emphasize that they often coordinate and defer to each other to process caseloads efficiently, rather than engage in regular adversarial conflict (Eisenstein and Jacob, 1977; Feeley, 1979; Emerson, 1983; Van Cleve, 2016). Courts can thus be understood as local social orders where actors with diverse commitments collaborate to get the work of the court done, creating templates for argumentation, going rates for different types of cases, and case-handling norms (Ulmer and Kramer, 1998; Ulmer, 2019). Micro-level studies further highlight the collective effort needed to manage a court's institutional complexity (McPherson and Sauder, 2013). Actors with varying roles and professional backgrounds creatively employ and mutually recognize each others' home institutional logics to promote their interests and find common ground. Accordingly, a court's existing practices and decision patterns reflect its working local solution to balancing multiple institutional demands and goals.

I contend that predictive algorithms can intervene in these existing local social arrangements by injecting an additional voice or perspective into multi-actor decision-making processes. This added presence opens new social possibilities for alliances, mediation, and distance between actors that can reconfigure dynamics within a decision-making group (Simmel, 1902). Predictive scores and recommendations can create symbolic openings (Christin, 2020) or supply rhetorical material that courtroom actors can harness to tell stories about legal subjects (Lynch, 2019) and advance their own ends (Ulmer and Kramer, 1998). Algorithms can further telegraph expectations about the terms on which actors should engage with one another (Traeger et al., 2020).

As the preceding section suggests, however, predictive algorithms are not infinitely malleable,

neutral resources. Instead, they may engender what I refer to as *selective empowerment*, in which the predictions they supply afford greater advantage to actors whose roles or perspectives resonate most with the underlying logic of the algorithm. In pretrial, actors could ostensibly harness quantified summaries of defendants to bolster a variety of arguments (Lynch, 2019), but risk assessments usually endorse a risk management logic prioritizing public safety that mirrors a prosecutorial perspective.They may thus be more effective when used to advocate for more restrictive pretrial conditions (Winter and Clair, 2023), as their design constrains the range of arguments they readily support (McPherson and Sauder, 2013). In this way, predictive algorithms could shift power at the micro-level, strengthening the position of prosecutors relative to the defense.

Risk assessments may, moreover, diminish judges' reliance on prosecutors and defense attorneys, traditionally the primary parties who shape how a decision appears before them (Page and Scott-Hayward, 2022). Risk assessments offer new possibilities for interactional rupture by adding a seemingly independent, impartial perspective that may contradict both the defense and prosecution. Access to a risk assessment may offer judges political insulation (Hannah-Moffat, 2013; Werth, 2017; Albright, 2023) and make it easier for judges to legitimize decisions that diverge from the positions of either side.

More broadly, risk assessments and similar predictive decision aids can create different opportunities and constraints for actors to align decisions with their perspectives and imperatives. Negotiations, conflicts, and compromises among decision stakeholders working with predictive tools can mold the moral foundation of decisions, even as the enlistment of these tools might disguise value judgments as impersonal technical exercises.

4 Pretrial Decision-Making and The Public Safety Assessment

Pretrial risk assessments are designed to inform decisions about whether and under what conditions an arrested person may be released before trial. While pretrial options vary, nearly all U.S. jurisdictions enable judges to set cash bail, a sum of money that must be deposited with the court

for a defendant to be released (NCSL, 2023). In most states, judges can also detain an individual without the possibility of release in cases involving serious charges. Alternatively, defendants may be released without financial conditions, with courts merely requiring that defendants sign an agreement to return to court. Judges may also separately impose non-monetary release conditions, such as electronic monitoring, no-contact orders, or drug testing.

Defendants at the pretrial stage are legally innocent, and the Supreme Court has viewed pretrial decision-making as a regulatory measure to protect public interests rather than a penal measure to punish alleged crimes (Smith, 2019; Page and Scott-Hayward, 2022). Traditionally, the legal basis for cash bail was providing collateral to ensure court appearance. However, Congress and state legislatures began to amend bail statutes in the 1970s and 1980s to allow public safety as a basis for detention. This shift aligned with rising public concern about crime and existing informal practices at the time, but legal scholars have warned that these changes can lead to harsher decisions as courts conflate flight and public safety concerns (Gouldin, 2016).

Despite bail statutes that favor release or the least restrictive conditions necessary to prevent nonappearance and pretrial crime, current pretrial practices often cast defendants in a state of liminal guilt (Winter and Clair, 2023). Cash bail and pretrial detention are widespread and disproportionately applied to racially minoritized and low-income defendants (Arnold et al., 2018; Didwania, 2021). The broader cash bail system is deeply flawed, tying freedom to financial circumstances while failing to guarantee court appearances effectively (Ouss and Stevenson, 2023). Pretrial detention, moreover, exacts a heavy toll on defendants' legal outcomes, relationships, economic security, political participation, and well-being (Sugie and Turney, 2017; Dobbie et al., 2018; Wakefield and Andersen, 2020; McDonough et al., 2022)

Pretrial risk assessments have been positioned as potential correctives for over-reliance on cash bail, unnecessary detention, and racial biases in decision-making (Kleinberg et al., 2018; Jorgensen and Smith, 2021). They date back to the 1960s (Ares et al., 1963) but have surged in popularity: one in four people in the U.S. lived in jurisdictions that use them in 2017—up from one in ten in 2013 (PJI, 2017). Pretrial risk assessments vary in design and the degree to which they structure

decision-making (Skeem and Monahan, 2011). Contemporary tools increasingly rely on predictive modeling, stimulating renewed attention to their technical characteristics, such as their predictive performance and potential for racial bias (Angwin and Larson, 2016; Desmarais et al., 2021; Goel et al., 2021).

What is known about how pretrial risk assessments impact decision-making in practice? A substantial older body of literature on risk assessments in criminal justice focused primarily on tools used in sentencing and parole, suggests that they do not entirely replace the traditional methods professionals use to evaluate legal subjects (for a review, see Werth, 2019a). Notably, several ethnographic studies have found that risk assessments may be actively resisted or ignored (Werth, 2017; Brayne and Christin, 2021). Yet, meta-analyses (Viljoen et al., 2019) and surveys of actors involved in pretrial decision-making (DeMichele et al., 2019; Terranova et al., 2020) indicate uptake of risk assessment varies across tools and local contexts. Where pretrial risk assessments are consulted, their impacts are mixed: some studies report negligible overall effects (Imai et al., 2023; Copp et al., 2022) while others show reduced reliance on monetary conditions (Lowder et al., 2021) that erode over time (Stevenson, 2018). Beyond overall effects, the literature indicates that higher scores and more stringent algorithmic recommendations lead to harsher decisions (Stevenson, 2018; Albright, 2023). Courts may apply risk assessments inconsistently, however, showing greater leniency toward white defendants (Copp et al., 2022; Albright, 2023; Zottola et al., 2023), women (Imai et al., 2023), and wealthier defendants (Skeem et al., 2020).

This study expands on this existing literature on pretrial risk assessments in several ways. Using data from the first randomized controlled trial testing the tools' impact (Greiner et al., 2020; Imai et al., 2023), this study is better positioned to isolate the causal effect of risk assessment provision than past quantitative studies. These earlier studies have relied on matching observably similar cases or over-time comparisons, where the risk of confounding is high because risk assessments are frequently implemented alongside other reforms (Viljoen et al., 2019). Additionally, by coupling this experimental design with court hearing transcripts, I can directly observe and examine how risk assessments shape deliberation and decision-making processes, which are obscured or inferred

indirectly in studies relying on administrative records or interviews. This enables me to assess the procedural concerns raised in legal and theoretical scholarship on risk assessments and contribute to the nascent qualitative sociological literature on U.S. bail processes (Ottone and Scott-Hayward, 2018; Winter and Clair, 2023).

4.1 Public Safety Assessment

This article focuses on the Public Safety Assessment (PSA), the most popular pretrial risk assessment in the U.S. Created with the support of Arnold Ventures, a private philanthropy, the PSA purports to help judges better calibrate decisions to a defendant's "objective" risk (Desmarais and Lowder, 2019). To date, four states and over 90 cities and counties have implemented the PSA, covering over 82 million people in the U.S. (Advancing Pretrial Policy and Research, 2023).

The PSA produces two 1-to-6 scores using administrative court records: one for risk of failing to appear in court (FTA) and another for risk of re-arrest during pretrial release (formerly "New Criminal Activity," now "New Criminal Arrest" or NCA). It also flags defendants deemed likely to commit a violent crime ("New Violent Criminal Arrest" or NVCA). These outputs are based on subsets of nine factors reflecting the defendant's age, the nature of the current charges, other pending charges, court appearance history, prior felony and misdemeanor convictions, and prior incarceration (detailed list in the Appendix).

These inputs were selected based on their predictive power in a dataset that pooled records from over 300 U.S. jurisdictions (DeMichele et al., 2020). Inputs are assigned whole-number weights, which are summed and scaled to determine the final score. The scoring system shared across all implementation sites is shown in Appendix Figure A1.

Each combination of scores maps on to a recommendation determined by locally developed rubrics called Decision-Making Frameworks (DMFs). Figure 1 presents Dane County's DMF. Low FTA and NCA scores correspond to recommendations for pretrial release without monetary conditions, while higher scores correspond to recommendations for higher cash bail. In many jurisdictions, additional local rules recommend tougher conditions for specific serious offenses.

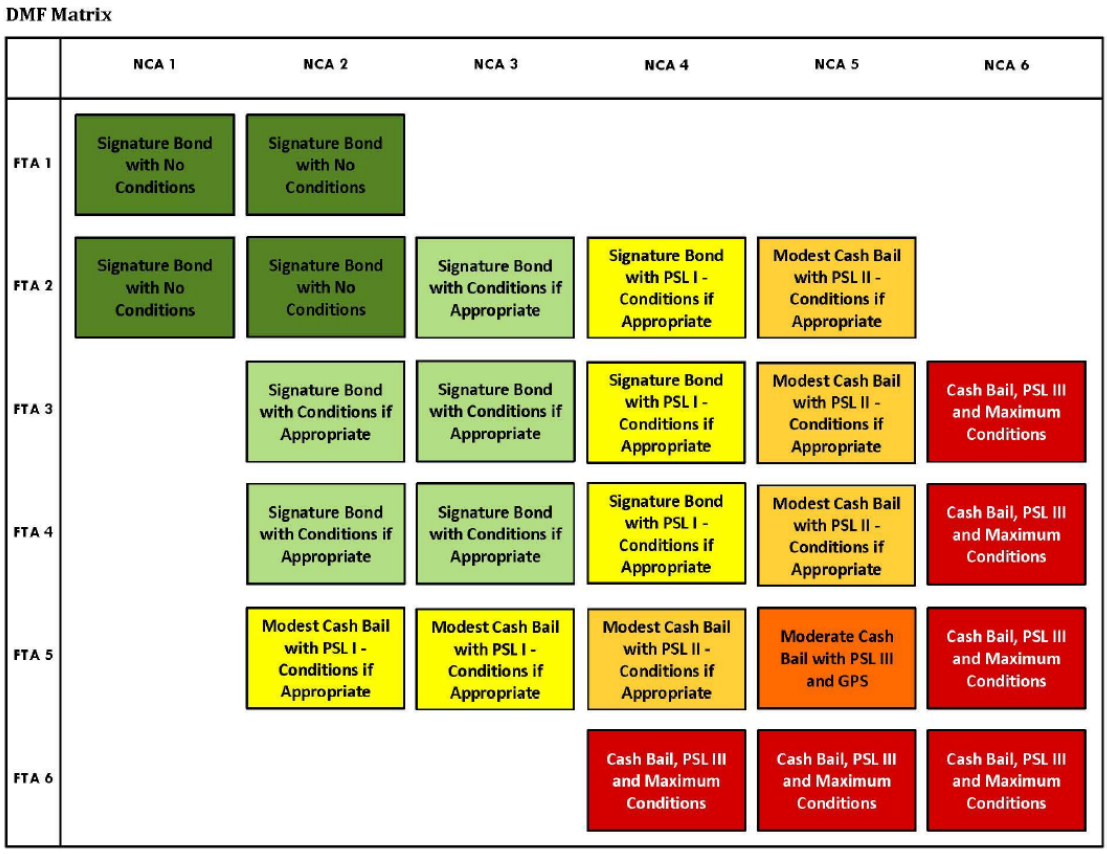


Figure 1: **Dane County Decision Recommendation Rubric**: Recommendations proceed from the least restrictive in dark green to the most restrictive in red. NCA and FTA represent risk assessment scores for New Criminal Activity and Failure to Appear respectively. PSL stands for pretrial supervision level, with higher supervision levels imposing stronger non-monetary conditions, such as more frequent telephone or in-person check-ins with pretrial officers.

The PSA's relative simplicity and transparency may diverge from common portrayals of predictive algorithms as complex and opaque, but it adheres to the core structure of the class: it is an algorithm, "a set of rules that precisely define a sequence of operations," (Stone, 1972) that is optimized to predict a chosen target outcome based on patterns in historical data (Breiman, 2001; Kleinberg et al., 2015). Moreover, the PSA is representative of a broad swath of consequential decision-support systems developed and used in the public and nonprofit sectors. Many predictive tools in the criminal legal system, education, child welfare, and housing continue to rely on variants of linear regression or simple decision trees rather than more complex machine learning methods (Sarver et al., 2015; Picard-Fritsche et al., 2017; Saxena et al., 2020; Vaithianathan and Kithulgoda,

2020; Soland et al., 2020).[4] Further, as the PSA affects millions of U.S. residents, it is valuable both for studying how predictive algorithms operate in practice and for its policy implications.

5 Study Setting

I study the impact of the PSA in Dane County, Wisconsin. Home to the University of Wisconsin-Madison, Dane County is a Democratic stronghold in a politically diverse state with a strong history of progressive organizing on criminal justice issues. While the county is predominantly white, with around 5% of the population identifying as Black or African American and 6% identifying as Hispanic or Latino as of 2019 (US Census Bureau, 2019), Black individuals comprise around half of the county jail population (Austin and Ocker, 2021). Recent local initiatives have pushed for pretrial justice reforms to reduce this racial inequality and better align the local administration of pretrial justice with best practices (Greiner et al., 2020). The PSA was one such reform.

Wisconsin law primarily offers judges two options: cash bail (defendants must post cash for release) or a signature bond (release without a cash deposit).[5] Statewide, judges are instructed to operate on a presumption of release. Cash bail should only be imposed if and in the amount deemed necessary to assure future court appearances. Although judges are not legally supposed to take public safety concerns into account in setting monetary conditions, practices vary and often depart from this law on the books (Lavigne et al., 2018). For instance, prosecutors and judges often argue that being charged with a serious violent crime that is likely to result in significant incarceration can create incentives for defendants to flee, blending concerns about public safety and court appearance.

[4]This situation may persist as public sector entities often have limited data infrastructure to support more demanding methods (Soland et al., 2020) and emerging evidence suggests that for a variety of important outcomes, the performance gains from more complex computational procedures over simple rules may be modest (FICO Decisions, 2018; Jung et al., 2020; Salganik et al., 2020)

[5]Wisconsin law offers mechanisms to detain defendants before case disposition without the possibility of release, but they are burdensome and not exercised in practice.

5.1 Local Procedures and PSA Implementation

In Dane County, the PSA is used in Initial Appearances and Bail/Bond Hearings for individuals in custody at the time of the proceedings. Initial Appearances, typically the first hearing after a new arrest, are held within 48 hours of arrest. Bail/Bond Hearings occur when an event in an ongoing case triggers a review of pretrial conditions, such as when a defendant is arrested on a warrant following a missed court appearance.

Hearings take place in a cramped courtroom in the county jail.[6] Around twenty to forty cases are processed daily (Greiner et al., 2020), with most lasting around four to eight minutes each. Most arrested individuals are represented by the public defender assigned to handle cases that day. They are generally advised to remain silent to avoid self-incrimination. As a result, hearings mainly involve discussions among the prosecutor, defense attorney, and commissioner.[7]

While hearings vary in structure, they typically follow the order of events depicted in Figure 2. Commissioners begin by reviewing the criminal complaint, a document prepared by the district attorney's office that outlines the charges against the defendant. They then ask the prosecution and defense to argue for particular pretrial release conditions. Commissioners might follow up with additional questions or proceed directly to a decision. PSA reports in Dane County are usually distributed shortly before the day's hearings, leaving little time for prosecutors, defense attorneys, and commissioners to review the materials before court. Figure 3 offers a sample PSA report.

6 Research Design and Data

This study builds on a randomized controlled trial (RCT) in Dane County, WI. The RCT randomized whether the PSA report for a given case was shared with the judge, prosecutor, and defense attorney: PSA reports were distributed in even-numbered cases (treated group) but withheld in odd-numbered

[6] In some jurisdictions, pretrial decisions are made without a formal hearing. In Kentucky, for instance, they are made via a telephone call between a judge and pretrial services officer (Albright, 2023).

[7] The primary judicial officers who conduct these hearings are called court commissioners, but I use "commissioner" and "judge" interchangeably in this article.

Figure 2: **Structure of a typical hearing in Dane County:** This figure presents a condensed version of the most central speech act sequence (the centro-type) of my sample of hearing transcripts – the hearing that is the most similar to all other hearings – identified based on a sequence analysis of sentences in transcripts (Abbott and Tsay, 2000). I use optimal matching here but other matching approaches returned similar centro-types.

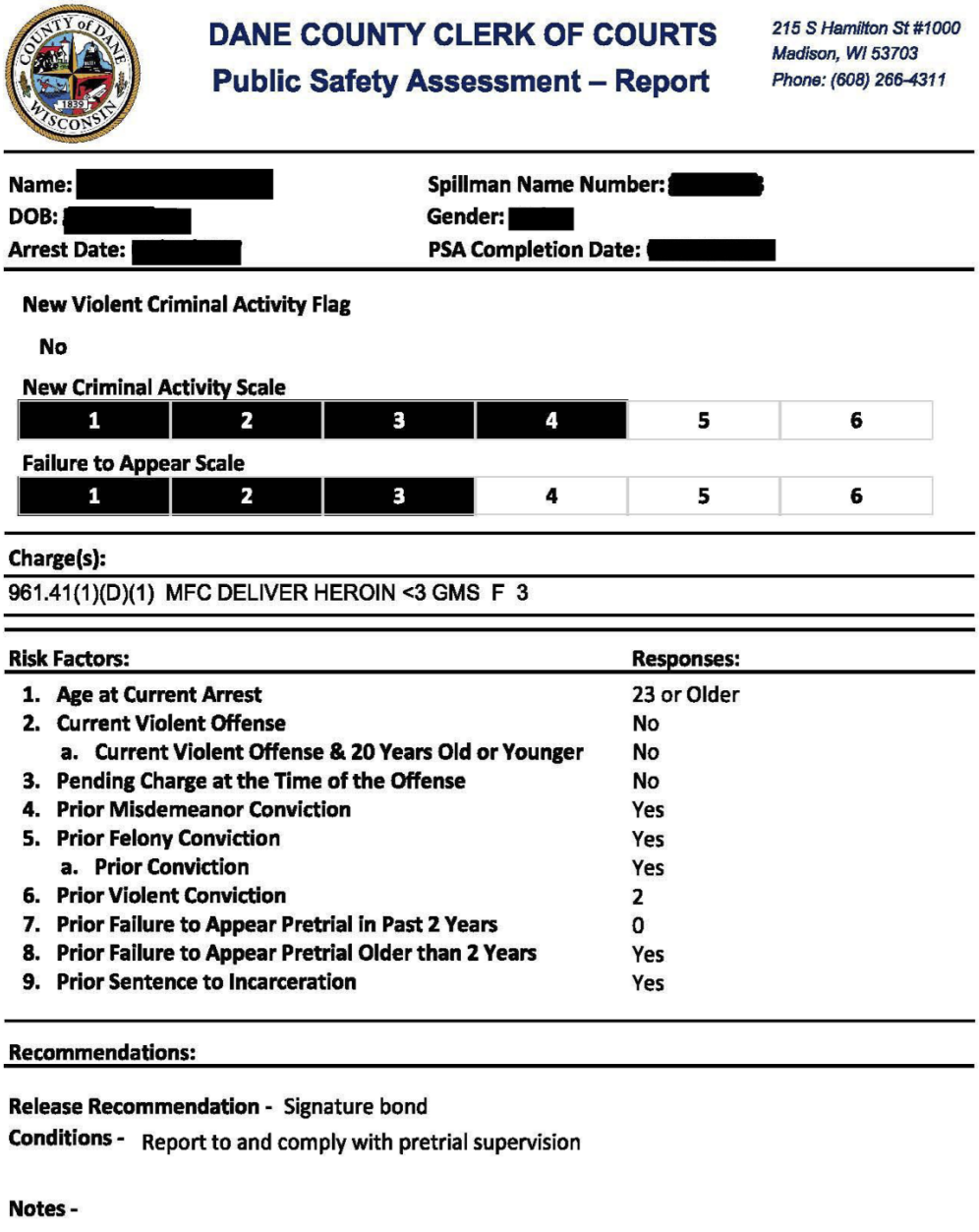
CLERK:	State of Wisconsin versus [redacted], Case No. [redacted]
PROSECUTOR:	State appears by [redacted]
DEFENSE ATTY:	Mr. [redacted] appears personally, in custody, with Attorney [redacted]. We have a copy of the Criminal Complaint, waive its reading...
COMMISSIONER:	... Well, I'll note probable cause has already been found for the charges alleged, and we'll proceed to bail.
PROSECUTOR:	Your Honor, since the defendant gave a Chicago address, State would just request some cash to assure continued appearances...
COMMISSIONER:	Bail recommendation from the defense?
DEFENSE ATTY:	Your Honor, Mr. [redacted]...doesn't have a large amount of money to post for bail. Clearly he sees how important it will be going forward to keep in touch and return for his court dates...
COMMISSIONER:	All right...I don't know if there is a ton of incentive for flight on these charges given how these cases are typically handled...I'm just not going to order cash bail when this is maybe a case that ends with a fine... As to time limits?
DEFENSE ATTY:	We'll enforce them, your Honor.
COMMISSIONER:	We're adjourned.
CLERK:	Prelim will be [redacted] in Branch [redacted]...

cases (control group). The randomized availability of PSA reports creates comparison groups that, in expectation, should be similar in every respect except whether the PSA report was shared, allowing scholars to isolate the causal effect of providing PSA information.

The RCT was implemented by the Access to Justice Lab at Harvard Law School with the initial aim of evaluating how the PSA shapes final case outcomes like failures to appear, pretrial crime, and pretrial detention (Griffin, 2016). The present study is an independent project that leverages this RCT as a strategic research site to observe how risk assessments influence courtroom deliberation and decision-making. It follows the model of past research augmenting policy evaluation RCTs with additional qualitative and quantitative data to examine mechanisms and evaluate impacts on additional outcomes of sociological significance (e.g., Rosenblatt and DeLuca, 2012; De La O, 2013; Urbina, 2020; DeLuca et al., 2023).

Through a data-sharing agreement with the original RCT study team, I gained access to data on the treatment status and PSA reports of all cases from the first year of the RCT (June 1, 2017 - May

Figure 3: **Example of a Dane County PSA Report** Source: Access to Justice Lab RCT Evaluation of the Implementation of the PSA-DMF System in Dane County, WI Interim Report



DANE COUNTY CLERK OF COURTS
Public Safety Assessment – Report

215 S Hamilton St #1000
Madison, WI 53703
Phone: (608) 266-4311

Name: [redacted] **Spillman Name Number:** [redacted]
DOB: [redacted] **Gender:** [redacted]
Arrest Date: [redacted] **PSA Completion Date:** [redacted]

New Violent Criminal Activity Flag

No

New Criminal Activity Scale

1	2	3	4	5	6

Failure to Appear Scale

1	2	3	4	5	6

Charge(s):

961.41(1)(D)(1) MFC DELIVER HEROIN <3 GMS F 3

Risk Factors:	Responses:
1. Age at Current Arrest	23 or Older
2. Current Violent Offense	No
a. Current Violent Offense & 20 Years Old or Younger	No
3. Pending Charge at the Time of the Offense	No
4. Prior Misdemeanor Conviction	Yes
5. Prior Felony Conviction	Yes
a. Prior Conviction	Yes
6. Prior Violent Conviction	2
7. Prior Failure to Appear Pretrial in Past 2 Years	0
8. Prior Failure to Appear Pretrial Older than 2 Years	Yes
9. Prior Sentence to Incarceration	Yes

Recommendations:

Release Recommendation - Signature bond
Conditions - Report to and comply with pretrial supervision

Notes -

31, 2018). I used this information to collect the core data in the present study: a sample of court hearing transcripts containing verbatim accounts of courtroom discussions in treated and control cases. I combine these transcripts with analyses of administrative court records from the Dane County Clerk's Office and Wisconsin Circuit Court Access on defendant backgrounds, case details,

and pretrial conditions imposed. I also supplement these data with field observations and interviews with court staff and commissioners conducted in January 2020, which inform my empirical strategy and guide my interpretation of the results.

I analyze transcripts as reflective of the courtroom decision-making process. While one might worry that courtroom discussions merely reflect surface-level performances or post hoc justifications, several contextual factors lend credibility to my approach. First, during my field work, commissioners told me that they usually have little time to review materials prior to court, so they typically gather information, reason about cases, and make decisions live during hearings. Second, courtroom actors in Dane County had minimal reason to perform for external audiences. Outside the occasional high-profile case that brought in local news crews, there were no systematic court observers and hearings were usually sparsely attended. My transcript collection effort commenced more than two years after the hearings in question. Courtroom actors had no reason to anticipate being observed at the time. These factors combined suggest that transcripts offer insights into the deliberative processes behind decisions.

Transcript Data Collection Details

I collaborated with local court reporters to assemble a sample of hearing transcripts from RCT cases. In Dane County, one primary court reporter was responsible for 83% of Initial Appearances, with two alternates covering days when the primary reporter was out of office. Unfortunately, one alternate court reporter passed away before the start of this study, making it impossible to access transcripts for 5% of hearings.

From the outset, the two court reporters who staffed most hearings during the study period informed me that they had limited capacity to fulfill a large volume of transcript requests. Transcript creation is labor-intensive, requiring court reporters to manually convert shorthand stenographic notes into full written text, often outside of work hours.[8] Given this constraint, I decided to concentrate on a well-defined sub-population – the 62% of cases in which the highest charge was a

[8]In Wisconsin, court reporters are paid directly for any transcripts they produce at the statewide rate of $2.75 per page.

felony rather than a misdemeanor – about which I could make stronger inferences, rather than spread transcript requests more thinly across a broader population. I focus on felony cases for two main reasons. First, data showed that signature bonds were historically granted in 86% of misdemeanor cases and 70% of felony cases (County of Dane, 2018). This suggested an increased risk of ceiling effects in misdemeanor cases that could mechanically limit the PSA's effect and reduce the study's applicability to other settings with different base rates. Second, judges in Wisconsin set pretrial conditions for all defendants charged with felonies but only some defendants charged with misdemeanors – those unable to post the standard bail amounts outlined in the state's Uniform Bond Schedule and those charged with select serious misdemeanors. Focusing on felony cases thus homes in on a cleaner set of cases where judges always set pretrial conditions (see the Supplemental Material for suggestive evidence that the main results likely also hold in misdemeanor cases). I further narrowed the sampling frame to focus on cases with Black or white defendants (96% of cases) that were handled by the three judges who oversaw most Initial Appearances (99%). Finally, the sampling frame excluded cases filed under temporary case numbers (9% of observations) due to restrictions on data access and five cases missing administrative data.

I drew a stratified random sample of cases from this sampling frame (255 treated, 251 control). I stratified on defendant race, sex, judge, and time since implementation, oversampling female defendants, earlier cases, and cases handled by a judge who oversaw fewer hearings.[9] All quantitative results reported in this article are weighted to account for this stratified sampling, ensuring they are representative of felony cases in the RCT.

Covariate Balance

Although randomization creates comparable treatment and control groups on average, it does not guarantee balance in any individual sample. To assess balance along observed covariates, Table 1 compares the treated and control groups in the transcript sample, while Appendix Table A1 summarizes balance in the broader administrative data. The tables show that the samples are generally well-balanced along most observed covariates, although there is a moderate degree

[9] I took this approach to facilitate additional analyses beyond those presented in this article.

Table 1: **Transcript Sample Characteristics / Covariate Balance Table** Standardized mean differences are calculated following Austin (2009) for binary and continuous variables and Yang and Dalton (2012) for categorical variables. Values displayed are the average of pairwise comparisons between the treated and control samples.

	Control	**Treatment**	**SMD**
N	251	255	
Defendant race = White (%)	130 (51.8)	136 (53.3)	0.031
Defendant sex = Male (%)	189 (75.3)	190 (74.5)	0.018
Defendant age (mean (SD))	32.73 (11.49)	32.71 (11.47)	0.002
Judge (%)			0.055
Judge A	107 (42.6)	106 (41.6)	
Judge B	102 (40.6)	101 (39.6)	
Judge C	42 (16.7)	48 (18.8)	
Severity of highest felony charge (%)			0.027
A-C	18 (7.3)	20 (7.9)	
D-E	14 (5.7)	15 (5.9)	
F-G	54 (21.9)	54 (21.3)	
H-I	161 (65.2)	165 (65.0)	
# Cases (mean (SD))	1.18 (0.70)	1.22 (0.61)	0.068
# Unique charges (mean (SD))	2.60 (1.55)	2.70 (1.58)	0.064
# Violent charges (mean (SD))	0.65 (1.03)	0.56 (0.93)	0.086
# Bail jumping charges (mean (SD))	0.01 (0.09)	0.00 (0.06)	0.053
Failure to Appear Score (mean (SD))	2.93 (1.40)	3.09 (1.35)	0.112
New Criminal Activity Score (mean (SD))	3.58 (1.43)	3.70 (1.34)	0.087
New Violent Criminal Activity = True (%)	53 (21.1)	50 (19.6)	0.037
Recommendation (%)			0.097
Signature bond	139 (55.4)	138 (54.1)	
Low cash	24 (9.6)	25 (9.8)	
Moderate cash	8 (3.2)	13 (5.1)	
High cash	80 (31.9)	79 (31.0)	

of imbalance in PSA outputs. To address concerns about potential imbalance, the Supplemental Material presents alternate estimates for this paper's main quantitative findings with controls for important covariates, yielding results with similar significance levels and effect sizes.

7 Empirical Strategy

I organize the analyses into two parts. Part I examines how PSA availability shapes the relative power of prosecutors, defense attorneys, and judges to secure their desired pretrial conditions and

steer the degree of scrutiny given to different cases. Part II assesses how the PSA influences the relative salience of competing pretrial goals in courtroom deliberation. In both parts, I employ a form of methodological bricolage (Bonikowski and Nelson, 2022) within the overarching framework of the RCT, triangulating between qualitative and quantitative evidence to examine the algorithm's impact on decision-making.

Throughout, the unit of analysis is a hearing for an individual defendant, consistent with the scope of a single transcript. Eighty-seven percent of hearings address a single criminal case, while the remainder address two or more cases. This reflects the fact that when judges set pretrial conditions in a new case following an arrest, they also revisit pretrial conditions in other ongoing cases involving the same defendant. Thus, examining outcomes at the hearing level aligns the analyses with the real-world context of decision-making, in which courts make pretrial decisions across all of a defendant's pending cases at once.[10]

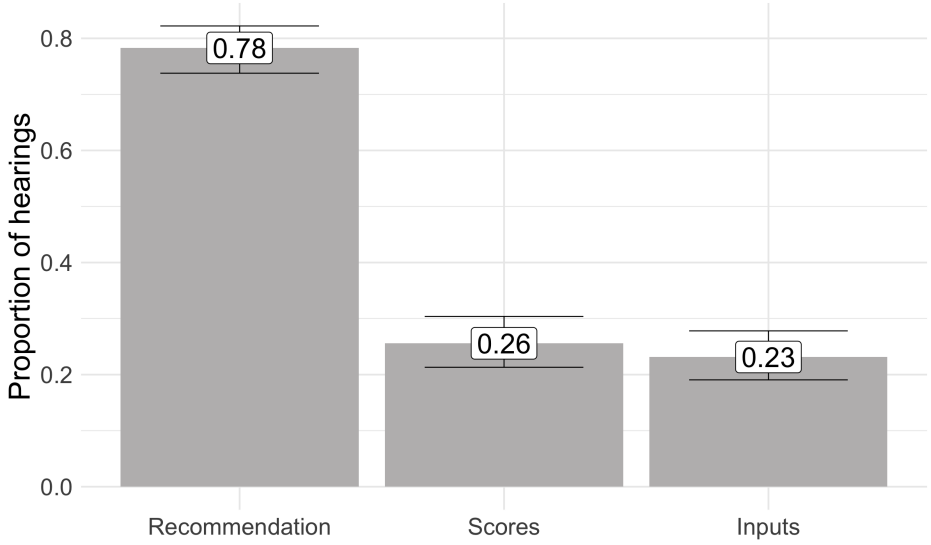
7.1 Part 1: Assessing how the PSA shapes decisions and shifts power

In Part I, I study whether and how the PSA influences decisions and reconfigures power at the micro level in courtroom decision-making. Following a minimal definition of power as the capacity to carry out one's will in a relationship even in the face of resistance (Weber, 1978), I analyze how the addition of an algorithmic perspective shapes courtroom actors' ability to influence i) whether defendants are required to post cash bail to be released before trial and ii) how much attention and scrutiny is directed at different cases.

I specifically focus on the PSA's recommendations for whether to impose cash bail, which link the PSA's risk scores to a concrete position on a case. These recommendations can selectively empower actors by offering new possibilities for alliances or distance between decision-makers as they support or undercut courtroom actors' own positions. As illustrated in Figure 4, PSA recommendations are the most-discussed component of PSA reports in Dane County.

[10] Another reasonable approach is to make an individual case the unit of analysis and examine only decision outcomes for the focal case used for randomization. The Supplemental Material shows that this approach produces results identical to those reported in the main text.

Figure 4: **How Often PSA Report Elements Are Discussed**. This figure shows the proportion of treated hearings (n = 255) in the transcript sample in which different elements of PSA reports were explicitly referenced.



I analyze how PSA recommendations shape whether prosecutors and defense attorneys succeed at securing their requested pretrial conditions, as well as how often judges depart from the joint requests of the prosecution and defense. To do so, I manually recorded the prosecution's and defense's requested pretrial conditions in each case (for additional coding details, see the Supplemental Material). I then assess how the PSA's impact varies as the PSA intervenes in the courtroom to create six combinations of positions on the appropriateness of cash bail. As Figure 5 illustrates, in treated hearings, PSA's recommendation may support a consensus between the prosecution and defense, disagree with a consensus, or side with one party over the other when the prosecution and defense disagree. By contrast, in control hearings, courtroom actors only observe the positions of the two parties. Researchers, however, can identify which of the six combinations of positions the PSA *would have* created had it been shared with the court. I use these latent combinations as comparison groups. To aid with interpretation, Table 2 clarifies how the effect of the PSA may be interpreted for each subgroup.

Figure 5: **Possible Relationships between PSA Recommendation, Prosecution Requests, and Defense Requests**

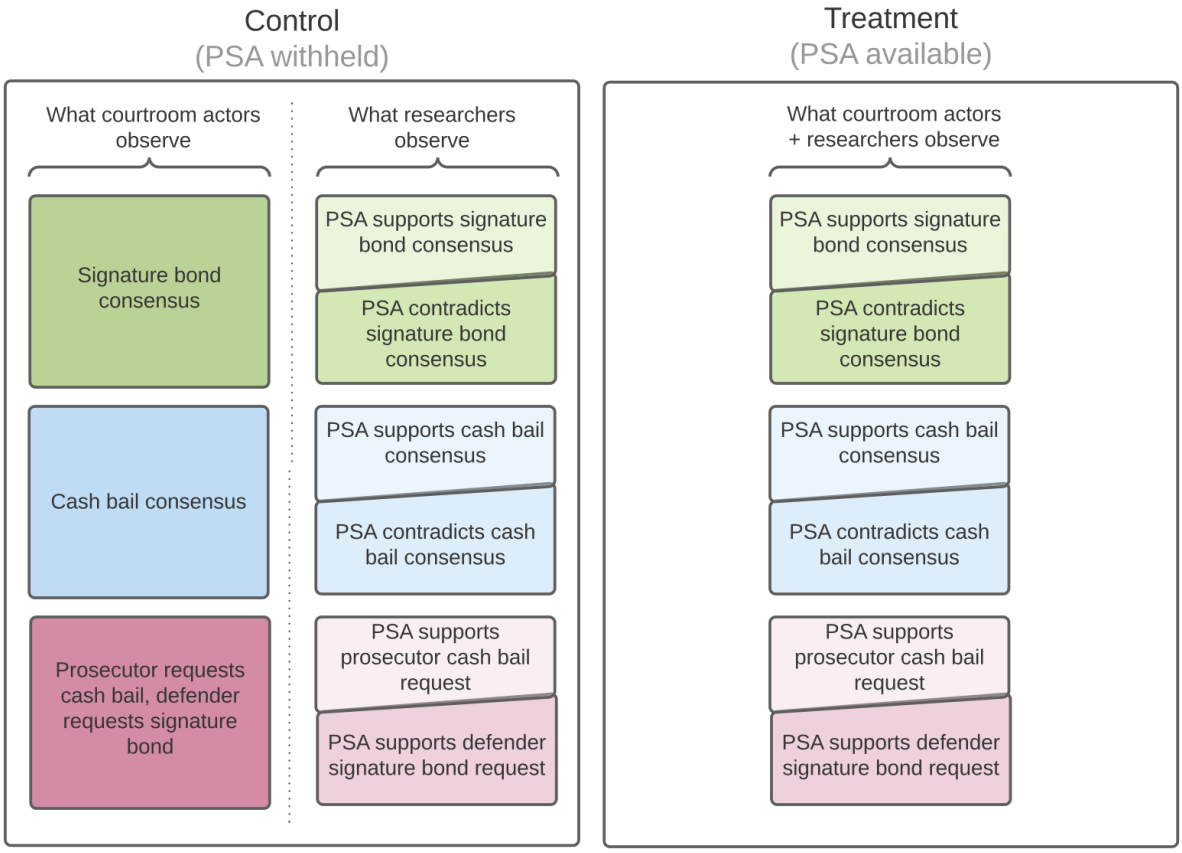


Table 2: **How to interpret the effect of providing vs withholding PSA reports under different combinations of prosecution requests, defense requests, and PSA recommendations**

Group	**Request/Recommendation**			**What effect of providing vs withholding the PSA represents**
	Prosecution	**Defense**	**PSA**	Difference in whether the judge in hearing *i* would a) set cash bail or b) direct added attention to the case if...
PSA supports signature bond consensus	Signature bond	Signature bond	Signature bond.	a joint request for a signature bond is advanced with the support of a PSA signature bond recommendation vs if it is advanced without PSA support
PSA contradicts signature bond consensus	Signature bond	Signature bond	Cash bail	a joint request for a signature bond is contradicted by a PSA cash bail recommendation vs if the request goes uncontradicted
PSA supports cash bail consensus	Cash bail	Cash bail	Cash bail	a joint request for cash bail is advanced with the support of a PSA cash bail recommendation vs if it is advanced without PSA support
PSA contradicts cash bail consensus	Cash bail	Cash bail	Signature bond	a joint request for cash bail is contradicted by a PSA signature recommendation vs if the request goes uncontradicted
PSA supports prosecution cash bail request	Cash bail	Signature bond	Cash bail	a prosecution request for cash bail that the defense attorney opposes is advanced with the support of a PSA cash bail recommendation vs if it is advanced without PSA support
PSA supports defense signature bond request	Cash bail	Signature bond	Signature bond	a defense request for signature bond that the prosecution opposes is advanced with the support of a PSA signature bond recommendation vs if it is advanced without PSA support

This analysis relies on a key assumption: seeing the PSA recommendation does not change what the prosecution and defense request. This assumption is plausible given the way the PSA was implemented in Dane County. Court staff normally printed PSA reports right before the afternoon's hearings, leaving courtroom actors little time to review them in advance (Greiner et al., 2020). Prosecutors and defense attorneys, meanwhile, typically prepared their positions ahead of time without the input of the PSA.[11] This assumption allows courtroom actors to update how they engage with each other and *frame* their requests given the PSA, but requires that they do not alter *what* they request. I present descriptive statistics consistent with this assumption in the next section. The Supplemental Material additionally discusses how violations of this assumption would bias the results and presents additional robustness checks.

7.1.A Part 1A: Measuring Effect on Cash Bail

Data on whether cash bail was imposed in each case come from Wisconsin Circuit Court Access, the state's administrative court records database. In hearings addressing multiple cases with the same defendant, I recorded the most restrictive condition imposed across the cases as the hearing's outcome. That is, if a judge orders cash bail in one case but a signature bond in another, I record that hearing as having resulted in cash bail.[12]

I estimate the average treatment effect (ATE) of PSA access on setting cash bail by calculating differences-in-proportions between treated and control cases. I further estimate conditional average treatment effects (CATEs) for my main analytic subgroups using differences-in-proportions with weights to account for stratified random sampling. As the results involve a number of comparisons, I applied the Benjamini-Hochberg procedure to adjust reported p-values and control the expected rate of false discoveries.

[11] In the district attorney's office, Assistant District Attorneys were tasked with determining prosecution requests ahead of the day's hearings. The prosecutor in the courtroom usually read out these prepared requests.

[12] In practice, judges are nearly always consistent across cases. As the Supplemental Materials show, measuring outcomes based just on the focal case used in randomization yields identical results to measuring outcomes at the hearing level. The quantitative results are not sensitive to this decision.

7.1.B Part 1B: Measuring Effect on Attention and Scrutiny Across Cases

I pair the preceding analysis of the PSA's influence on what courts decide with an examination of how those outcomes were achieved, focusing on the attention and scrutiny that different cases receive. When and how does the PSA help courtroom actors streamline or shut down debate? Direct added scrutiny toward defendants? Achieve their desired outcomes with less effort?

In this subsection, I quantitatively summarize the PSA's effect on the amount of deliberation across hearings and use qualitative coding to unpack how courtroom actors leverage the PSA to influence deliberation. In the quantitative analysis, I measure the amount of deliberation and debate in hearings in a few ways: number of words spoken (across a hearing and in judges' rationales for their decisions) and linguistic markers of cognitive processing, as measured using Linguistic Inquiry Word Count (LIWC) (Pennebaker et al., 2015). LIWC is a widely used dictionary-based method for detecting words and phrases associated with different linguistic styles, psychological states, and thought processes (Tausczik and Pennebaker, 2010). Cognitive processing refers to the extent to which a text's language reflects complex analytical thinking. As with my analysis of impacts on cash bail, I estimate the CATEs of the PSA on these deliberative outcomes across my main analytic subgroups using differences-in-means, with weighting to account for stratified sampling and p-values adjusted using the Benjamini-Hochberg procedure to account for multiple testing.

To examine how the PSA affected courtroom deliberations, I employed a multi-phase qualitative coding process (Charmaz, 2006). Initially, I used open coding to identify all instances where actors invoked the PSA and capture the various ways actors weaved the tool into their statements. I then categorized these instances into broader themes, paying close attention to patterns in which actors leveraged the PSA, the circumstances surrounding its use, and its impact on the deliberative dynamics.

7.2 Part 2: Examining the PSA's Influence on the Salience of Competing Imperatives

Part 2 investigates how PSA access shapes the relative salience of different pretrial goals in courtroom discussions and compares the discussion topics most associated with setting cash bail in the treated and control groups. In particular, in Subsection 2A, I examine whether access to PSA reports alters how often burdens, non-appearance, and public safety are discussed. To measure discussions of these pretrial concerns, I hand-coded instances in transcripts where courtroom actors appealed to them.[13] For burdens, for example, I recorded any time an actor discussed the appropriateness of cash bail given an arrested individual's financial means or the potential impact of pretrial detention on the individual's ability to care for loved ones, maintain a job, or build a strong defense. I estimate the CATE of the PSA on discussions of each pretrial concerns by PSA recommendation, calculating differences-in-proportions with the same weighting and multiple testing adjustments as in Part 1.

In subsection 2B, I then evaluate whether the discussion topics that PSA access makes more salient are associated with judges' final decisions of whether to require cash bail. My approach, however, was not to assume that varying discussions of pretrial goals necessarily link to different decisions. Instead, I take a more open-ended inductive approach that also allows the possibility for other topics to inform decisions.

In particular, I convert the text of judges' decision rationales into sequences of one, two, and three consecutive words (unigrams, bigrams, and trigrams). This produces a large vocabulary of words and phrases spoken across hearings. I then use penalized logistic regression, specifically LASSO, to identify the words and short phrases that are most predictive of setting cash bail in treated and control hearings. In this setup, my binary outcome variable is whether cash bail was set and my predictors are unigrams, bigrams, and trigrams. LASSO works by adding a penalty term to a standard logistic regression, which shrinks the coefficients of less informative predictors to exactly

[13] I counted defendants' words under "Defense" to account for the rare occasions when defendants contributed information without their lawyers as an intermediary. Additional coding details are in the Supplemental Material.

zero (for additional details on text pre-processing, model tuning, and model performance see the Supplemental Material). The result is a model where only the most predictive words or phrases are retained. I run separate models on treated and control hearings, allowing me to compare the phrases most strongly associated with cash bail decisions with and without the PSA. I then closely read transcripts where those predictive phrases appear to understand the contexts in which they are spoken and their link to final decisions.

8 Descriptive Statistics

Figure 6 displays how often the prosecution and defense request cash bail rather than a signature bond with and without the PSA. Prosecutors request cash bail in around half of cases, while defense attorneys request cash bail in only around 1 in 10 cases. Consistent with the PSA not influencing what courtroom actors request, there are no statistically significant differences in requests between treated and control hearings.

Figure 6: **Cash bail request rates by actor.** (n_T = 255; n_C = 251)

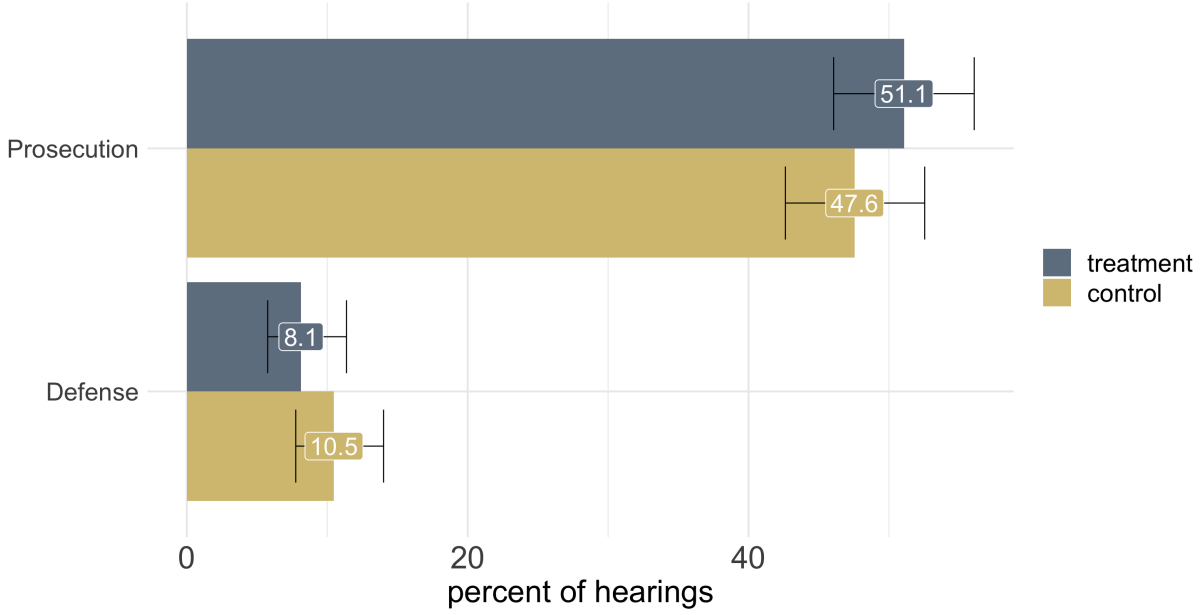
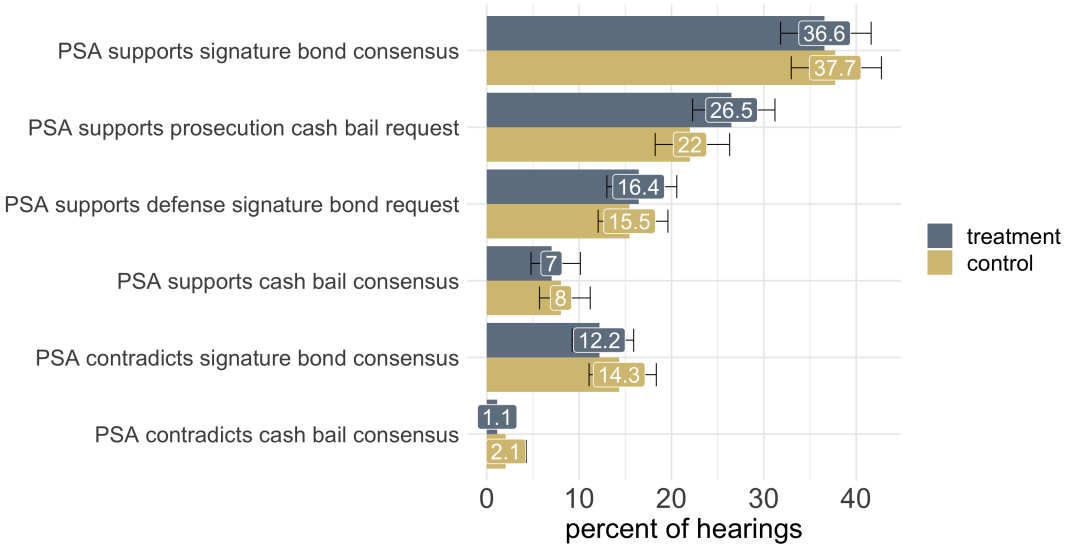


Figure 7 displays the prevalence of different combinations of relationships between the PSA recommendation, prosecution request, and defense request in a case. On the whole, PSA recommendations align with prosecutors' requests more often than they do with defense attorney's requests. When there is disagreement between the prosecution and defense, the PSA sides with the prosecution more often than the defense.

The PSA recommendation additionally introduces disagreement in around 12-14% of cases by recommending cash bail despite joint requests for a signature bond. More rarely, the PSA adds disagreement by recommending a signature bond while the prosecution and defense request cash bail. Given the rarity of this analytic group, I will not formally examine the PSA's impact in these cases in this paper.

Consistent with the PSA not influencing what courtroom actors request, there are no statistically significant differences in requests between treated and control hearings.

Figure 7: **Prevalence of comparison groups.** (n_T = 255; n_C = 251)



9 Results

9.1 Part 1: How the PSA shapes decisions and shifts power

9.1.A 1A: PSA's effect on cash bail

Figure 8 presents an overview of the impact of PSA report provision on the use of cash bail in the full administrative data sample.[14] The left panel displays Average Treatment Effects with 95% confidence intervals, while the right panel displays rates of cash bail in the treated and control groups. While PSA provision does not have a statistically significant effect averaging across all hearings (top panel, p = 0.15), it has heterogeneous effects by PSA recommendation (bottom panel). Among cases where the PSA recommends cash bail (n_T = 314, n_C = 306), PSA access increases the use of cash bail by an estimated 12 percentage points (p = 0.007). Meanwhile, among cases where the PSA recommends a signature bond (n_T = 381, n_C = 378), there is insufficient evidence to

[14]These results align with those in an earlier paper based on the same RCT, Imai et al. (2023), which also fails to detect a significant overall effect of PSA reports on cash bail. This paper, which focuses on different overarching research questions, diverges from Imai et al. (2023) by further testing for effect heterogeneity by PSA recommendation. The Supplemental Material shows that this effect heterogeneity replicates in Imai et al. (2023)'s data despite some differences in analytic choices between the two papers (mostly notably, I exclude misdemeanor cases while Imai et al. (2023) exclude PSAs triggered by re-arrests), which I detail further in the Supplemental Material.

conclude that PSA access has an effect ($p = 0.43$). Although the 95% confidence intervals are wide, the results suggest that it is unlikely that the true effect of PSA signature bond recommendations would be negative and of a sufficient magnitude to counterbalance the positive effect of PSA cash bail recommendations.

Figure 8: **Effect of PSA on Use of Cash Bail, Overall.** Based on administrative data; $n_T = 695$ and $n_C = 684$.

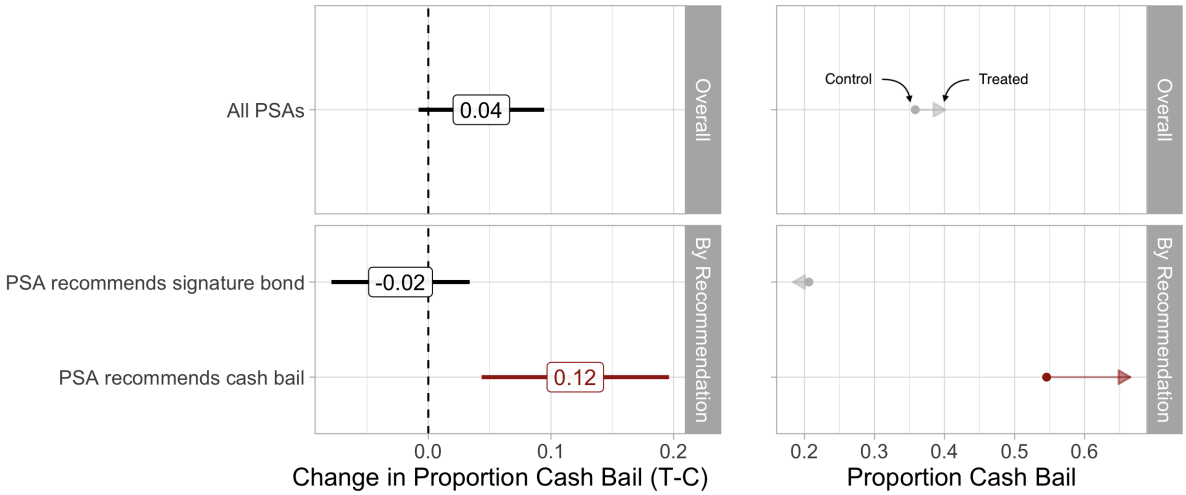
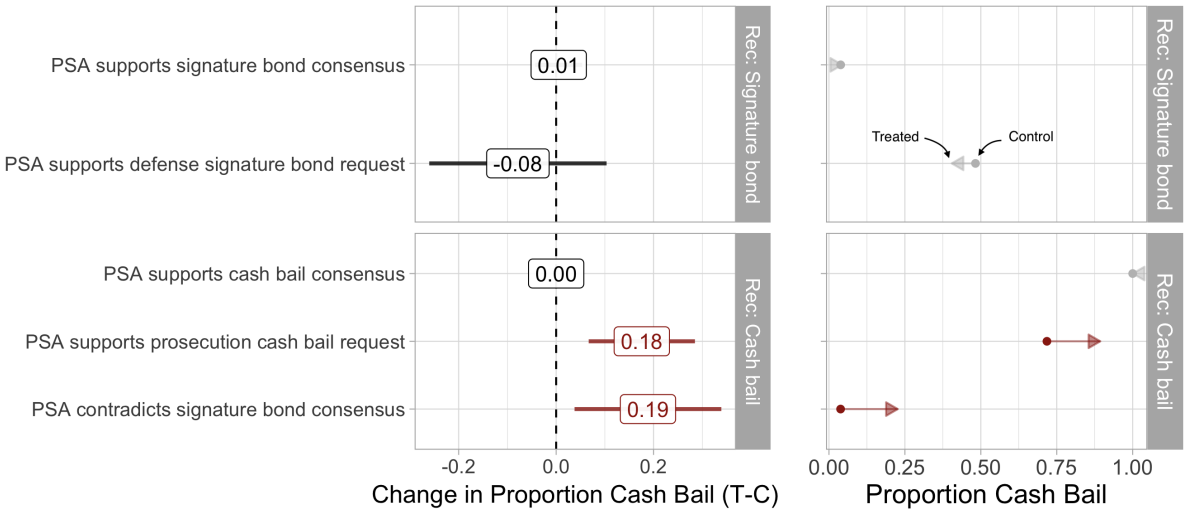


Figure 9: **Effect of PSA on Use of Cash Bail, by Relationship with Prosecution and Defense Requests.** Based on transcript data; $n_T = 255$ and $n_C = 251$. See Table 3 for subgroup sample sizes.



To investigate whether the PSA reconfigures actors' ability to align decisions with their stated positions, Figure 9 displays the estimated effects of PSA provision on the use of cash bail in analytic subgroups defined by how the PSA recommendation compares to prosecution and defense requests. Table 3 presents p-values and subgroup sample sizes.

When the prosecution and defense disagree on the appropriateness of cash bail, PSA recommendations backing prosecution requests for cash bail increase the use of cash bail by 18 percentage points, from 71.8% to 89.4% of cases. By contrast, there is insufficient evidence to conclude that PSA recommendations supporting defense requests for a signature bond have a comparable offsetting effect in the opposite direction. While I may be underpowered to detect a small negative effect for this group, the results suggest it is unlikely for the true effect to be of a comparable magnitude to mirror that of a PSA recommendation siding with a prosecution request for cash bail. These results suggest that PSA reports bolster the positions of prosecutors more than those of defense attorneys when there is disagreement.

Meanwhile, when PSA recommendations align with a consensus between the prosecution and defense, PSA report access has minimal impact. Regardless of PSA access, all cases where the PSA would agree with a cash bail consensus end with cash bail being set. In cases where the PSA

agrees with a signature bond consensus, PSA provision similarly has little impact. This is primarily because commissioners only required cash bail in 4% of such cases in the control group, making substantial reductions in the use of cash bail in the treated group mechanically impossible.

However, when PSA recommendations introduce fresh disagreement and contradict a signature bond consensus, they increase cash bail decisions by 19 percentage points, from 3.8% to 22.7%. This suggests that PSA recommendations facilitate commissioners diverging from the bounds that the prosecution and defense laid out, departing from their normal practice of deferring to their requests. Together, these results indicate that PSA cash bail recommendations are more influential than signature bond recommendations and that PSAs may be especially empowering for prosecutors and judges in hearings where actors disagree on the right course of action.

9.1.B 1B: How PSA access influences attention and scrutiny across cases

This section combines qualitative and quantitative analyses of how access to PSA reports intervenes in the courtroom to shape the amount of attention and scrutiny defendants receive. Complementing the preceding analysis of decision outcomes, it compares the same analytic subgroups, summarizing how actors mobilize the PSA in different scenarios and assessing the degree to which the PSA steers attention in directions favorable to different actors.

Table 3 displays the tool's impact on a range of deliberation-related outcomes in the main analytic subgroups.

Table 3: **Effect of PSA Access on Cash Bail and Deliberation Outcomes:** presents Average Treatment Effect estimates by analytic group. Standard errors are in brackets. Cognitive Processing is measured in terms of the number of linguistic markers consistent with cognitive processing according to LIWC (Pennebaker et al., 2015). p-values adjusted for multiple testing using the Benjamini-Hochberg procedure.

Analytic group	**Cash bail**	**Words Spoken**		**Cog. Processing**	**N**	
		Total	Judge Rationale		Control	Treated
PSA supports signature bond consensus	0.01 (0.024)	⇑ 99* (36.7)	⇑ 35** (11.8)	⇑ 15.4** (5.3)	95	92
PSA supports defense signature bond request	-0.08 (0.09)	-41 (67.5)	-34 (29.7)	-11.7 (10.8)	38	43
PSA supports cash bail consensus	0 (0)	⇓ -211* (91.4)	4 (28.2)	⇓ -24.1* (10.7)	20	19
PSA supports prosecution cash bail request	⇑ 0.18** (0.055)	⇓ -137* (64.9)	-4 (17.1)	⇓ -24.5** (8.5)	57	65
PSA contradicts signature bond consensus	⇑ 0.19* (0.074)	⇑ 251** (71.0)	⇑ 99*** (20.0)	⇑ 31.6** (10.4)	35	32

* p < .05, ** p < .01, *** p < .001

When there is disagreement and the PSA recommendation sides with a prosecution cash bail request, PSA access streamlines decision-making, leading to less debate and evidence of cognitive processing. Considered together with the finding that PSA access increases the chance of cash bail in these cases, the results suggest that PSA access not only helps prosecutors secure their desired conditions more often but also helps them do so more easily. By contrast, when the PSA sides with a defense signature bond request, I find no statistically significant effects on deliberation outcomes. This suggests that PSAs supporting the defense may have minimal impact on both attention and decision outcomes.

The contrasting effectiveness of PSAs for the prosecution and defense when the two sides disagree is made starker by differences in the amount of effort the actors invest in incorporating the PSA into their arguments. Qualitative coding of rhetorical strategies reveals that defense attorneys, tend to engage with PSA reports more deeply than do prosecutors. When prosecutors appeal to the PSA, the report typically assumes a minimal role in their arguments. If the PSA recommendation aligns with their ask, prosecutors often merely note that their request is "consistent with," "follows," or "agrees with" the PSA. If it does not align, prosecutors typically acknowledge the departure and

explain it in a brief, perfunctory manner.

Meanwhile, defense attorneys tend to incorporate the PSA more centrally into their arguments. They draw attention PSA recommendations and other elements of PSA reports that place their client in a favorable light, such as their client's court appearance record (e.g., "This assessment did not note any failures to appear within the last two years or failure to appear older than two years") or favorable scores (e.g., "He scores a one on the failure to appear score"). Importantly, when the PSA aligns with their request for a signature bond, defense attorneys explicitly encourage judges to trust the tool, assuring them "you can rely on the PSA," opining that the PSA "correctly weighs" the defendant's risk, stating that the PSA's recommendation is "appropriate," and explicitly asking them to follow the signature bond recommendation. The PSA's relative ineffectiveness when it sides with the prosecution rather than the defense is thus particularly striking as it takes place *despite* defense attorneys' substantial efforts to harness the PSA as a rhetorical resource.

When the PSA supports a signature bond consensus, it leads to small increases in overall hearing length, judge rationales, and cognitive processing. Qualitative analyses indicate that these increases are concentrated in cases where judges still set cash bail. With the PSA, judges in such cases typically elaborate more on the reasons why their decision departs from not only the prosecution and defense but also the PSA.

When the PSA supports a cash bail consensus, it has no impact on decisions – all result in cash bail regardless of PSA access – but streamlines arriving at that result. With the PSA, hearings are shorter and entail less cognitive processing. Judges cut debate off earlier, sometimes employing the PSA directly to signal they have heard enough through transitions like "Okay, well, I will note that there is a Public Safety Assessment that also recommends cash bail" or "Well, I agree with the Public Safety Assessment."

Finally, when the PSA contradicts a consensus signature bond request, it substantially increases the deliberation and scrutiny directed at cases. Access to the report leads to a second look at cases that would otherwise be processed as uncontroversial signature bonds. Without the PSA, judges commonly follow the prosecution and defense, regularly drawing attention to "joint requests" as

part of their rationales (e.g. "I'm willing to go along with the joint request for a signature bond"). Indeed, judges explicitly acknowledge their typical practice of deferring to these requests. As one stated, "I'm not one who typically jumps the joint recommendations by the parties who have more information than I have."

PSA cash bail recommendations, however, prompt further scrutiny of these cases. To understand the shift, it is helpful to first illustrate how cases in the control group typically proceed. The following is a representative extract of the modal control hearing:

> **COMMISSIONER:** Great. What's the State's position on bail?
>
> **PROSECUTOR** Signature bond with the condition of no controlled substances or drug paraphernalia without a valid prescription.
>
> **DEFENSE ATTORNEY:** No objection.
>
> **THE COURT**: So ordered....

Unlike these brief, straightforward hearings, treated group hearings are longer and exhibit substantially more cognitive processing. Close reading of transcripts from this subgroup reveal two major shifts in courtroom debate that are consistent with the PSA empowering judges while undermining the requests of the prosecution and defense. First, in hearings with the PSA, judges began to quiz prosecutors on the discrepancy between their requests for a signature bond and the PSA cash bail recommendation, pairing questions like "have you reviewed the Public Safety Assessment?" with "did you want to weigh in at all for the reasons for a signature bond?" These interjections signaled to the prosecution and defense that cash bail might be on the table. Prosecutors were then moved to defend their request, while defense attorneys provided fuller arguments for a signature bond.

Second, whereas judges regularly cast control group decisions as so self-evident as to not require any explanation, judges described comparable treated group decisions as more fraught. More than tripling the length of their rationales, judges described more of their decisions as "a close call" or "right on the edge" between a signature bond and cash bail. When they set cash bail, an action they more often took in response to PSA cash bail recommendations, judges frequently cited the PSA

in lengthy explanations of why they thought the prosecution and defense were underestimating the defendant's "true" riskiness. For instance, in one hearing, the judge introduces new details from the defendant's history and uses the PSA to pivot to his read of the case:

> I'm concerned about a signature bond in this case. There is a Public Safety Assessment that has a violent criminal activity flag flapping to the extent that it can on a piece of paper. It has the absolute highest level for new criminal activity risk. It's the middle of the road for failure to appear risk and also recommends cash bail. As I understand it Mr. [Defendant] is on parole.... He's now alleged to be on felony bond and engaging in a new battery charge. If my quick look of [the state court records database] is correct... with his release to extended supervision he was arrested and charged with possession of a firearm as a convicted felon, possession of narcotic drugs... Ultimately, all of this convinces me that the evidence-based Public Safety Report is the way to go here and I'm going to set cash bail.

Judges thus mobilize PSA cash bail recommendations to more comfortably depart from joint requests, placing affected defendants under greater suspicion than they would have been absent the PSA.

In sum, these results offer further evidence that the PSA may help prosecutors more than it helps defense attorneys, enabling prosecutors to more easily secure their desired conditions. Defense attorneys, despite investing more energy in highlighting PSA details favorable to their clients, fail to secure greater influence over deliberations and final outcomes. Judges, meanwhile, are empowered by the PSA as it offers a rhetorical resource to exert greater control over courtroom debate and justify harsher pretrial conditions than those requested by the prosecution and defense.

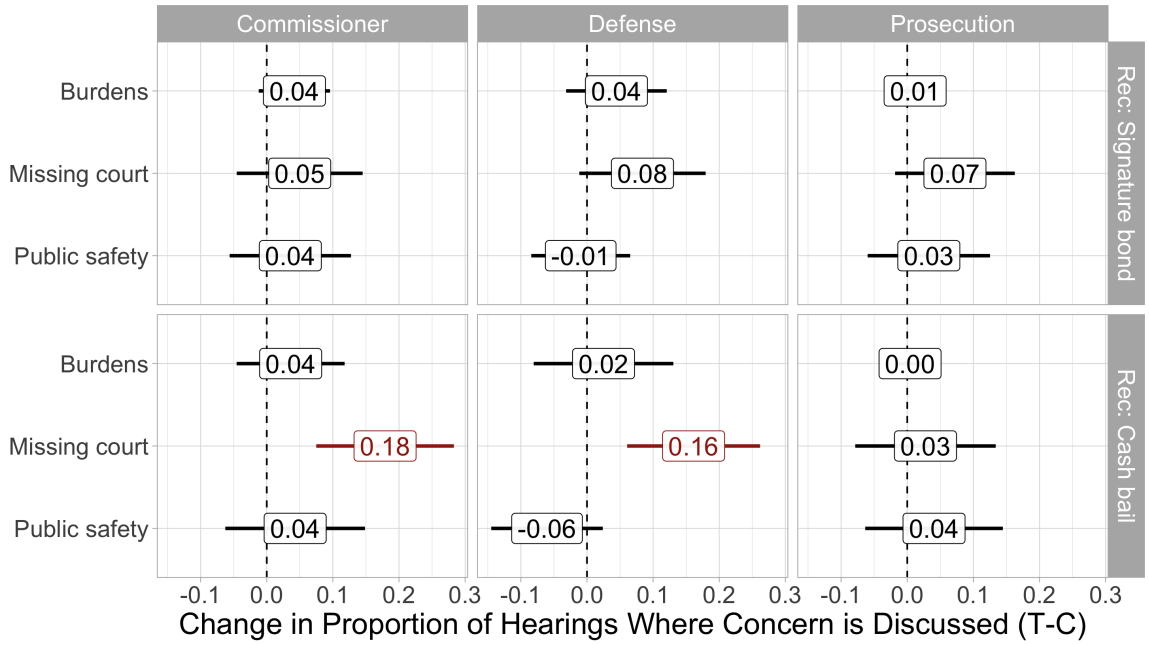
9.2 How the PSA influences the salience of competing imperatives

9.2.A Effect on discussions of different pretrial concerns

In this section, I explore a reason why risk assessments may tend to empower prosecutors and judges more than defense attorneys: risk assessments steer the terms of debate toward the outcomes that the tools predict, funneling goals in ways that disfavor the defense.

Figure 10 plots the PSA's effect on how often different courtroom actors discuss non-appearance, public safety, and burdens on defendants by PSA recommendation.

Figure 10: **Discussion of Pretrial Concerns by Actor and PSA Recommendation.** For hearings with a signature bond recommendation, n_T = 138 and n_C = 139. For hearings with a cash bail recommendation, n_T = 117 and n_C = 112.



Consistent with earlier results, the top row of the figure indicates that signature bond recommendations do not substantially alter how often different pretrial concerns are discussed. Additionally, the PSA, regardless of recommendation, prompts little change in the concerns that prosecutors emphasize (third column). This latter finding is consistent with the risk assessment presenting little challenge to prosecutors' normal practices because the tool's design generally aligns with a prosecutorial perspective.

By contrast, PSA cash bail recommendations lead the defense and commissioners to discuss non-appearance risk more frequently ($p = 0.016$ and $p = 0.015$ respectively). Given PSA reports, defense attorneys more often advocate for signature bonds by doubling down on their clients' ties to the community, arguing for instance, "he's lived here his whole life," "He has children ages 9 and 3, parents, siblings...all his family lives in this area," or "He's been working at [local gaming establishment] for about 5 years and his ties to the community therefore are strong." When the PSA report shows a missed court appearance, defense attorneys often explain circumstances behind the lapse that might reduce their client's perceived moral culpability or suggest that history is unlikely to be repeated. They highlight personal emergencies ("it's my understanding that the bail jumping comes from [my client] having to appear at his father's funeral."), medical circumstances ("...he had trouble obtaining his prescription for psychotropic medications.... He wasn't on his medication, and you know, things get discombobulated... so that's why he missed that prior court appearance. I don't think it was anything that he did on purpose."), or the defendant not being informed of the court date ("he didn't receive the mail notifications").

Defense attorneys also directly leverage low Failure to Appear scores and favorable court appearance histories summarized in PSA reports to argue that their client is a good bet for showing up for court and should therefore be granted a signature bond. For instance, one public defender contested a prosecutor's request for cash bail in the following manner:

> We oppose cash bail, your Honor. [The defendant] has lived in this area his entire life. I'm not hearing about recent missed court dates. In fact, the Public Safety Assessment reports that there haven't been any missed court dates in the past two years.... Additionally, concerns about compliance with conditions of bail aren't a basis for cash bail. That's to assure appearance in court."

Here the defense uses PSA information to emphasize the defendant's solid court appearance history and, notably, reminds the court that the statutory legal basis for setting cash bail in Wisconsin is to assure court appearance. This latter move may help explain why the PSA increases discussions of appearance risk but not public safety risk despite offering predictive scores for both. It suggests

that enforcing a distinction between non-appearance and public safety risk, two often-conflated concepts (Gouldin, 2016), may better align decision-making with the legal basis for cash bail.

Yet it remains an open question how increasing the salience of non-appearance shapes decisions. Do appeals like the one detailed previously work? While defense attorneys tend to draw attention to favorable court appearance information, centering non-appearance concerns may, on net, inadvertently cast defendants in a harsher light. This may be especially the case as Figure 10 indicates the PSA prompts no offsetting increase in discussions of the burdens that pretrial conditions could place on defendants. The next section assesses the possibility that tipping the balance of objectives influences decision outcomes.

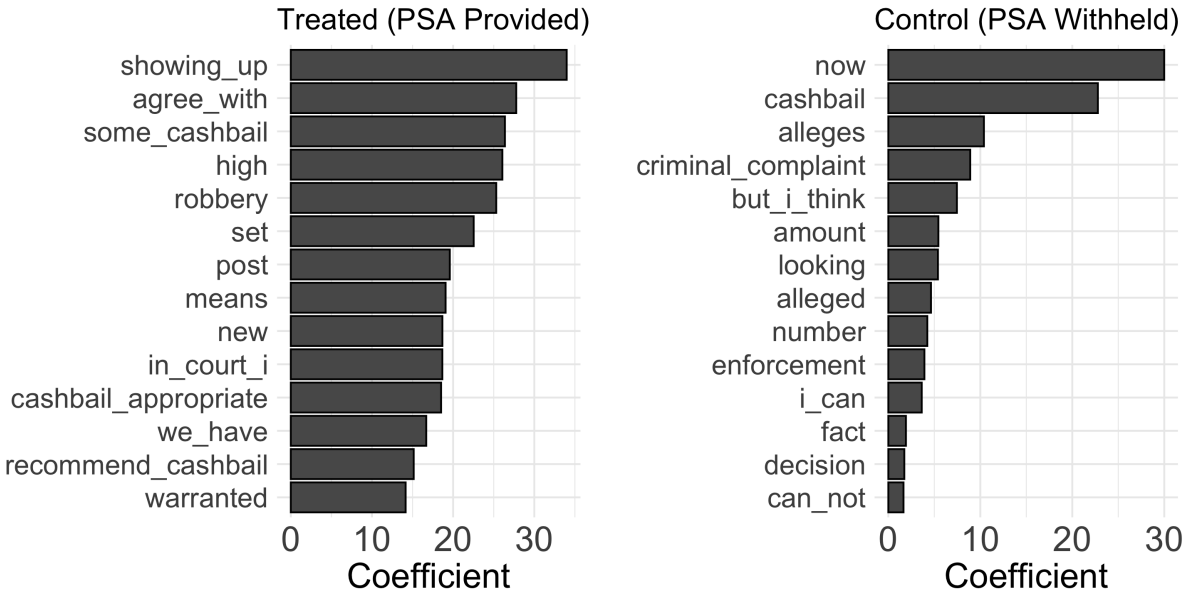
9.2.B Comparing discussion topics linked to cash bail decisions in treated vs control hearings

This section examines what elements of judges' statements pre-decision are most predictive with cash bail being set, comparing the words and phrases that predict cash bail decisions with and without the PSA. Figure 11 shows the most predictive phrases in each group. Positive coefficients indicate that the phrase is associated with a greater chance of the hearing ending in a cash bail decision. To interpret this figure, I discuss below how the predictive phrases appear in broader conversational context and note that some phrases take on multiple meanings.

Unsurprisingly, the figure shows that phrases directly related to setting cash bail – "cash bail," "some cash bail," and "set" – strongly predict cash bail decisions in both groups. Discussions of appropriate bail dollar amounts, such as amount ("I am not going to set an excessive amount"), number ("the number mentioned by the State is reasonable"), and high ("it does not need to be as high as the State has asked for"), also predict cash bail.

A few key differences between the treated and control results suggest how the PSA shapes the considerations that contribute to cash bail decisions. A first difference is in the appearance of PSA-related phrases as strong predictors in treated hearings. Notably, "recommend cash bail," which is nearly always preceded by "the PSA" is associated with a higher chance of cash bail, again indicating the potency of cash bail recommendations (by contrast, the lasso model did not select

Figure 11: **Top phrases predictive of cash bail decision** (n_T = 255; n_C = 251) Coefficients come from penalized logistic regression models run separately on treated and control hearings (LASSO with penalties of λ_t = 0.0344 and λ_c = 0.0599 in the treated and control groups respectively). The models attain an average predictive accuracy of 77% and AUC of 0.90 on the test set. See Online Supplement for additional details.



"recommends signature bond" as a meaningful predictor of decisions). Judges add statements like "well, I will note that there is a PSA that also recommends cash bail" and 'I also put some weight on the PSA that recommends cash bail" when enumerating their reasons for setting cash bail. The highly predictive word "high" also appears in conjunction with PSA-related statements, typically as part of comments on a defendant's elevated risk. For instance, judges note "the PSA indicates high risk" or "He is on the high end of the scale, five out of six for failure to appear... suggesting he is a risk to not appear and that cash bail is appropriate."

A second difference suggests that the concerns animating decisions differ in treated and control hearings: discussions of criminal activity and public safety are more central to control group decisions, while discussions of court appearance concerns are more central to treated group decisions. In control hearings, more top phrases reflect a focus on the details of the allegations: "alleges", "criminal complaint", and "alleged." Other phrases commonly appear in statements suggesting

that a defendant is likely to commit further offenses given their criminal history or other pending cases, including "now" ("and now we have this allegation of pretty serious new violent assaultive offenses") and "number" ("cash bail would make sense given the number of charges and number of cases").

Two additional words predictive of cash bail in the control group are associated with arguments that blend concerns about public safety and court appearance. The first, "looking," largely appears in talk of the severity of the charges and punishments a defendant is facing, such as when judges state, "he is looking at the mandatory minimum three years in prison... I think there is a lot of incentive not to show up for that" or "I am certainly not going to take any chances given the nature of what he is looking at here." These discussions are premised on the idea that the heavy penalties associated with more serious crimes create more incentive for flight. The second word, "enforcement" consistently features in judges' comments on defendants' perceived compliance with law enforcement officers. Local judges position behavior around law enforcement as indicative of one's general respect for the law, which they believe shapes a defendant's risk of criminal activity and missing court. For example, a judge in one case observed, "Here we have somebody who essentially was not complying with law enforcement, and when people do that I start to worry about whether they are going to show up in court and follow other conditions of bail," while another remarked that a defendant's behavior with law enforcement "that's alleged in this Complaint strikes at the very heart of ... a civilized society, but it also strikes at the heart of whether he's willing to comply with these bail conditions and his likelihood of appearance...." In both cases, these words reflect the mixing of concerns about crime and missed court.

By contrast, in treated hearings, we observe phrases strongly predictive of cash bail that are more tightly focused on concerns about court appearance rather than public safety or a mixture of the two. "Showing up" and "in court" are among the top predictive phrases, with judges observing that a defendant has a "pretty poor track record of showing up in court" or stating "you are somebody that I ought to worry about...showing up in court." These concerns, aligned with the formal legal basis for cash bail in Wisconsin, are in turn linked to the perceived need for monetary conditions ("so it

does seem like a case for some cash bail to assure Mr. [Defendant]'s appearance in court"). While the PSA does not completely upend concerns about public safety – "robbery" appears among the top phrases, reflecting discussions of serial robberies prevalent at the time – it appears to strengthen the association between court appearance concerns and cash bail decisions. This is significant, as the previous section showed that the PSA led defense attorneys and commissioners to talk more often about non-appearance risk. This result thus suggests that the increased salience of this concern may, on net, work against defendants seeking signature bonds; centering concerns about non-appearance risk strongly predicts cash bail.

Other contrasts offer suggestive evidence that the PSA facilitates cash bail decisions by altering interactional possibilities for achieving the result. In treated hearings, "agree with" highly predicts cash bail. In such hearings, judges observe how their cash bail decision aligns with the positions of the prosecution or the PSA (e.g. "I agree with the State that at this point the cash bail is appropriate"), or acknowledge that they have taken some elements of the defense's arguments into account (e.g. "I agree with Attorney [Public Defender], the incentive for flight isn't great in this particular case. I'm going to set it relatively low."). Meanwhile, in the control group, "but I think" is highly predictive. This phrase is generally used as a transitional pivot, demarcating judges' own read of the case and asserting their agency. For instance, a judge in one case stated, "I will take into consideration the public defender eligibility and his ties to the area, but I think he needs some money to understand the importance of showing up." These findings suggest that in treated hearings, the PSA allows judges to align with others rhetorically when imposing cash bail, while in control hearings, judges must draw greater attention to their own judgment.

A final important result in Figure 11 is that "warranted" and "cash bail [is] appropriate" predict cash bail, but only in treated hearings. Close reading reveals that "warranted" nearly always appears as part of statements that "cash bail is warranted" or "the state's request [for cash] is warranted." It typically follows or precedes summaries of information casting defendants in a negative light, and sometimes features the PSA ("I think some cash bail is warranted... The PSA has him as a five out of six on the failure to appear scale, six out of six on the criminal activity scale.").

Similarly, judges suggest that defendants deserve cash bail by asserting its appropriateness ("given the allegations here, it is pretty clear that cash bail is appropriate"). Notably, judges often invoked the PSA to support these assertions For instance, one judge remarked, "I agree with the state's ... characterization of the PSA report. I also agree with its position that some cash bail is appropriate here." Overall, these statements suggest that the PSA not only shapes decisions but also enables judges to more readily position cash bail as just deserts.

10 Discussion

This article demonstrated how predictive algorithms can produce an array of interconnected asymmetric impacts on the ground through the case of a popular pretrial risk assessment algorithm. Combining qualitative and quantitative data from a randomized controlled trial, I examined how the risk assessment shaped courtroom deliberation, the dynamics among courtroom actors, and decisions about cash bail. The results showed that algorithmic information indicating a person is high risk is more influential: recommendations for cash bail are more central to deliberations and steer the harshness of decisions more than recommendations for signature bonds. This pattern takes root as the PSA unevenly empowers decision-making actors, helping prosecutors more than defense attorneys secure their desired outcomes. The PSA further erodes judges' reliance on the two sides to define the space of plausible decisions, offering an additional voice that enables judges to break from the norm of staying within the confines of prosecution and defense requests. These dynamics are shaped by pretrial risk assessments predicting outcomes that align with the concerns of prosecutors and judges. I showed how PSA reports shift the terms of debate, increasing the salience of missing court, one of the outcomes the algorithm predicts. This shift aligned deliberations better with the legal basis for cash bail in Wisconsin. But without accompanying increases in discussions of factors like ability-to-pay that might moderate the use of punitive pretrial conditions, this change created openings to cast arrested people in a more unfavorable light and facilitated harsher decisions.

This article illustrates how predictive algorithms can amplify the concerns that the tools predict. It bridges studies of the sociopolitical factors that shape their design and their use in practice (Passi and Barocas, 2019; Benjamin, 2019; Brayne, 2020; Joyce et al., 2021; Brayne and Christin, 2021) by tracing the downstream practical consequences of an upstream technical choice required to build predictive models: selecting specific outcomes to predict. This choice encodes a particular outlook, demarcating, as many quantification projects do, what ought to matter most (Espeland and Stevens, 1998; Espeland and Vannebo, 2007). In the case of pretrial decision-making, this article reveals how predicting missed court dates empowers actors whose roles center on that concern, redefining the decision-making situation and influencing final decisions. For economic sociologists and quantification scholars who have documented the enduring role of narratives in moral evaluation (Kiviat, 2017; Werth, 2017; Lynch, 2019), this finding further suggests that predictive algorithms can shape *which* narratives become most relevant to tell in settings of limited time and information.

These results also build on exchanges between two traditions of socio-legal scholarship on how actuarialism and quantification shape criminal legal practices (Rothschild-Elyassi, 2022). Consistent with macro-level accounts that actuarial reforms distort the goals of criminal legal institutions (Feeley and Simon, 1992; Harcourt, 2007), this article shows how risk assessment instruments alter decision-making to center the goals they project more strongly. Although thinking about people in terms of risk may already be deeply entrenched in criminal legal institutions, giving actors specific information on a defendant's risk of adverse outcomes, like missing court, can still tilt decisions in more punitive directions. At the same time, this study also aligns with past work on local practices (Hannah-Moffat et al., 2009; Werth, 2019a; Brayne and Christin, 2021), showing that risk assessments leave room for professional discretion and other goals. It suggests, however, that scholars should not only document the co-existence of multiple decision-making logics but also trace the relative balance among them. This working balance can alter decisions, such as by increasing attention to non-appearance, which facilitated the increased use of cash bail in the jurisdiction studied.

This paper additionally contributes to cross-disciplinary literature on how predictive models

shape public policy and public sector institutions (Eubanks, 2018; Johnson and Rostain, 2020; Levy et al., 2021). Answering calls to study predictive tools in organizational context (Selbst et al., 2019; Joyce et al., 2021; Brayne and Christin, 2021), I examine how they add a new voice to multi-actor decision deliberations, reconfiguring interactional dynamics. Rather than viewing predictive algorithms as nudges for individual decision-makers, this paper's approach draws on studies that apply an organizational lens to courtroom communities (McPherson and Sauder, 2013; Ulmer, 2019) and follows a tradition of studying how technologies influence organizational authority relations and interactions when introduced (Barley, 1986; Orlikowski, 2000; Christin, 2020). This emphasis on interactions between decision-making actors with distinct roles, moreover, aligns with recent literature on street-level bureaucracy highlighting the importance of vertical and lateral relations in understanding contemporary poverty governance (Seim, 2017; Barnard, 2023).

Examining these dynamics, this article shows how algorithms can relocate power at the micro level, complementing macro-level analyses of how emerging technologies concentrate power among technical elites (Burrell and Fourcade, 2021). It further joins efforts to demonstrate how technological choices embed substantive policy choices (Passi and Barocas, 2019; Mulligan and Bamberger, 2019; Levy et al., 2021; Johnson and Zhang, 2022). By symbolically asserting the importance of a pre-selected aim and empowering aligned actors, off-the-shelf predictive tools can alter institutions' sense of their mission and the foundations of their policy choices. Within the adversarial system at the foundation of U.S. criminal law, the observed empowerment of prosecutors over defense attorneys by risk assessments raises concerns that such interventions can create an imbalance between protecting the rights of the accused and the interests of the public and the state. As scholars continue to assess whether the interplay between humans and algorithms amplifies or reduces biases in public sector decision-making, this study encourages a closer look at which actors can most effectively harness algorithmic information and how micro-interactional shifts in power influence the pursuit of justice.

Overall, from a policy impact evaluation perspective, this article indicates that simulations of risk assessment algorithms' policy impact based on perfect human compliance (e.g. Kleinberg

et al., 2018) may lead to overly optimistic conclusions. Risk assessment information can be asymmetrically applied, with reports suggesting that a defendant is a high risk being afforded more attention and influence than ones indicating a low risk. Thus, tools like the PSA may not reliably deliver on their promised benefits.

This article additionally offers several methodological contributions. It showcases how randomized controlled trials can be combined with additional administrative or digital trace data to address the critique that experiments can obscure the social processes linking interventions to outcomes of interest (Mize and Manago, 2022). While pinpointing causal mechanisms remains a methodological challenge, harnessing the growing availability of digitized records, especially textual records of institutional encounters and deliberations, can help scholars better trace causal pathways.

For scholars of court processes and judicial decision-making, this article demonstrates how scholars can extract patterns in deliberation and attention from court hearing transcripts or similar records of legal proceedings (see also Dalke, 2023). These methods complement traditional qualitative approaches such as courtroom observation, interviews, and ethnography, which remain crucial for capturing off-record conversations, backstage processes, and embodied interactions that shed further light on the meanings and motivations behind court actors' strategies and decisions. Scholars can moreover apply the computational text analysis methods used here to analogous data in other domains, like care team deliberations in child welfare or health care. As digital records of institutional encounters proliferate — from traffic stops to doctor-patient interactions — here is significant potential for scholars to use them to study how institutions interact with and assess individuals (Voigt et al., 2017).

10.1 Limitations and Directions for Future Research

This research closely investigates a single site to identify processes by which predictive algorithms intervene in pretrial decision-making. U.S. criminal courts, however, vary in their legal and political contexts, caseload characteristics, court cultures, and local procedures, which can shape the impact

of quantification efforts like algorithmic risk assessments (Ulmer and Kramer, 1998; Espeland and Vannebo, 2007; Viljoen et al., 2019; Hood and Schneider, 2019). Two attributes of Dane County are especially important to the generalizability of the asymmetric effects observed. First, Dane County assures public defender representation for defendants at initial appearances. Since access to defense counsel typically improves pretrial outcomes for defendants (Anwar et al., 2023), the asymmetries observed might be more pronounced in jurisdictions without such guarantees, where defendants may be less well-positioned to defend themselves against any added suspicion a risk assessment may introduce.

Second, Dane County is a Democratic stronghold in a swing state. On the one hand, this could make the observed asymmetries more likely in Dane County than in more conservative areas. Judges in the county already granted signature bonds to many defendants at baseline, leaving them less room to be more lenient. On the other hand, we could also see Dane County as a place where the asymmetries are less likely, given that local officials faced political pressure to moderate their use of cash bail from local advocacy groups and formal initiatives to improve pretrial justice locally. Future comparative research examining impacts in other locales should analyze how these and other local factors shape risk assessments' impact.

Specifics of local PSA implementation may also have affected the PSA's impact. While previous studies have indicated that courtroom actors often ignore risk assessments (Brayne and Christin, 2021; Pruss, 2023), this study finds substantial engagement with the tool. One reason the tool secured more buy-in from local actors may have been that local actors were trained on it and involved in the design process, such as in constructing the Decision-Making Framework. The randomized-controlled trial itself may also have influenced local engagement. Withholding half of the PSA reports could have increased the salience of those provided. Awareness that court decisions would be scrutinized for research could also have modified actors' behavior. Further, control hearings do not fully reflect a world without risk assessments. Instead, they reflect a world where actors have already been exposed to how the risk assessment structures pretrial decision-making. Thus, the RCT may increase attention to risk assessments, but it may also attenuate the

tools' observed impact, as control group hearings have already been "treated" with the underlying framework that the PSA propagates.

Several limitations related to the lack of data may be addressed in future research. First, this study examines only felony cases. While the Supplemental Material offers some evidence that the asymmetric impact of cash bail and signature bond recommendations also applies in misdemeanor cases, additional research is needed to examine heterogeneity in the processes I discuss by offense severity. Second, the data cover only one year of implementation. Since the effects of risk assessment algorithms may wane over time (Stevenson, 2018), additional research is needed to evaluate long-term effects and whether practices the tools prompt are institutionalized (Chun and Sauder, 2022). Third, this study's sample size is insufficient to cleanly evaluate differential effects for defendants of different ethnoracial backgrounds or gender identities. While preliminary analyses are in the Supplemental Material, additional data is needed to draw definitive conclusions.

More broadly, this study raises questions on whether goal funneling and selective empowerment observed in one decision-making arena may spill over into other organizational tasks or domains. The most immediate way this might happen is when multiple decisions are made about people at once or in sequence, such as in child welfare, education, and health care. In pretrial, notably, actors make decisions not only about cash bail and detention but also about non-monetary release conditions (e.g., electronic monitoring, mandatory drug testing, no-contact orders) and referrals to social services and diversion programs. If goal funneling also enhances the salience of predicted goals in related decision problems, predictive algorithms may have effects on decisions beyond the focal ones commonly scrutinized in research. Beyond routine decision-making, goal funneling and selective empowerment could further affect larger-scale strategic decisions like budgeting, governance, program support, and the adoption of future reforms. Tracing such downstream consequences may be challenging and will likely require a mix of methodological approaches. However, widening the lens where we look for impact may be necessary to fully capture the ripple effects of deploying predictive algorithms in complex organizations.

This study additionally underscores the need for more research on how predictive algorithms

might lead organizations to take risk-averse actions that place heavier burdens on individuals. This study showed this dynamic in pretrial decision-making, but similar patterns may appear in other domains where some types of inferential mistakes impose greater legitimacy costs. For instance, in tax enforcement and public benefits administration, failing to detect fraud often draws significant public scrutiny. Conversely, when individuals are erroneously accused of wrongdoing, such cases often receive less attention, even though the accused often face substantial consequences like legal scrutiny and denied benefits. Similarly, in child welfare, the reputational repercussions of mistakenly leaving a child in a dangerous situation might be greater than those of many unnecessary investigations or out-of-home placements. In settings with diverse goals and differentially visible errors, risk-based algorithms can heighten scrutiny of cases labeled as high risk as decision-makers seek to avoid the most costly mistakes. At the same time, these tools could also allow decision-makers to take more lenient or less aggressive approaches if algorithms can shield them from blame (Hannah-Moffat, 2013; Werth, 2017; Albright, 2023). Future research should explore the relative strength of these forces across sites.

Zooming out further, it may be that the organizations most inclined to adopt algorithmic decision tools in the first place are those driven by fears of specific errors. And the tools they adopt may mirror established ways of thinking about problems (Barabas et al., 2020), relying on data systems that, in what they do and do not know, highlight particular institutional interests (Scott, 1998; Gitelman, 2013; Brayne, 2020; Levy et al., 2021). Future studies should examine the spread of algorithmic decision aids, deepening our understanding of the institutional actors most influenced by the data imperative (Fourcade and Healy, 2017) and the members of society most exposed to algorithmic decision-making (Eubanks, 2018; Rona-Tas, 2020). Investigating how decision aids are constructed, paying attention to design options that are ruled out or dismissed, can shed further light on whether these tools reinforce existing institutional practices or challenge them (Benjamin, 2019).

Another line of research could explore how predictive tools could be redesigned or reimagined. Scholars could examine whether modifying how these tools frame a decision (Rivera and Tilcsik,

2019; Zottola et al., 2023) alters how people interact with them, such as choosing target outcomes that frame people in terms of successes rather than failures, showcasing statistical uncertainty to open space for individualized deliberation, and anchoring algorithmic scores in real-world prevalences — such as highlighting that most people whom the PSA deems high risk do attend court (DeMichele et al., 2020) — to correct misconceptions. Research could also focus on alternative tools that invert the typical downward gaze of prediction models by predicting outcomes like police misconduct, judicial bias, or landlord negligence (Barabas et al., 2020; Brayne, 2020; Meyer et al., 2022). This could help advance sociological theories of quantification and digital technologies by moving beyond model cases and reveal additional ways data can facilitate institutional decision-making or enhance accountability (Johnson and Rostain, 2020; Abebe et al., 2020).

References

Abbott, Andrew and Angela Tsay. 2000. "Sequence Analysis and Optimal Matching Methods in Sociology: Review and Prospect." *Sociological Methods & Research* 29:3–33.

Abebe, Rediet, Solon Barocas, Jon Kleinberg, Karen Levy, Manish Raghavan, and David G. Robinson. 2020. "Roles for Computing in Social Change." *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* pp. 252–260.

Advancing Pretrial Policy and Research. 2023. "PSA Map."

Ajunwa, Ifeoma. 2023. *The Quantified Worker: Law and Technology in the Modern Workplace*. Cambridge: Cambridge University Press.

Albright, Alex. 2023. "The Hidden Effects of Algorithmic Recommendations." Working Paper.

Angwin, Julia and Jeff Larson. 2016. "Machine Bias." *ProPublica* .

Anthony, Callen, Beth A. Bechky, and Anne-Laure Fayard. 2023. ""Collaborating" with AI: Taking a System View to Explore the Future of Work." *Organization Science* .

Anwar, Shamena, Shawn Bushway, and John Engberg. 2023. "The Impact of Defense Counsel at Bail Hearings." *Science Advances* 9:eade3909.

Ares, Charles E., Anne Rankin, and Herbert Sturz. 1963. "The Manhattan Bail Project: An Interim Report on the Use of Pre-Trial Parole." *New York University Law Review* 38:67–95.

Arnold, David, Will Dobbie, and Crystal S. Yang. 2018. "Racial Bias in Bail Decisions." *The Quarterly Journal of Economics* 133:1885–1932.

Austin, James and Roger Ocker. 2021. "Dane County Jail Race and Ethnicity Disparity Analysis." Technical report.

Austin, Peter C. 2009. "Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples." *Statistics in Medicine* 28:3083–3107.

Barabas, Chelsea, Colin Doyle, JB Rubinovitz, and Karthik Dinakar. 2020. "Studying up: Reorienting the Study of Algorithmic Fairness around Issues of Power." In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pp. 167–176.

Barley, Stephen R. 1986. "Technology as an Occasion for Structuring: Evidence from Observations of CT Scanners and the Social Order of Radiology Departments." *Administrative Science Quarterly* 31:78–108.

Barnard, Alex V. 2023. *Conservatorship: Inside California's System of Coercion and Care for Mental Illness*. Columbia University Press.

Barocas, Solon and Andrew D. Selbst. 2016. "Big Data's Disparate Impact." *California Law Review* 104:671–732.

Benjamin, Ruha. 2019. "Discriminatory Design, Liberating Imagination." In *Captivating Technology: Race, Carceral Technoscience, and Liberatory Imagination in Everyday Life*, edited by Ruha Benjamin. Durham: Duke University Press.

Berrick, Jill Duerr. 2017. *The Impossible Imperative: Navigating the Competing Principles of Child Protection*. New York, NY: Oxford University Press.

Blair-Loy, Mary, Olga V. Mayorova, Pamela C. Cosman, and Stephanie I. Fraley. 2022. "Can Rubrics Combat Gender Bias in Faculty Hiring?" *Science* 377:35–37.

Bonikowski, Bart and Laura K. Nelson. 2022. "From Ends to Means: The Promise of Computational Text Analysis for Theoretically Driven Sociological Research." *Sociological Methods & Research* 51:1469–1483.

Bosk, Emily A. 2020. "Iron Cage or Paper Cage? The Interplay of Worker Characteristics and Organizational Policy in Shaping Unequal Responses to a Standardized Decision-Making Tool." *Social Problems* 67:654–676.

boyd, danah and Kate Crawford. 2012. "Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon." *Information, Communication & Society* 15:662–679.

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

Brayne, Sarah and Angèle Christin. 2021. "Technologies of Crime Prediction: The Reception of Algorithms in Policing and Criminal Courts." *Social Problems* 68:608–624.

Breiman, Leo. 2001. "Statistical Modeling: The Two Cultures." *Statistical Science* 16:199–215.

Burrell, Jenna and Marion Fourcade. 2021. "The Society of Algorithms." *Annual Review of Sociology* 47:213–237.

Charmaz, Kathy. 2006. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. SAGE.

Chouldechova, Alexandra, Diana Benavides-Prado, Oleksandr Fialko, and Rhema Vaithianathan. 2018. "A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions." In *Conference on Fairness, Accountability and Transparency*, pp. 134–148.

Christin, Angèle. 2020. "The Ethnographer and the Algorithm: Beyond the Black Box." *Theory and Society* 49:897–918.

Chu, James. 2021. "Cameras of Merit or Engines of Inequality? College Ranking Systems and the Enrollment of Disadvantaged Students." *American Journal of Sociology* 126:1307–1346.

Chun, Hyunsik and Michael Sauder. 2022. "The Logic of Quantification: Institutionalizing Numerical Thinking." *Theory and Society* 51:335–370.

Copp, Jennifer E., William Casey, Thomas G. Blomberg, and George Pesta. 2022. "Pretrial Risk Assessment Instruments in Practice: The Role of Judicial Discretion in Pretrial Reform." *Criminology & Public Policy* 21:329–358.

County of Dane, WI. 2018. "Frequency of Signature Bonds in Dane County Criminal Cases: 2012-2016." Technical report, Madison, WI.

Dalke, Isaac. 2023. "I Come before You a Changed Man: "Insight," Compliance, and Refurbishing Penal Practice in California." *Law & Social Inquiry* pp. 1–31.

De La O, Ana L. 2013. "Do Conditional Cash Transfers Affect Electoral Behavior? Evidence from a Randomized Experiment in Mexico." *American Journal of Political Science* 57:1–14.

DeLuca, Stefanie, Lawrence F. Katz, and Sarah C. Oppenheimer. 2023. ""When Someone Cares About You, It's Priceless": Reducing Administrative Burdens and Boosting Housing Search Confidence to Increase Opportunity Moves for Voucher Holders." *RSF: The Russell Sage Foundation Journal of the Social Sciences* 9:179–211.

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

DeMichele, Matthew, Peter Baumgartner, Michael Wenger, Kelle Barrick, and Megan Comfort. 2020. "Public safety assessment: Predictive utility and differential prediction by race in Kentucky." *Criminology & Public Policy* 19:409–431.

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

Desmarais, Sarah L and Evan M Lowder. 2019. "Pretrial Risk Assessment Tools: A Primer for Judges, Prosecutors, and Defense Attorneys." Technical report.

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* 48:398–420.

Didwania, Stephanie. 2021. "Discretion and Disparity in Federal Detention." *Northwestern University Law Review* 115:1261–1336.

Dobbie, Will, Jacob Goldin, and Crystal S. Yang. 2018. "The Effects of Pre-Trial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges." *American Economic Review* 108:201–240.

Dobbin, Frank, Daniel Schrage, and Alexandra Kalev. 2015. "Rage against the Iron Cage The Varied Effects of Bureaucratic Personnel Reforms on Diversity." *American Sociological Review* 80:1014–1044.

Eisenstein, James and Herbert Jacob. 1977. *Felony Justice: An Organizational Analysis of Criminal Courts*. Boston: Little, Brown, and Co.

Emerson, Robert M. 1983. "Holistic Effects in Social Control Decision-Making." *Law & Society Review* 17:425–455.

Engler, Alex. 2021. "Enrollment Algorithms Are Contributing to the Crises of Higher Education." Technical report, The Brookings Institution.

Espeland, Wendy and Michael Sauder. 2007. "Rankings and Reactivity: How Public Measures Recreate Social Worlds." *American Journal of Sociology* 113:1–40.

Espeland, Wendy Nelson and Mitchell L. Stevens. 1998. "Commensuration as a Social Process." *Annual Review of Sociology* 24:313–343.

Espeland, Wendy Nelson and Berit Irene Vannebo. 2007. "Accountability, Quantification, and Law." *Annual Review of Law and Social Science* 3:21–43.

Eubanks, Virginia. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. 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.

Feeley, Malcolm M. and Jonathan Simon. 1992. "The New Penology: Notes on the Emerging Strategy of Corrections and Its Implications." *Criminology* 30:449–474.

FICO Decisions. 2018. "Machine Learning and FICO Scores." Technical report.

Fong, Kelley. 2020. "Getting Eyes in the Home: Child Protective Services Investigations and State Surveillance of Family Life." *American Sociological Review* 85:610–638.

Fourcade, Marion and Kieran Healy. 2017. "Seeing like a market." *Socio-Economic Review* 15:9–29.

Gitelman, Lisa. 2013. *Raw Data Is an Oxymoron*. The MIT Press.

Goel, Sharad, Ravi Shroff, Jennifer Skeem, and Christopher Slobogin. 2021. "The accuracy, equity, and jurisprudence of criminal risk assessment." In *Research Handbook on Big Data Law*.

Gottfredson, Stephen D. and Laura J. Moriarty. 2006. "Statistical Risk Assessment: Old Problems and New Applications." *Crime & Delinquency* 52:178–200.

Gouldin, Lauryn. 2018. "Defining Flight Risk." *The University of Chicago Law Review* 85:677–742.

Gouldin, Lauryn P. 2016. "Disentangling Flight Risk from Dangerousness." *Brigham Young University Law Review* p. 63.

Green, Ben and Yiling Chen. 2021. "Algorithmic Risk Assessments Can Alter Human Decision-Making Processes in High-Stakes Government Contexts." In *Proceedings of the ACM on Human-Computer Interaction*, volume 5, pp. 1–33.

Greiner, D James, Ryan Halen, Matthew Stubenberg, and Christopher L. Griffin. 2020. "Randomized Control Trial Evaluation of the Implementation of the PSA-DMF System: Interim Report." Technical report.

Griffin, Christopher L. 2016. "Dane County PSA Study Analysis Plan."

Hannah-Moffat, Kelly. 2013. "Actuarial Sentencing: An "Unsettled" Proposition." *Justice Quarterly* 30:270–296.

Hannah-Moffat, Kelly, Paula Maurutto, and Sarah Turnbull. 2009. "Negotiated Risk: Actuarial Illusions and Discretion in Probation." *Canadian Journal of Law and Society* 24:391–409.

Harcourt, Bernard E. 2007. *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age*. University of Chicago Press.

Heimer, Carol A. 2001. "Cases and Biographies: An Essay on Routinization and the Nature of Comparison." *Annual Review of Sociology* 27:47–76.

Hirschman, Daniel, Ellen Berrey, and Fiona Rose-Greenland. 2016. "Dequantifying diversity: affirmative action and admissions at the University of Michigan." *Theory and Society* 45:265–301.

Hirschman, Daniel and Emily Adlin Bosk. 2020. "Standardizing Biases: Selection Devices and the Quantification of Race." *Sociology of Race and Ethnicity* 6:348–364.

Hood, Katherine and Daniel Schneider. 2019. "Bail and Pretrial Detention: Contours and Causes of Temporal and County Variation." *RSF: The Russell Sage Foundation Journal of the Social Sciences* 5:126–149.

Imai, Kosuke, Zhichao Jiang, D James Greiner, Ryan Halen, and Sooahn Shin. 2023. "Experimental Evaluation of Algorithm-Assisted Human Decision-Making: Application to Pretrial Public Safety Assessment*." *Journal of the Royal Statistical Society Series A: Statistics in Society* 186:167–189.

Jacobs, Abigail Z. and Hanna Wallach. 2021. "Measurement and Fairness." In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pp. 375–385.

Johnson, Rebecca A. and Tanina Rostain. 2020. "Tool for Surveillance or Spotlight on Inequality? Big Data and the Law." *Annual Review of Law and Social Science* 16:453–472.

Johnson, Rebecca Ann and Simone Zhang. 2022. "What Is the Bureaucratic Counterfactual? Categorical versus Algorithmic Prioritization in U.S. Social Policy." In *2022 ACM Conference on Fairness, Accountability, and Transparency*, pp. 1671–1682.

Jorgensen, Isabella and Sandra Smith. 2021. "The Current State of Bail Reform in the United States: Results of a Landscape Analysis of Bail Reforms across All 50 States." Working Paper RWP21-033, Harvard Kennedy School.

Joyce, Kelly, Laurel Smith-Doerr, Sharla Alegria, Susan Bell, Taylor Cruz, Steve G. Hoffman, Safiya Umoja Noble, and Benjamin Shestakofsky. 2021. "Toward a Sociology of Artificial Intelligence: A Call for Research on Inequalities and Structural Change." *Socius* 7:1–11.

Jung, Jongbin, Connor Concannon, Ravi Shroff, Sharad Goel, and Daniel G. Goldstein. 2020. "Simple Rules to Guide Expert Classifications." *Journal of the Royal Statistical Society Series A: Statistics in Society* 183:771–800.

Kalev, Alexandra. 2014. "How You Downsize Is Who You Downsize: Biased Formalization, Accountability, and Managerial Diversity." *American Sociological Review* 79:109–135.

Kiviat, Barbara. 2017. "The art of deciding with data: evidence from how employers translate credit reports into hiring decisions." *Socio-Economic Review* .

Kiviat, Barbara. 2019. "The Moral Limits of Predictive Practices: The Case of Credit-Based Insurance Scores." *American Sociological Review* 84:1134–1158.

Kiviat, Barbara. 2023. "The Moral Affordances of Construing People as Cases: How Algorithms

and the Data They Depend on Obscure Narrative and Noncomparative Justice." *Sociological Theory* 41:175–200.

Kleinberg, Jon, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan. 2018. "Human Decisions and Machine Predictions." *The Quarterly Journal of Economics* 133:237–293.

Kleinberg, Jon, Jens Ludwig, Sendhil Mullainathan, and Ziad Obermeyer. 2015. "Prediction Policy Problems." *American Economic Review* 105:491–495.

Laufer, Benjamin, Thomas Krendl Gilbert, and Helen Nissenbaum. 2023. "Optimization's Neglected Normative Commitments." In *2023 ACM Conference on Fairness, Accountability, and Transparency*, pp. 50–63.

Lavigne, Michele, Bonnie Hoffman, and Renee Spence. 2018. "The Wisconsin Bail Manual." Technical report, National Association of Criminal Defense Lawyers, Washington, DC.

Lebovitz, Sarah, Hila Lifshitz-Assaf, and Natalia Levina. 2022. "To Engage or Not to Engage with AI for Critical Judgments: How Professionals Deal with Opacity When Using AI for Medical Diagnosis." *Organization Science* 33:126–148.

Levy, Karen, Kyla E. Chasalow, and Sarah Riley. 2021. "Algorithms and Decision-Making in the Public Sector." *Annual Review of Law and Social Science* 17:309–334.

Lipsky, Michael. 2010. *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services*. New York: Russell Sage Foundation, 30th anniversary edition.

Lowder, Evan M., Carmen L. Diaz, Eric Grommon, and Bradley R. Ray. 2021. "Effects of Pretrial Risk Assessments on Release Decisions and Misconduct Outcomes Relative to Practice as Usual." *Journal of Criminal Justice* 73:101754.

Lynch, Mona. 2019. "The Narrative of the Number: Quantification in Criminal Court." *Law & Social Inquiry* 44:31–57.

March, James and Herbert Simon. 1958. *Organizations*. John Wiley and Sons.

Maurutto, Paula and Kelly Hannah-Moffat. 2006. "Assembling Risk and the Restructuring of Penal Control." *The British Journal of Criminology* 46:438–454.

McDonough, Anne, Ted Enamorado, and Tali Mendelberg. 2022. "Jailed While Presumed Innocent: The Demobilizing Effects of Pretrial Incarceration." *The Journal of Politics* 84:1777–1790.

McPherson, Chad Michael and Michael Sauder. 2013. "Logics in Action: Managing Institutional Complexity in a Drug Court." *Administrative Science Quarterly* 58:165–196.

Meyer, Mikaela, Aaron Horowitz, Erica Marshall, and Kristian Lum. 2022. "Flipping the Script on Criminal Justice Risk Assessment: An Actuarial Model for Assessing the Risk the Federal Sentencing System Poses to Defendants." In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, pp. 366–378.

Mize, Trenton D. and Bianca Manago. 2022. "The Past, Present, and Future of Experimental Methods in the Social Sciences." *Social Science Research* 108:102799.

Mulligan, Deirdre K. and Kenneth A. Bamberger. 2019. "Procurement As Policy: Administrative Process for Machine Learning." *Berkeley Technology Law Journal* 37:773–851.

NCSL. 2023. "Pretrial Release: Financial Conditions of Release." Technical report, National Conference of State Legislatures, Denver, CO.

Obermeyer, Ziad, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. 2019. "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations." *Science* 366:447–453.

O'Malley, Pat. 1992. "Risk, Power and Crime Prevention." *Economy and Society* 21:252–275.

Orlikowski, Wanda J. 2000. "Using Technology and Constituting Structures: A Practice Lens for Studying Technology in Organizations." *Organization Science* 11:404–428.

Ottone, Sarah and Christine S. Scott-Hayward. 2018. "Pretrial Detention and the Decision to Impose Bail in Southern California." *Criminology, Criminal Justice, Law and Society* 19:24–43.

Ouss, Aurélie and Megan Stevenson. 2023. "Does Cash Bail Deter Misconduct?" *American Economic Journal: Applied Economics* 15:150–182.

Page, Joshua and Christine S. Scott-Hayward. 2022. "Bail and Pretrial Justice in the United States: A Field of Possibility." *Annual Review of Criminology* 5:91–113.

Pager, Devah and Hana Shepherd. 2008. "The Sociology of Discrimination: Racial Discrimination in Employment, Housing, Credit, and Consumer Markets." *Annual Review of Sociology* 34:181–209.

Passi, Samir and Solon Barocas. 2019. "Problem Formulation and Fairness." In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, pp. 39–48.

Pennebaker, James W., Ryan L. Boyd, Kayla Jordan, and Kate Blackburn. 2015. "The Development and Psychometric Properties of LIWC2015." .

Picard-Fritsche, Sarah, Michael Rempel, Jennifer A Tallon, Julian Adler, and Natalie Reyes. 2017. "Demystifying Risk Assessment." Technical report, Center for Court Innovation.

PJI. 2017. "The State of Pretrial Justice in America." Technical report, Pretrial Justice Institute, Baltimore, Maryland.

Porter, Theodore M. 1995. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton: Princeton University Press.

Pruss, Dasha. 2023. "Ghosting the Machine: Judicial Resistance to a Recidivism Risk Assessment Instrument." In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, pp. 312–323.

Rivera, Lauren A. and András Tilcsik. 2019. "Scaling Down Inequality: Rating Scales, Gender Bias, and the Architecture of Evaluation." *American Sociological Review* 84:248–274.

Robinson, David G and Logan Koepke. 2019. "Civil Rights and Pretrial Risk Assessment Instruments." Technical report, Upturn.

Rona-Tas, Akos. 2020. "Predicting the Future: Art and Algorithms." *Socio-Economic Review* 18:893–911.

Rosen, Eva, Philip M. E. Garboden, and Jennifer E. Cossyleon. 2021. "Racial Discrimination in Housing: How Landlords Use Algorithms and Home Visits to Screen Tenants." *American Sociological Review* 86:787–822.

Rosenblatt, Peter and Stefanie DeLuca. 2012. ""We Don't Live Outside, We Live in Here": Neighborhood and Residential Mobility Decisions Among Low-Income Families†." *City & Community* 11:254–284.

Rothschild-Elyassi, Gil. 2022. "The Datafication of Law: How Technology Encodes Carceral Power and Affects Judicial Practice in the United States." *Law & Social Inquiry* 47:55–94.

Salganik, Matthew J., Ian Lundberg, Alexander T. Kindel, Caitlin E. Ahearn, Khaled Al-Ghoneim, Abdullah Almaatouq, Drew M. Altschul, Jennie E. Brand, Nicole Bohme Carnegie, Ryan James Compton, Debanjan Datta, Thomas Davidson, Anna Filippova, Connor Gilroy, Brian J. Goode, Eaman Jahani, Ridhi Kashyap, Antje Kirchner, Stephen McKay, Allison C. Morgan, Alex Pentland, Kivan Polimis, Louis Raes, Daniel E. Rigobon, Claudia V. Roberts, Diana M. Stanescu, Yoshihiko Suhara, Adaner Usmani, Erik H. Wang, Muna Adem, Abdulla Alhajri, Bedoor AlShebli, Redwane Amin, Ryan B. Amos, Lisa P. Argyle, Livia Baer-Bositis, Moritz Büchi, Bo-Ryehn Chung, William Eggert, Gregory Faletto, Zhilin Fan, Jeremy Freese, Tejomay Gadgil, Josh Gagné, Yue Gao, Andrew Halpern-Manners, Sonia P. Hashim, Sonia Hausen, Guanhua He, Kimberly Higuera, Bernie Hogan, Ilana M. Horwitz, Lisa M. Hummel, Naman Jain, Kun Jin, David Jurgens, Patrick Kaminski, Areg Karapetyan, E. H. Kim, Ben Leizman, Naijia Liu, Malte Möser, Andrew E. Mack, Mayank Mahajan, Noah Mandell, Helge Marahrens, Diana Mercado-Garcia, Viola Mocz, Katariina Mueller-Gastell, Ahmed Musse, Qiankun Niu, William Nowak,

Hamidreza Omidvar, Andrew Or, Karen Ouyang, Katy M. Pinto, Ethan Porter, Kristin E. Porter, Crystal Qian, Tamkinat Rauf, Anahit Sargsyan, Thomas Schaffner, Landon Schnabel, Bryan Schonfeld, Ben Sender, Jonathan D. Tang, Emma Tsurkov, Austin van Loon, Onur Varol, Xiafei Wang, Zhi Wang, Julia Wang, Flora Wang, Samantha Weissman, Kirstie Whitaker, Maria K. Wolters, Wei Lee Woon, James Wu, Catherine Wu, Kengran Yang, Jingwen Yin, Bingyu Zhao, Chenyun Zhu, Jeanne Brooks-Gunn, Barbara E. Engelhardt, Moritz Hardt, Dean Knox, Karen Levy, Arvind Narayanan, Brandon M. Stewart, Duncan J. Watts, and Sara McLanahan. 2020. "Measuring the Predictability of Life Outcomes with a Scientific Mass Collaboration." *Proceedings of the National Academy of Sciences* 117:8398–8403.

Sarver, Christian, Kort Prince, Jessica Seawright, and Robert Butters. 2015. "A Review of Brief Risk Assessment Tools Validated for Use in Correctional Settings." Technical report, Utah Criminal Justice Center.

Saxena, Devansh, Karla Badillo-Urquiola, Pamela J. Wisniewski, and Shion Guha. 2020. "A Human-Centered Review of Algorithms Used within the U.S. Child Welfare System." In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–15, Honolulu HI USA. ACM.

Scott, James C. 1998. *Seeing like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven, CT London: Yale University Press.

Seim, Josh. 2017. "The Ambulance: Toward a Labor Theory of Poverty Governance." *American Sociological Review* 82:451–475.

Selbst, Andrew D., Danah Boyd, Sorelle A. Friedler, Suresh Venkatasubramanian, and Janet Vertesi. 2019. "Fairness and Abstraction in Sociotechnical Systems." In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, FAT* '19, pp. 59–68, New York, NY, USA. ACM. event-place: Atlanta, GA, USA.

Simmel, Georg. 1902. "The Number of Members as Determining the Sociological Form of the Group." *American Journal of Sociology* 8:1–46.

Simon, Jonathan. 1988. "The Ideological Effects of Actuarial Practices." *Law & Society Review* 22:30.

Skeem, Jennifer, Nicholas Scurich, and John Monahan. 2020. "Impact of Risk Assessment on Judges' Fairness in Sentencing Relatively Poor Defendants." *Law and Human Behavior* 44:51–59.

Skeem, Jennifer L. and John Monahan. 2011. "Current Directions in Violence Risk Assessment." *Current Directions in Psychological Science* 20:38–42.

Smith, Alison M. 2019. "U.S. Constitutional Limits on State Money-Bail Practices for Criminal Defendants." Technical Report R45533, Congressional Research Service, Washington, D.C.

Soland, James, Benjamin Domingue, and David Lang. 2020. "Using Machine Learning to Advance Early Warning Systems: Promise and Pitfalls." *Teachers College Record: The Voice of Scholarship in Education* 122:1–30.

Starr, Sonja B. 2014. "Evidence-Based Sentencing and the Scientific Rationalization of Discrimination." *Stanford Law Review* 66:803–72.

Stevenson, Megan T. 2018. "Assessing Risk Assessment in Action." *Minnesota Law Review* 103.

Stevenson, Megan T and Sandra G Mayson. 2021. "Pretrial Detention and the Value of Liberty." p. 61.

Stone, Harold S. 1972. *Introduction to Computer Organization and Data Structures*. McGraw-Hill.

Sugie, Naomi F. and Kristin Turney. 2017. "Beyond Incarceration: Criminal Justice Contact and Mental Health." *American Sociological Review* 82:719–743.

Tausczik, Yla R. and James W. Pennebaker. 2010. "The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods." *Journal of Language and Social Psychology* 29:24–54.

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* 47:927–942.

Traeger, Margaret L., Sarah Strohkorb Sebo, Malte Jung, Brian Scassellati, and Nicholas A. Christakis. 2020. "Vulnerable Robots Positively Shape Human Conversational Dynamics in a Human–Robot Team." *Proceedings of the National Academy of Sciences* 117:6370–6375.

Uhlmann, Eric Luis and Geoffrey L. Cohen. 2005. "Constructed Criteria: Redefining Merit to Justify Discrimination." *Psychological Science* 16:474–480.

Ulmer, Jeffery T. 2019. "Criminal Courts as Inhabited Institutions: Making Sense of Difference and Similarity in Sentencing." *Crime and Justice* 48:483–522.

Ulmer, Jeffery T. and John H. Kramer. 1998. "The Use and Transformation of Formal Decision-Making Criteria: Sentencing Guidelines, Organizational Contexts, and Case Processing Strategies*." *Social Problems* 45:248–267.

Urbina, Daniela R. 2020. "In the Hands of Women: Conditional Cash Transfers and Household Dynamics." *Journal of Marriage and Family* 82:1571–1586.

US Census Bureau. 2019. "Quick Facts: Dane County, Wisconsin." Technical report.

Vaithianathan, R and C.I. Kithulgoda. 2020. "Using Predictive Risk Modeling to Prioritize Services for People Experiencing Homelessness in Allegheny County: Methodology Paper for the Allegheny Housing Assessment." Technical report, Allegheny County Analytics.

Van Cleve, Nicole Gonzalez. 2016. *Crook County: Racism and Injustice in America's Largest Criminal Court*. Stanford University Press.

Van Cleve, Nicole Gonzalez and Lauren Mayes. 2015. "Criminal Justice Through "Colorblind" Lenses: A Call to Examine the Mutual Constitution of Race and Criminal Justice." *Law & Social Inquiry* 40:406–432.

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* 43:397–420.

Voigt, Rob, Nicholas P. Camp, Vinodkumar Prabhakaran, William L. Hamilton, Rebecca C. Hetey, Camilla M. Griffiths, David Jurgens, Dan Jurafsky, and Jennifer L. Eberhardt. 2017. "Language from Police Body Camera Footage Shows Racial Disparities in Officer Respect." *Proceedings of the National Academy of Sciences* 114:6521–6526.

Wakefield, Sara and Lars Andersen. 2020. "Pretrial Detention and the Costs of System Overreach for Employment and Family Life." *Sociological Science* 7:342–366.

Wang, Angelina, Sayash Kapoor, Solon Barocas, and Arvind Narayanan. 2023. "Against Predictive Optimization: On the Legitimacy of Decision-Making Algorithms That Optimize Predictive Accuracy." In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, p. 626.

Weber, Max. 1978. *Economy and Society: An Outline of Interpretive Sociology*. University of California Press.

Werth, Robert. 2017. "Individualizing Risk: Moral Judgement, Professional Knowledge and Affect in Parole Evaluations." *The British Journal of Criminology* 57:808–827.

Werth, Robert. 2019a. "Risk and Punishment: The Recent History and Uncertain Future of Actuarial, Algorithmic, and "Evidence-Based" Penal Techniques." *Sociology Compass* 13:e12659.

Werth, Robert. 2019b. "Theorizing the Performative Effects of Penal Risk Technologies: (Re)producing the Subject Who Must Be Dangerous." *Social & Legal Studies* 28:327–348.

Wilson, James. 1989. *Bureaucracy: What Government Agencies Do And Why They Do It*. New York, NY: Basic Books.

Winter, Alix S. and Matthew Clair. 2023. ""The Roughest Form of Social Work:" How Court Officials Justify Bail Decisions." *Criminology* .

Yang, Dongsheng and Jarrod E Dalton. 2012. "A unified approach to measuring the effect size between two groups using SAS®." *SAS Global Forum* p. 6.

Zacka, Bernardo. 2017. *When the State Meets the Street: Public Service and Moral Agency*. Cambridge, Massachusetts: The Belknap Press of Harvard University Press.

Zottola, Samantha A., Sarah L. Desmarais, D. Kamiya Stewart, Sarah E. Duhart Clarke, and John Monahan. 2023. "Pretrial Risk Assessment, Release Recommendations, and Racial Bias." *Criminal Justice and Behavior* 50:1255–1278.

A Appendix

A.1 Appendix Figures and Tables

Figure A1: **PSA Inputs, Weights, and Scaling** The top panel shows the weights assigned to each input, while the bottom panel illustrates how raw scores are converted into scaled scores.

PUBLIC SAFETY ASSESSMENT RISK FACTORS

RISK FACTOR	WEIGHTS
FAILURE TO APPEAR maximum total weight = **7 points**	
Pending charge at the time of the offense	No = 0 Yes = 1
Prior conviction	No = 0 Yes = 1
Prior failure to appear pretrial in past 2 years	0 = 0 1 = 2 2 or more = 4
Prior failure to appear pretrial older than 2 years	No = 0 Yes = 1
NEW CRIMINAL ACTIVITY maximum total weight = **13 points**	
Age at current arrest	23 or older = 0 22 or younger = 2
Pending charge at the time of the offense	No = 0 Yes = 3
Prior misdemeanor conviction	No = 0 Yes = 1
Prior felony conviction	No = 0 Yes = 1
Prior violent conviction	0 = 0 1 or 2 = 1 3 or more = 2
Prior failure to appear pretrial in past 2 years	0 = 0 1 = 1 2 or more = 2
Prior sentence to incarceration	No = 0 Yes = 2
NEW VIOLENT CRIMINAL ACTIVITY maximum total weight = **7 points**	
Current violent offense	No = 0 Yes = 2
Current violent offense & 20 years old or younger	No = 0 Yes = 1
Pending charge at the time of the offense	No = 0 Yes = 1
Prior conviction	No = 0 Yes = 1
Prior violent conviction	0 = 0 1 or 2 = 1 3 or more = 2

Source: Laura and John Arnold Foundation

JOURNAL

(a) Inputs and weights

FTA Raw Score	FTA 6 Point Scale	NCA Raw Score	NCA 6 Point Scale	NVCA Raw Score	NVCA Flag
0	1	0	1	0	No
1	2	1	2	1	No
2	3	2	2	2	No
3	4	3	3	3	No
4	4	4	3	4	Yes
5	5	5	4	5	Yes
6	5	6	4	6	Yes
7	6	7	5	7	Yes
		8	5		
		9-13	6		

(b) How raw scores are converted into scaled scores

Table A1: **Administrative Data Sample Characteristics / Covariate Balance Table** Standardized mean differences are calculated following Austin (2009) for binary and continuous variables and Yang and Dalton (2012) for categorical variables.

	Control	**Treatment**	**Std. mean difference**
N	684	695	
Defendant race (%)			0.068
Black	297 (43.4)	302 (43.5)	
White	362 (52.9)	363 (52.3)	
American Indian or Alaskan Native	5 (0.7)	5 (0.7)	
Asian/Pacific Islander	8 (1.2)	11 (1.6)	
Hispanic	8 (1.2)	11 (1.6)	
Unknown	4 (0.6)	2 (0.3)	
Defendant sex = Male (%)	547 (80.0)	564 (81.3)	0.033
Defendant age (mean (SD))	33.64 (11.84)	33.46 (11.59)	0.016
Judge (%)			0.061
Judge A	296 (43.3)	314 (45.2)	
Judge B	310 (45.3)	296 (42.6)	
Judge C	69 (10.1)	77 (11.1)	
Other Judges	9 (1.3)	8 (1.2)	
Severity of highest felony charge (%)			0.064
A-C	55 (8.3)	49 (7.3)	
D-E	37 (5.6)	31 (4.6)	
F-G	145 (21.8)	145 (21.5)	
H-I	427 (64.3)	449 (66.6)	
# Cases (mean (SD))	1.17 (0.71)	1.19 (0.54)	0.038
# Unique charges (mean (SD))	2.49 (1.41)	2.56 (1.44)	0.052
# Violent charges (mean (SD))	0.56 (0.96)	0.50 (0.87)	0.069
# Bail jumping charges (mean (SD))	0.53 (1.18)	0.50 (1.24)	0.030
Failure to Appear Score (mean (SD))	2.92 (1.41)	3.07 (1.37)	0.109
New Criminal Activity Score (mean (SD))	3.54 (1.48)	3.69 (1.39)	0.108
New Violent Criminal (%)	133 (19.4)	132 (19.0)	0.011
Activity = True			
Recommendation (%)			0.075
Signature bond	378 (55.3)	381 (54.8)	
Low cash	76 (11.1)	70 (10.1)	
Moderate cash	24 (3.5)	34 (4.9)	
High cash	206 (30.1)	210 (30.2)	

A.2 Local decision rules for PSA recommendations

Dane County instituted two local decision rules that govern the tool's recommendations for pretrial conditions beyond the standardized Decision-Making Framework rubric:

1. Rules that automatically lead to recommendations for high cash bail and intensive pretrial supervision. These rules are triggered in cases in which the arrested individual was i) extradited from another state, ii) charged with a serious offense (escape, murder/homicide, felony first or second degree sexual assault, or armed robbery), or iii) charged with a violent offense and evaluated by the PSA as likely to commit a violent offense if released.

2. Rules that automatically increase the restrictiveness of the rubric's recommendation by one level. Charges involving domestic abuse, stalking, robbery, violations of restraining orders, kidnapping, arson, or use of a firearm lead to an automatic one-level increase in the restrictiveness of the recommendation. For instance, if the rubric recommends low cash bail in a case that involves one of the those charges, the decision rule would lead to a final recommendation of moderate cash bail.